On Tail Equivalence to the Supremum: From Lévy to Compound Renewal Processes

Nikolaos Karapiperakis

School of Mathematical Sciences

Lancaster University

A thesis submitted for the degree of

Doctor of Philosophy

May, 2026

Abstract

The supremum of a stochastic process is an important object in applied probability, with applications in insurance, finance, queueing theory, and risk analysis. Although its distribution is typically intractable, in many settings the tail of the supremum is asymptotically equivalent to that of the underlying process, a phenomenon we call *tail equivalence to the supremum*. In this thesis we study certain cases where such equivalence holds.

By highlighting a certain property, which we term *synergistic at extremes*, we show that processes attaining extreme values through the combined effect of all increments admit simpler and more intuitive arguments. This perspective allows us to give alternative proofs of existing results on tail equivalence for Lévy processes and to treat special cases such as the perturbed ruin process.

The main contribution of the thesis is to extend the theory beyond the Lévy setting to compound renewal processes, removing the assumption of exponential waiting times. The key step is to establish extreme lower-tail asymptotics for sums of i.i.d. random variables, using exponential tilting and Edgeworth-type expansions. This leads to tail equivalence results for a broad class of compound renewal processes with light-tailed jumps.

Acknowledgements

I would like to sincerely thank my supervisor, Professor Dmitry Korshunov, for his guidance and support. I truly appreciate it.

I am grateful to my parents, Maria and Kostas, without whom I would never have reached the stage of becoming a PhD student.

I also thank my girlfriend Anna for her constant support through all the years of my studies, and Nick for being at the next desk and making things lighter.

Finally, I want to thank my friend Mike for bringing to my attention a research question for the independent project in the appendix and for the discussions we had.

Declaration

I declare that the work presented in this thesis is, to the best of my knowledge and belief, original and my own work, unless otherwise stated, cited, or commonly known. The material has not been submitted, either in whole or in part, for a degree at this, or any other university. This thesis does not exceed the maximum permitted word length of 80,000 words including appendices and footnotes, but excluding the bibliography. A rough estimate of the word count is: 25073

Nikolaos Karapiperakis

Contents

Chapter 1

Introduction

The supremum of stochastic processes is an important object in applied probability. In a wide range of applications — insurance, finance, queueing systems, and risk theory — the largest value attained by a process over time determines the key performance or risk measure of interest. In finance, for instance, barrier and lookback options are priced by functionals of the running maximum, while in queueing models the waiting time distribution depends directly on the supremum of an associated random walk.

The stochastic processes typically used in these applications are often Lévy processes, such as compound Poisson processes, or more generally compound renewal processes. Understanding the distributional properties of the supremum of such processes, and in particular its tail behaviour, is therefore central to many applications. However, obtaining closed-form expressions for the distribution of the supremum is typically very difficult.

Under suitable assumptions, an important phenomenon observed across several classes of processes is that the supremum often has the *same tail order* as the underlying process itself. More precisely, for large thresholds x, the probability that the supremum exceeds x is asymptotically equivalent to the probability that the process exceeds x at a fixed time. Formally,

$$\mathbb{P}\big(\overline{X}_t > x\big) \sim c\,\mathbb{P}(X_t > x)\,, \qquad x \to \infty,$$

where $\overline{X}_t = \sup_{0\leq s\leq t} X_s$ and $c > 0$ is a constant. Here, for positive functions f and g, the notation

$$f(x) \sim g(x), \qquad x \to \infty,$$

means that

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1.$$

Such a relation is extremely useful. It allows one to study the distribution of the supremum indirectly through the distribution of the process, thereby providing asymptotic information in situations where the exact distribution of the supremum is analytically intractable.

This relation, which we refer to as *tail equivalence to the supremum*, is the main object of this thesis. In particular, we ask under what assumptions the above equivalence holds. The assumptions explored concern both the class of the process and the behaviour of its tails.

The study of this kind of phenomenon goes back to Lévy's classical result [35], namely the reflection principle:

$$\mathbb{P}\big(\overline{B}_t > x\big) \;=\; 2\,\mathbb{P}(B_t > x)\,, \qquad \text{for every } t > 0,$$

where $(B_t)_{t\geq 0}$ denotes standard Brownian motion. The above is a strict equality, in contrast to the asymptotic results that followed later.

Subsequently, many authors investigated the topic by focusing on specific classes of Lévy processes with certain tail behaviour. We present these results in detail later in Section 3.2. Most notably, Willekens [44] treated Lévy processes with long tails, proving tail equivalence to the supremum for general Lévy processes. Braverman, who systematically studied the light-tail case in a series of papers [16, 12, 15, 14], established tail equivalence to the supremum for compound Poisson processes with light-tailed[1] jumps. In addition, Albin [1] treated the case where the tails of the process are exponential.

We observe that some existing proofs rely on elements that are not strictly necessary, while at the same time a certain condition/property emerges as key to the phenomenon in the light-tailed case. We call processes that possess this property *synergistic at extremes* and base a particular elementary method for proving asymptotic equivalence to the supremum on this property.

In particular, we show that this property implies (and is equivalent, in the Lévy

[1]Note that his definition of light tails differs from the standard one. See Definition 2.2.3.

case) that a process attains extreme values by using all of its increments. This is the analogue for processes of the principle known for sums of light-tailed random variables as the *conspiracy* or *synergy principle*. Using the fact that such processes achieve extreme values in this specific way, we obtain simple and elementary proofs of previous results and treat some special cases such as the generalised ruin process.

Most importantly, we point out that the property of independent and stationary increments is stronger than necessary: regeneration at hitting times already suffices. This allows us to extend the results beyond the Lévy setting to compound renewal processes with light-tailed jumps.

In particular, the main contributions of this thesis are as follows:

(a) We establish tail equivalence to the supremum for a broad class of compound renewal processes with light-tailed jumps and non-negative linear drift (Theorem 4.1.9). This provides a direct extension of Braverman's result for compound Poisson processes [14] and, to the best of our knowledge, constitutes the first such result outside the Lévy class.

(b) We derive sharp extreme lower-tail asymptotics for the renewal time T_n as $n \to \infty$, for each fixed $s > 0$, obtaining an explicit asymptotic expansion for $\mathbb{P}(T_n \leq s)$ that includes both the precise prefactor and the first subexponential correction term (Proposition 4.3.1).

(c) We prove tail equivalence to the supremum for the perturbed ruin process, that is, for the sum of a compound Poisson process, a linear drift, and an independent Brownian component. To the best of our knowledge, this result does not appear in the existing literature (Proposition 3.4.22).

(d) We introduce a simple method that yields elementary proofs of tail equivalence to the supremum statements (see Section 3.4.6). This approach enables the extension to the non-Lévy setting described in (a) and also provides a unified and conceptually transparent derivation of several classical results, including tail equivalence for Brownian motion with drift, Lévy processes with exponential-type tails, and compound Poisson processes with light-tailed jumps. The resulting proofs are shorter and more transparent than the original arguments and clarify the structural reason why tail equivalence holds.

We begin in Chapter 2 with the basic theory relevant to this thesis. This includes the asymptotic notation used throughout, an introduction to Lévy processes and compound renewal processes, and a review of definitions and notions related to the tails of distributions.

In Chapter 3 we turn to the problem of asymptotic tail equivalence for Lévy processes. We begin by reviewing existing results. We then introduce a key structural property and analyse its consequences, which form the basis of our method for proving tail equivalence to the supremum. We call processes satisfying this property *synergistic at extremes*. Building on this framework, we provide alternative proofs for several important special cases—namely Brownian motion with drift, compound Poisson processes with light-tailed jumps in the sense of Braverman, and Lévy processes with exponential-type tails—and, in addition, establish an original result for the perturbed ruin process. Finally, we summarise the method and outline possible directions for further research.

Chapter 4 contains the main contributions of this thesis, where we turn to compound renewal processes. We remove the assumption of exponential waiting times and develop tools analogous to those introduced for Lévy processes synergistic at extremes, adapted here to processes that regenerate at hitting times. To show that such processes are synergistic at extremes, and hence to apply our method, we must establish sharp extreme lower-tail asymptotics; this is the subject of Section 4.3. This is achieved by exponentially tilting the waiting times and analysing the resulting exponential integral. For this final step, we use an Edgeworth expansion for triangular arrays of random variables, developed in detail in Section A.1.

Finally, in Appendix B we present a short independent research project that is nevertheless relevant to our thesis. We develop logarithmic asymptotics for the tails of the area under a powered random walk having jumps with semi-exponential tails during its first excursion (Theorem B.3.1). The topic is related to what we have done so far, but the techniques are very different, as we use a Large Deviation Principle approach influenced by [5].

Chapter 2

Preliminaries

2.1 Asymptotics

2.1.1 Asymptotic notation: O, o and $\sim$

We recall some standard asymptotic notation that will be used throughout the thesis; see [36].

Definition 2.1.1 (Big-O)**.** *Let $f(z)$ and $g(z)$ be two complex-valued functions defined in a set D of the complex plane.*

(i) ***(At a point).*** *Suppose the closure of D contains a point z_0. We write*

$$f(z) = O(g(z)) \qquad \text{as } z \to z_0 \text{ from } D$$

if there is a number $\delta > 0$ such that

$$|f(z)| \leq C|g(z)|, \qquad z \in D \text{ with } 0 < |z - z_0| < \delta,$$

for some constant $C > 0$. That is, f is bounded in magnitude by a fixed multiple of g for all $z \in D$ that lie sufficiently close to z_0.

(ii) ***(At infinity).*** *Suppose $f(z)$ and $g(z)$ are defined in an unbounded set D of the complex plane. We write*

$$f(z) = O(g(z)) \qquad \text{as } z \to \infty \text{ from } D$$

if there is a number $M > 0$ such that

$$|f(z)| \leq C|g(z)|, \qquad z \in D \text{ with } |z| > M,$$

for some constant $C > 0$. That is, f is bounded in magnitude by a fixed multiple of g for all $z \in D$ that are sufficiently large.

Definition 2.1.2 (Little-o)**.** *Let $f(z)$ and $g(z)$ be two complex-valued functions defined in a set D of the complex plane.*

(i) ***(At a point).*** *Suppose the closure of D contains a point z_0. We write*

$$f(z) = o(g(z)) \qquad \text{as } z \to z_0 \text{ from } D$$

if for any $\varepsilon > 0$ there exists $\delta(\varepsilon) > 0$ such that

$$|f(z)| \leq \varepsilon|g(z)| \qquad \text{whenever } z \in D \text{ and } 0 < |z - z_0| < \delta(\varepsilon).$$

That is, f is smaller in magnitude than any multiple of g for $z \in D$ sufficiently close to z_0.

(ii) ***(At infinity).*** *Suppose $f(z)$ and $g(z)$ are defined in an unbounded set D. We write*

$$f(z) = o(g(z)) \qquad \text{as } z \to \infty \text{ from } D$$

if for any $\varepsilon > 0$ there exists $M(\varepsilon) > 0$ such that

$$|f(z)| \leq \varepsilon|g(z)| \qquad \text{whenever } z \in D \text{ and } |z| > M(\varepsilon).$$

That is, f is smaller in magnitude than any multiple of g for $z \in D$ sufficiently large.

In most cases, we do not need to explicitly refer to the set D, as it will be understood from the context.

Remark 2.1.3. *In practice, the following equivalent characterization of the little-o notation is often more useful and follows directly from the definition (e.g., by repeatedly applying the definition choosing $\varepsilon = 1/n$ in (i)).*

Suppose there exists $\mu > 0$ *such that* $g(z) \neq 0$ *whenever* $0 < |z - z_0| < \mu$ *and* $z \in D$. *Then*

$$f(z) = o(g(z)) \quad \text{as } z \to z_0 \text{ from } D$$

if and only if

$$\lim_{\substack{z \to z_0 \\ z \in D}} \frac{f(z)}{g(z)} = 0.$$

Note that the function g *is allowed to vanish at* z_0*, but not at other nearby points.*

In particular, we will often use the notation $o(1)$ and $O(1)$. We write $f(x) = o(1)$ as $x \to x_0$ if $f(x) \to 0$ as $x \to x_0$, and $f(x) = O(1)$ as $x \to x_0$ if $f(x)$ remains bounded for x sufficiently close to x_0 (or for x sufficiently large when $x_0 = \infty$).

Intuitively, the notation $f(x) = O(g(x))$ means that the magnitude of f is eventually at most of the same order as that of g, while $f(x) = o(g(x))$ means that f is eventually negligible compared to g. Here, "eventually" means for x sufficiently close to x_0 when considering the limit $x \to x_0$, or for x sufficiently large when considering the limit $x \to \infty$.

It is immediate from the definitions that $f(x) = o(g(x))$ implies $f(x) = O(g(x))$. Moreover, the little-o and big-O relations are transitive. Namely, if $f(x) = o(g(x))$ and $g(x) = o(h(x))$, then $f(x) = o(h(x))$. Similarly, if $f(x) = O(g(x))$ and $g(x) = O(h(x))$, then $f(x) = O(h(x))$ (see [36, Ch. 1, p. 20]).

Another notation we use frequently is the following.

Definition 2.1.4 (Asymptotic equivalence)**.** *Let* f *and* g *be functions defined in a set* D*. We write*

$$f(x) \sim g(x) \quad \text{as } x \to x_0$$

if

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 1.$$

Remark 2.1.5. *Note that the above relation can equivalently be written as*

$$f(x) = g(x)(1 + o(1)) \qquad \text{as } x \to x_0.$$

2.1.2 Types of asymptotic results: exact and logarithmic

We distinguish between different levels of precision in asymptotic approximations. In particular, we consider logarithmic asymptotics, exact asymptotics, and, more informally, refined (or sharp) asymptotics, as they appear in our results. This distinction helps clarify the level of precision achieved by the asymptotic descriptions.

For simplicity, we state the definitions for functions as $x \to \infty$. The same notions can be defined for sequences (a_n) as $n \to \infty$, and similarly for limits $x \to x_0$.

Definition 2.1.6 (Logarithmic asymptotics)**.** *Let f and g be positive functions defined in a neighbourhood of infinity. We say that f and g have the same* logarithmic asymptotics *as $x \to \infty$ if*

$$\log f(x) \sim \log g(x), \qquad x \to \infty.$$

Definition 2.1.7 (Exact asymptotics)**.** *Let f and g be functions defined in a neighbourhood of infinity with $g(x) \neq 0$ for all sufficiently large x. We say that f and g have the same* exact asymptotics *as $x \to \infty$ if*

$$f(x) \sim g(x), \qquad x \to \infty.$$

In our context, f will represent a tail probability function, that is,

$$f(x) = \mathbb{P}(X > x)$$

for some random variable X, while g may either be of the same form for another random variable or be a function describing the asymptotic behaviour of f.

Similarly, sequences may arise in the form

$$a_n = \mathbb{P}(X_n > c_n)\,,$$

where $(X_n)_{n\geq 0}$ is a sequence of random variables, $(c_n)_{n\geq 0}$ a sequence of reals, and $n \to \infty$.

Remark 2.1.8 (Exact asymptotics imply logarithmic asymptotics)**.** *Exact asymptotics imply logarithmic asymptotics under the mild condition that $g(x) \not\to 1$. Indeed, if $f(x) \sim$*

$g(x)$ and $f(x), g(x) > 0$ for sufficiently large x, then

$$f(x) = g(x)(1 + o(1)),$$

and hence

$$\log f(x) = \log g(x) + o(1).$$

Therefore,

$$\frac{\log f(x)}{\log g(x)} \to 1.$$

In our setting this condition is automatically satisfied, since both f and g typically represent tail probabilities and therefore converge to 0 as $x \to \infty$.

Let us explore some examples to illustrate the difference between these notions.

Example 2.1.9. *Let*

$$g(x) = e^{-x}, \qquad f(x) = x^a e^{-x}, \qquad a \in \mathbb{R}.$$

Then

$$\frac{\log f(x)}{\log g(x)} = \frac{-x + a \log x}{-x} = 1 - \frac{a \log x}{x} \to 1,$$

so g provides the logarithmic asymptotics of f. However, it does not provide the exact asymptotics unless $a = 0$, since

$$\frac{f(x)}{g(x)} = x^a \to \begin{cases} 0, & a < 0, \\ \infty, & a > 0. \end{cases}$$

We provide two interpretations of logarithmic asymptotics and how they differ from exact asymptotics. The first interpretation highlights that logarithmic asymptotics approximate a function f only up to a relative error that -in contrast to exact ones- may not vanish. Nevertheless, they provide valuable information since they identify the leading exponential term governing the decay (or growth) of f.

Remark 2.1.10 (Logarithmic approximation: non–vanishing relative error)**.** *Exact asymptotics provide an approximation with a vanishing relative error. Indeed, if $f(x) \sim$*

g(x), then

$$f(x) = g(x)(1 + o(1)),$$

so g approximates f up to a multiplicative relative error $o(1)$ that vanishes as $x \to \infty$.

In contrast, logarithmic asymptotics provide an approximation with a vanishing relative error only in the logarithmic scale. That is, if

$$\log f(x) \sim \log g(x),$$

then

$$\log f(x) = (1 + o(1)) \log g(x).$$

To interpret this relation, write

$$g(x) = e^{\phi(x)},$$

for some function $\phi(x)$ such that $\phi(x) \to -\infty$. Then

$$\log f(x) = \phi(x)(1 + o(1)),$$

and therefore

$$f(x) = e^{\phi(x)(1+o(1))} = e^{\phi(x)+o(\phi(x))} = g(x)e^{o(\phi(x))}. \tag{2.1.1}$$

Hence the relative error between f and g is

$$\frac{f(x) - g(x)}{f(x)} = 1 - e^{-o(\phi(x))}.$$

Since a function that is $o(\phi(x))$ does not necessarily converge to 0, the relative error need not vanish and may even diverge.

Remark 2.1.11 (Logarithmic approximation: leading exponential term)**.** *Equation* (2.1.1)*, and in particular its second equality, highlights another interpretation of logarithmic asymptotics. A logarithmic asymptotic approximation of f, here $g(x) = e^{\phi(x)}$, identifies the leading exponential term $\phi(x)$ governing the decay of f.*

The previous Example 2.1.9 illustrates this phenomenon. Writing

$$f(x) = x^a e^{-x} = e^{-x+a\log x},$$

We see that $g(x) = e^{-x}$ *captures the leading exponential term* $-x$*, while the term* $a\log x = o(x)$ *represents a lower–order correction that is not captured by logarithmic asymptotics. More generally, if*

$$f(x) = p(x)e^{h(x)}, \qquad \log p(x) = o(|h(x)|),$$

then $g(x) = e^{h(x)}$ *provides the logarithmic asymptotics of* f*, but it does not in general provide the exact asymptotics unless* $p(x) \to 1$*.*

Example 2.1.12 (Cramér-type logarithmic asymptotics)**.** *Let* (X_i) *be i.i.d. real-valued random variables such that*

$$\varphi(t) = \mathbb{E}e^{tX_1} < \infty, \qquad t \in \mathbb{R}.$$

Let $S_n = \sum_{i=1}^n X_i$*. Then, for any* $a > \mathbb{E}X_1$*, Cramér's theorem (see [31, Ch. 1, Sec. 3]) states that*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq an) = -I(a),$$

where

$$I(a) = \sup_{t\in\mathbb{R}}\{at - \log\varphi(t)\}$$

is the associated rate function.

Thus

$$\log \mathbb{P}(S_n \geq an) = -I(a)n\,(1 + o(1))\,, \tag{2.1.2}$$

or

$$\log \mathbb{P}(S_n \geq an) \sim \log e^{-I(a)n}.$$

In particular, this provides logarithmic asymptotics for the probability $\mathbb{P}(S_n \geq an)$*.*

Remark 2.1.13 (Characterization of our results)**.** *In relation to our results, Theorem 4.1.9 provides exact asymptotics, while Theorem B.3.1 gives logarithmic asymptotics.*

On the other hand, Proposition 4.3.1 yields exact asymptotics when the positive parameter b *satisfies* $b \geq 1$*, and refined logarithmic asymptotics when* $0 < b < 1$*.*

By refined *we mean that the result identifies additional orders beyond the leading exponential term of decay. In contrast to standard logarithmic asymptotic approximations (see Remark 2.1.11) or classical large deviation results such as Cramér's theorem (see Example 2.1.12), which determine only the leading exponential rate of decay, our asymptotics also capture lower–order corrections (see also Remark 4.3.3).*

2.2 Tails of distributions

In this section we discuss several notions describing the tail behaviour of probability distributions, and introduce classes of random variables that will appear later in the thesis. Our exposition of the basic notions largely follows [38].

Before proceeding, we define the notion of a *tail probability.* For a random variable X with distribution function F, the (right) *tail probability function*, also called the *survival function*, is defined as

$$\overline{F}(x) := 1 - F(x) = \mathbb{P}(X > x)\,.$$

2.2.1 Light-tailed distributions

Definition 2.2.1 (Light tails)**.** *A random variable X is said to be* light-tailed *if its moment generating function is finite in a neighbourhood of the origin, that is,*

$$\exists \theta > 0 \quad \textit{such that} \quad M(\theta) := \mathbb{E}\big[e^{\theta X}\big] < \infty.$$

Random variables that do not satisfy the above condition are called *heavy-tailed* and are covered in Subsection 2.2.2.

As we discuss in the next remark, a random variable has light tails when its tail probability function decays at least exponentially fast. For this reason the exponential distribution plays a natural boundary between light-tailed and heavy-tailed distributions.

Remark 2.2.2 (Equivalent characterization of light-tailed random variables)**.** *A random variable X satisfies Definition 2.2.1 if and only if the tail of X is bounded above by an exponential function. In other words, X is light-tailed if and only if there exists $\theta_0 > 0$ such that*

$$\limsup_{x\to\infty} \frac{\overline{F}(x)}{e^{-\theta_0 x}} < \infty. \tag{2.2.1}$$

Proof. For the "only if" direction, observe that if $\mathbb{E}\left[e^{\theta_0 X}\right] < \infty$ for some $\theta_0 > 0$, then by Markov's inequality,

$$\overline{F}(x) = \mathbb{P}(X > x) = \mathbb{P}\big(e^{\theta_0 X} > e^{\theta_0 x}\big) \leq e^{-\theta_0 x}\mathbb{E}\left[e^{\theta_0 X}\right], \qquad x \geq 0.$$

Thus (2.2.1) holds, and the tail probability decays at least exponentially.

The converse implication is proved later in Remark 2.2.11. □

Typical examples of light-tailed random variables include those with Gaussian, Exponential, or Gamma distributions, as well as those with Weibull distribution with shape parameter $\alpha \geq 1$. All bounded random variables are also light-tailed, for instance those with uniform or beta distribution. In addition, many widely used discrete random variables, such as those with geometric, Poisson, or binomial distribution, also belong to this class. The justification follows immediately by checking that their moment generating function is finite in a neighbourhood of the origin (see also [38, p. 15] for further discussion on the Weibull case).

We now introduce a definition used by Braverman in several of his papers (see [13, 14, 15]), which will also be used frequently throughout this thesis. Braverman refers to such random variables simply as "light". To avoid confusion with the classical notion introduced above, we will refer to them here as *Braverman-light.*

Definition 2.2.3 (Braverman-light tails)**.** *We say that the distribution of a random variable X has a* ***Braverman-light tail*** *if one of the following conditions holds:*

$$\mathbb{P}(X > x) > 0 \quad \text{for all } x > 0, \quad \text{and} \quad \lim_{x\to\infty} \frac{\mathbb{P}(X_1 > x)}{\mathbb{P}(X_1 + X_2 > x)} = 0; \tag{2.2.2}$$

$$\mathbb{P}(X > 0) > 0 \quad \text{and} \quad X \leq C \quad \text{for some constant } C, \tag{2.2.3}$$

where X_1 and X_2 are independent copies of X.

Most of the usual light-tailed distributions are also Braverman-light. In fact, all the examples listed above as light-tailed belong to the Braverman-light class as well. Bounded random variables are Braverman-light immediately from condition (2.2.3). We now briefly illustrate this for several classical distributions.

Example 2.2.4 (Gaussian distribution)**.** *Let $X \sim N(0, \sigma^2)$. Then $X_1 + X_2 \sim N(0, 2\sigma^2)$.*

Using the standard Gaussian tail asymptotics (see [30, p. 98]),

$$\mathbb{P}(X > x) \sim \frac{\sigma}{x\sqrt{2\pi}} e^{-x^2/(2\sigma^2)}, \qquad \mathbb{P}(X_1 + X_2 > x) \sim \frac{\sqrt{2}\sigma}{x\sqrt{2\pi}} e^{-x^2/(4\sigma^2)},$$

and therefore

$$\frac{\mathbb{P}(X > x)}{\mathbb{P}(X_1 + X_2 > x)} \sim \frac{1}{\sqrt{2}} e^{-x^2/(4\sigma^2)} \to 0.$$

Thus the Gaussian distribution is Braverman-light.

Example 2.2.5 (Gamma distribution)**.** *Let* $X \sim \Gamma(\alpha, \lambda)$*, so that*

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \qquad x > 0.$$

Then $X_1 + X_2 \sim \Gamma(2\alpha, \lambda)$*, and therefore*

$$\mathbb{P}(X > x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_x^\infty t^{\alpha-1} e^{-\lambda t}\, dt,$$

while

$$\mathbb{P}(X_1 + X_2 > x) = \frac{\lambda^{2\alpha}}{\Gamma(2\alpha)} \int_x^\infty t^{2\alpha-1} e^{-\lambda t}\, dt.$$

Hence

$$\frac{\mathbb{P}(X > x)}{\mathbb{P}(X_1 + X_2 > x)} = \frac{\Gamma(2\alpha)}{\Gamma(\alpha)\lambda^\alpha} \cdot \frac{\int_x^\infty t^{\alpha-1} e^{-\lambda t}\, dt}{\int_x^\infty t^{2\alpha-1} e^{-\lambda t}\, dt}.$$

Now both numerator and denominator tend to 0 *as* $x \to \infty$*, so by l'Hôspital's rule,*

$$\lim_{x\to\infty} \frac{\int_x^\infty t^{\alpha-1} e^{-\lambda t}\, dt}{\int_x^\infty t^{2\alpha-1} e^{-\lambda t}\, dt} = \lim_{x\to\infty} \frac{x^{\alpha-1} e^{-\lambda x}}{x^{2\alpha-1} e^{-\lambda x}} = \lim_{x\to\infty} x^{-\alpha} = 0.$$

Therefore,

$$\lim_{x\to\infty} \frac{\mathbb{P}(X > x)}{\mathbb{P}(X_1 + X_2 > x)} = 0,$$

and so the Gamma distribution is Braverman-light.

Example 2.2.6 (Weibull distribution)**.** *Let* X *have a Weibull*(α, β) *distribution, so that*

$$\mathbb{P}(X > x) = e^{-(\beta x)^\alpha}, \qquad x \geq 0.$$

First consider the case $\alpha > 1$. Since

$$\mathbb{P}(X_1 + X_2 > x) \geq \mathbb{P}(X_1 > x/2,\ X_2 > x/2)\,,$$

we obtain

$$\mathbb{P}(X_1 + X_2 > x) \geq \mathbb{P}(X > x/2)^2 = e^{-2(\beta x/2)^\alpha}.$$

Therefore,

$$\frac{\mathbb{P}(X > x)}{\mathbb{P}(X_1 + X_2 > x)} \leq \frac{e^{-(\beta x)^\alpha}}{e^{-2(\beta x/2)^\alpha}} = \exp(-(\beta x)^\alpha + 2(\beta x/2)^\alpha)\,.$$

Since

$$-(\beta x)^\alpha + 2(\beta x/2)^\alpha = -(\beta x)^\alpha \big(1 - 2^{1-\alpha}\big),$$

and $1 - 2^{1-\alpha} > 0$ for $\alpha > 1$, it follows that

$$\frac{\mathbb{P}(X > x)}{\mathbb{P}(X_1 + X_2 > x)} \to 0, \qquad x \to \infty.$$

If $\alpha = 1$, then X is exponentially distributed, equivalently $X \sim \Gamma(1, \beta)$, and this case is covered by Example 2.2.5.

Remark 2.2.7 (Braverman-light r.v.'s form a strict subclass of light-tailed)**.** *Although most typical examples of light-tailed random variables are also Braverman-light, the latter class forms a strict subclass of the former.*

First we explain why Braverman-light random variables are necessarily light-tailed. Let X have Braverman-light tails and suppose X is unbounded. If X is bounded, then it is immediately light-tailed since its moment generating function $M(\theta)$ is finite for all $\theta \in \mathbb{R}$.

By [14, Lemma 2], if X is Braverman-light then the positive part $X^{(+)} := \max(X, 0)$ is also Braverman-light. In particular,

$$\lim_{x\to\infty} \frac{\mathbb{P}\Big(X_1^{(+)} > x\Big)}{\mathbb{P}\Big(X_1^{(+)} + X_2^{(+)} > x\Big)} = 0,$$

where $X_1^{(+)}, X_2^{(+)}$ are i.i.d. copies of $X^{(+)}$.

However, the above behaviour is impossible for heavy-tailed distributions. Indeed, for *every non-negative heavy-tailed random variable* Y *it holds that*

$$\limsup_{x\to\infty} \frac{\mathbb{P}(Y_1 > x)}{\mathbb{P}(Y_1 + Y_2 > x)} = \frac{1}{2}, \tag{2.2.4}$$

see [27, Theorem 2.12]. Hence $X^{(+)}$ *cannot be heavy-tailed.*

Since heavy-tailedness is a tail property (see [27, p. 8]) and X *and* $X^{(+)}$ *have identical right tails, it follows that* X *is also not heavy-tailed. Therefore* X *must be light-tailed.*

On the other hand, there exist distributions that are light-tailed but not Braverman-light as we show in the next two examples.

The following construction was used in [27, p. 16] to show that there exist light-tailed random variables satisfying (2.2.4). For our purposes, it provides an example of a light-tailed random variable that is not Braverman-light.

Example 2.2.8. *Let* X *be a discrete random variable with probability distribution* $(p_n)_{n\in\mathbb{N}}$ *supported on* $(x_n)_{n\in\mathbb{N}}$*, where* $x_0 = 1$ *and* $x_{n+1} > 2x_n$ *for every* n*.*

Let X_1, X_2 *be i.i.d. copies of* X*. Since the support grows faster than a factor of two, if* $X_1 \leq x_{n-1}$ *and* $X_2 \leq x_{n-1}$ *then*

$$X_1 + X_2 \leq 2x_{n-1} < x_n.$$

Hence the event $\{X_1 + X_2 > x_n - 1\} = \{X_1 + X_2 \geq x_n\}$ *can occur only if at least one of the variables exceeds* x_{n-1} *and therefore* $x_n - 1$*. Splitting according to whether* $X_1 > x_n - 1$ *or not yields*

$$\mathbb{P}(X_1 + X_2 > x_n - 1) = \mathbb{P}(X_1 > x_n - 1) + \mathbb{P}(\{X_1 \leq x_n - 1\} \cap \{X_2 > x_n - 1\})\,.$$

Using independence we obtain

$$\mathbb{P}(X_1 + X_2 > x_n - 1) = \mathbb{P}(X_1 > x_n - 1) + \mathbb{P}(X_1 \leq x_n)\,\mathbb{P}(X_2 > x_n - 1)\,.$$

Since $\mathbb{P}(X_1 \leq x_n) \to 1$ *as* $n \to \infty$*, it follows that*

$$\mathbb{P}(X_1 + X_2 > x_n - 1) \sim 2\,\mathbb{P}(X_1 > x_n - 1)\,, \qquad n \to \infty,$$

and therefore

$$\lim_{n\to\infty} \frac{\mathbb{P}(X_1 > x_n - 1)}{\mathbb{P}(X_1 + X_2 > x_n - 1)} = \frac{1}{2}.$$

and hence X cannot be Braverman-light.

On the other hand, taking $x_n = 3^n$ and $p_n = ce^{-3^n}$ yields a light-tailed distribution, since the tail probabilities decay faster than exponentially (see Remark 2.2.2). Interestingly, in this example the failure of Braverman-lightness is determined solely by the choice of the support.

The following distribution was constructed in [27, p. 45] to show that the condition defining subexponential distributions, namely (2.2.9), does not imply long-tailedness in general, unlike in the case of non-negative random variables[1]. For our purposes, this construction provides a convenient example of a continuous random variable that is light-tailed but not Braverman-light.

Example 2.2.9. *For $A \geq 0$, consider a random variable X with distribution F on $[-A, \infty)$ and survival function*

$$\overline{F}(x) = (x + A + 1)^{-2} e^{-(x+A)}, \qquad x \geq -A.$$

First, X is light-tailed. Indeed, since $(x + A + 1)^{-2} \sim x^{-2}$ as $x \to \infty$, we have

$$\overline{F}(x) \sim x^{-2} e^{-(x+A)} = e^{-x-2\log x - A}.$$

Thus $\overline{F}(x)$ decays at least exponentially fast, and hence X is light-tailed (see Remark 2.2.2).

On the other hand,

$$\mathbb{P}(X_1 + X_2 > x) = \int_{-\infty}^{\infty} \overline{F}(x - y) F(dy)$$

$$= \int_{-\infty}^{x/2} \overline{F}(x - y) F(dy) - \int_{x/2}^{\infty} \overline{F}(x - y)\, dF(y)$$

$$= 2 \int_{-\infty}^{x/2} \overline{F}(x - y) F(dy) + \left(\overline{F}(x/2)\right)^2,$$

[1] We formally introduce the aforementioned notions later in the thesis.

after integration by parts. We thus have, as $x \to \infty$,

$$\overline{F * F}(x) \sim 2e^{-x-A} \int_{-A}^{x/2} (x-y)^{-2} e^{y} F(dy) + o(\overline{F}(x))$$

$$\sim 2x^{-2} e^{-x-A} \int_{-A}^{x/2} e^{y} F(dy) + o(\overline{F}(x))$$

$$\sim 2\overline{F}(x) \int_{-A}^{\infty} e^{y} F(dy).$$

Hence

$$\lim_{x\to\infty} \frac{\mathbb{P}(X > x)}{\mathbb{P}(X_1 + X_2 > x)} = \frac{1}{2\displaystyle\int_{-A}^{\infty} e^{y} F(dy)} \neq 0, \tag{2.2.5}$$

and therefore X is not Braverman-light.

2.2.2 Heavy-tailed distributions

Heavy-tailed distributions arise naturally in many areas of everyday life, for instance in the distribution of wealth, insurance claim sizes, or service waiting times. Nevertheless, we are generally less familiar with such distributions than with light-tailed ones. Moreover, the term *heavy-tailed* is sometimes used rather loosely: it may refer to power-law behaviour, an increasing mean residual life, infinite variance, and so on. In what follows, we restrict ourselves to precise mathematical definitions of different levels and aspects of heavy-tailedness. Each definition provides a different intuition for the concept of heavy tails.

We also highlight the different mechanisms through which sums of heavy- and light-tailed random variables attain extreme values, namely the contrast between the *catastrophe principle* and the *synergy principle*. These ideas will play an important role later in our treatment of Lévy processes and compound renewal processes.

Our exposition is mainly based on [38].

Heavy-tailed

So far we have informally characterized heavy-tailed distributions as those that are not light-tailed. We now give a precise definition.

Definition 2.2.10 (Heavy tails)**.** *A random variable X is said to be* heavy-tailed ($X \in \mathcal{H}$) *if*

$$\mathbb{E}\big[e^{\theta X}\big] = \infty \qquad \text{for all } \theta > 0.$$

As expected, and as implied by the equivalent characterization of light-tailed random variables in the previous section, heavy-tailed random variables admit an equivalent definition in terms of the rate of decay of the tail probability function.

Remark 2.2.11 (Equivalent characterization of heavy-tailed random variables)**.** *A random variable is heavy-tailed if and only if its tail probability decays slower than any exponential function. More precisely, a random variable X with tail function $\overline{F}$ is heavy-tailed if and only if for all $\theta > 0$,*

$$\limsup_{x\to\infty} \frac{\overline{F}(x)}{e^{-\theta x}} = \infty. \tag{2.2.6}$$

Proof. We prove here the "only if" direction. The converse follows from Remark 2.2.2, where the equivalent characterization of light-tailed random variables was established.

Assume that X is heavy-tailed and argue by contradiction. Suppose that there exists $\theta_0 > 0$ such that

$$\limsup_{x\to\infty} \frac{\overline{F}(x)}{e^{-\theta_0 x}} < \infty.$$

Then, by the definition of the limit superior, there exist constants $C > 0$ and $x_0 > 0$ such that

$$\overline{F}(x) \le C e^{-\theta_0 x}, \qquad x \ge x_0.$$

Let $X^+ := \max\{X, 0\}$. Since $e^{\eta X} \le 1 + e^{\eta X^+}$, it suffices to show that $\mathbb{E}\Big[e^{\eta X^+}\Big] < \infty$ for some $\eta > 0$.

Now for every non-negative random variable Y it holds that

$$\mathbb{E}\big[e^{\eta Y}\big] = 1 + \eta \int_0^\infty e^{\eta x}\mathbb{P}(Y > x)\, dx,$$

see Fact A.2.1 in Section A.2. Applying this identity with $Y = X^+$ yields

$$\mathbb{E}\Big[e^{\eta X^+}\Big] = 1 + \eta \int_0^\infty e^{\eta x}\overline{F}(x)\, dx.$$

Fix any $\eta \in (0, \theta_0)$. Then

$$\int_0^\infty e^{\eta x}\overline{F}(x)\,dx = \int_0^{x_0} e^{\eta x}\overline{F}(x)\,dx + \int_{x_0}^\infty e^{\eta x}\overline{F}(x)\,dx.$$

The first integral is finite. For the second we use the exponential bound:

$$\int_{x_0}^\infty e^{\eta x}\overline{F}(x)\,dx \le C\int_{x_0}^\infty e^{-(\theta_0-\eta)x}\,dx < \infty.$$

Hence $\mathbb{E}\left[e^{\eta X^+}\right] < \infty$, and therefore $\mathbb{E}\left[e^{\eta X}\right] < \infty$, which contradicts the assumption that X is heavy-tailed.

Thus for every $\theta > 0$,

$$\limsup_{x\to\infty} \frac{\overline{F}(x)}{e^{-\theta x}} = \infty.$$

□

Loosely speaking, both characterizations express the same intuition: heavy-tailed distributions place more probability mass on extreme values than light-tailed distributions.

Examples of heavy-tailed distributions include the Weibull distribution with shape parameter $\alpha < 1$, the lognormal distribution, the Pareto distribution, the Cauchy distribution, and the Fréchet distribution. Further discussion of these examples can be found in [38, Ch. 1, Sec. 2].

Long-tailed

Long-tailed distributions are closely connected to the notion of the *residual life distribution*, which we introduce below.

Definition 2.2.12 (Residual life distribution)**.** *Let X be a real-valued random variable with distribution function F and tail probability function $\overline{F}(x) = 1 - F(x)$. For x with $\overline{F}(x) > 0$, the residual life distribution at threshold x is the law of $X - x$ conditioned on $\{X > x\}$. Its distribution function is*

$$R_x(y) \;:=\; 1 - \mathbb{P}(X > x+y \,|\, X > x) \;=\; 1 - \frac{\overline{F}(x+y)}{\overline{F}(x)}, \qquad y \ge 0. \tag{2.2.7}$$

Equivalently, its tail probability function is $\overline{R_x}(y) = \dfrac{\overline{F}(x+y)}{\overline{F}(x)}$.

Definition 2.2.13 (Long tails)**.** *A random variable X is said to be* long-tailed *($X \in \mathcal{L}$) if for every fixed $y \geq 0$,*

$$\lim_{x\to\infty} \overline{R_x}(y) = 1. \tag{2.2.8}$$

Long-tailed distributions form a subclass of heavy-tailed distributions; see [38, Ch. 4, Sec. 3]. In practice, most commonly encountered heavy-tailed distributions are also long-tailed, and constructing a heavy-tailed distribution that is not long-tailed is not entirely straightforward. In particular, all of the aforementioned examples of heavy-tailed distributions are long-tailed as well: the Weibull distribution with shape parameter $\alpha < 1$, the lognormal distribution, the Pareto distribution, the Cauchy distribution, and the Fréchet distribution.

The notion of long-tailedness provides another perspective on heavy-tailed behaviour. Roughly speaking, if a long-tailed random variable exceeds a large threshold x, it is likely to exceed it by an arbitrarily large amount. In other words, the defining feature of long-tailed (or heavy-tailed) random variables is not that they frequently take extreme values, but rather that when they do take a large value, it can be extremely large.

Another intuition comes from the concept of the *expected residual life*, denoted by

$$m(x) := \mathbb{E}[X - x \,|\, X > x] \,.$$

Most light-tailed distributions have $m(x)$ that is eventually decreasing, while most heavy-tailed distributions have $m(x)$ that is eventually increasing (see [38, Ch. 4, Sec. 2]). The memoryless property of the exponential distribution implies that it is the only distribution with constant $m(x)$, once again showing that it forms a natural boundary between light and heavy tails.

The class of long-tailed distributions makes the above heuristic precise. More specifically, whenever $\lim_{x\to\infty} m(x)$ exists, a random variable is long-tailed if and only if

$$\lim_{x\to\infty} m(x) = \infty$$

(see [38, p. 94]).

The behaviour of the expected residual life of a long-tailed distribution may at first seem counterintuitive. In many familiar situations, the longer we have already waited for an event, the shorter the remaining expected waiting time becomes. However, for long-tailed distributions the opposite behaviour may occur: the expected residual life increases with the elapsed time.

Subexponential distributions

Subexponential distributions are closely connected with convolutions and sums of i.i.d. random variables.

Definition 2.2.14 (Subexponential distributions). *A random variable X is said to be* subexponential *($X \in \mathcal{S}$) if it is long-tailed and*

$$\lim_{x\to\infty} \frac{\mathbb{P}(X_1 > x)}{\mathbb{P}(X_1 + X_2 > x)} = \frac{1}{2}, \tag{2.2.9}$$

where X_1, X_2 are independent copies of X.

In general, subexponentiality implies that

$$\overline{F^{*n}}(x) \sim n\,\overline{F}(x), \qquad x \to \infty, \tag{2.2.10}$$

where F^{*n} denotes the n-fold convolution of F. For non-negative random variables, the converse also holds, so that (2.2.10) is an equivalent definition of subexponentiality in that case; see [27, p. 52].

The class of subexponential distributions is a subclass of the heavy-tailed distributions, since every subexponential distribution is long-tailed. In fact, all the examples mentioned above as heavy-tailed are also subexponential: the Weibull distribution with shape parameter $\alpha < 1$, the lognormal distribution, the Pareto distribution, the Cauchy distribution, and the Fréchet distribution; see [38, p. 64].

For non-negative random variables, subexponentiality is defined directly through condition (2.2.9), since in that case it already implies long-tailedness [27, p. 44]. However, this implication does not hold in general when random variables may take negative values. That is the case with the distribution presented in Example 2.2.9, if we choose A such that the integral in (2.2.5) equals 1. For this reason the definition above explicitly includes the condition of long-tailedness in order to preserve this property.

Catastrophe vs. Synergy. The notion of the *single big jump*, or catastrophe principle, is prevalent among heavy-tailed distributions. In terms of sums of heavy-tailed i.i.d. random variables, this means that a large sum typically occurs due to the contribution of a single very large summand. For non-negative random variables, subexponential distributions form precisely the class for which this principle holds. Namely, when X is non-negative, condition (2.2.9) -and therefore (2.2.10)- is equivalent to

$$\mathbb{P}(X_1 + X_2 + \cdots + X_n > x) \sim \mathbb{P}(\max(X_1, X_2, \ldots, X_n) > x), \qquad x \to \infty. \tag{2.2.11}$$

This follows from (2.2.10) together with

$$\mathbb{P}(\max(X_1, \ldots, X_n) > x) = 1 - (1 - \mathbb{P}(X > x))^n \sim n\,\mathbb{P}(X > x), \qquad x \to \infty,$$

which holds for every random variable.

A non-negative random variable that satisfies (2.2.11) is said to satisfy the *catastrophe principle*; see [38].

In contrast, sums of i.i.d. light-tailed random variables typically attain large values through a synergistic effect between the summands, where no single summand is of the same order as the sum. In [38], a non-negative random variable is said to satisfy the *conspiracy principle* if

$$\mathbb{P}(\max(X_1, X_2, \ldots, X_n) > x) = o(\mathbb{P}(X_1 + X_2 + \cdots + X_n > x)), \qquad x \to \infty.$$

Most common light-tailed distributions satisfy the conspiracy principle. Examples include the exponential, Gaussian, and Weibull distributions with shape parameter $\alpha > 1$; see [38, p. 60].

Regularly varying distributions

Regularly varying distributions are closely related to the notion of a power law. Roughly speaking, their tail probability function $\overline{F}(x)$ behaves asymptotically like a power of x. To describe this precisely we first introduce the concept of regularly varying functions.

Definition 2.2.15 (Regularly varying function)**.** *A measurable function $f : (0, \infty) \to (0, \infty)$ is said to be* regularly varying of index $\alpha \in \mathbb{R}$*, written $f \in \mathcal{RV}(\alpha)$, if for every*

$\lambda > 0$

$$\lim_{x\to\infty} \frac{f(\lambda x)}{f(x)} = \lambda^{\alpha}.$$

If $\alpha = 0$, the function is called slowly varying.

Slowly varying functions may be thought of as functions that grow or decay asymptotically more slowly than any power of x. Typical examples include $\log x$, $\log\log x$, and constant functions.

A fundamental property of regularly varying functions is that they are asymptotically of power-law type: $f \in \mathcal{RV}(\alpha)$ if and only if there exists a slowly varying function L such that

$$f(x) = x^{\alpha} L(x).$$

Indeed, given $f \in \mathcal{RV}(\alpha)$ one may take $L(x) = f(x)/x^{\alpha}$.

Definition 2.2.16 (Regularly varying random variables)**.** *A random variable X with distribution function F is said to be* regularly varying with index $\alpha > 0$, *written $X \in \mathcal{RV}(\alpha)$ if its tail probability function satisfies*

$$\overline{F} \in \mathcal{RV}(-\alpha).$$

Equivalently,

$$\overline{F}(x) = x^{-\alpha} L(x),$$

for some slowly varying function L.

A non-negative regularly varying random variable with index $-\alpha$, i.e. $X \in \mathcal{RV}(-\alpha)$, has finite moments only up to order α. In particular, $\mathbb{E}[X^m] < \infty$ for $m < \alpha$, while $\mathbb{E}[X^m] = \infty$ for $m > \alpha$, see [38, p. 43].

Regularly varying random variables are closely connected with the notion of *asymptotic scale invariance.* Indeed, if $\overline{F} \in \mathcal{RV}(-\alpha)$, then for every $\lambda > 0$

$$\overline{F}(\lambda x) \sim \lambda^{-\alpha} \overline{F}(x), \qquad x \to \infty,$$

so that the shape of the tail remains asymptotically unchanged under multiplicative scaling.

Regular variation represents the strongest form of heavy-tailed behaviour encountered in this chapter. In particular, regularly varying distributions form a subclass of the subexponential distributions [27, p. 57]. Well-known examples include the Pareto distribution, the Cauchy distribution, and the Fréchet distribution. The Pareto distribution is the most typical example, since its tail follows an exact power law.

On the other hand, there are heavy-tailed distributions which are not regularly varying, such as the lognormal distribution and the Weibull distribution with shape parameter $\alpha < 1$; see [38, p. 64].

In conclusion, we have the strict inclusions

$$\mathcal{RV} \subset \mathcal{S} \subset \mathcal{L} \subset \mathcal{H}.$$

2.2.3 Exponential tails

The following class of probability distributions, commonly denoted by $\mathcal{L}(\alpha)$ in the literature, appears frequently in the study of convolution tails (as does the related class $\mathcal{S}(\alpha)$ introduced later); see, for example, [1, 17, 22, 21]. Following [17], we refer to such distributions as having *exponential tails*. The reason for this terminology will become clear after the definition.

Definition 2.2.17 (Exponential tails). *A random variable X is said to have* exponential tails *with index α ($X \in \mathcal{L}(\alpha)$) if*

$$\lim_{x\to\infty} \overline{R_x}(y) = \lim_{x\to\infty} \frac{\overline{F}(x+y)}{\overline{F}(x)} = e^{-\alpha y}, \qquad y \geq 0. \tag{2.2.12}$$

The exponential random variable $X \sim \text{Exp}(\alpha)$ is the prime example of a random variable in $\mathcal{L}(\alpha)$. In fact, it serves as the prototype of this class, since (2.2.12) holds for it with equality rather than merely asymptotically. This is precisely the well-known memoryless property of the exponential distribution.

Another important subclass is the class of long-tailed distributions $\mathcal{L}$. Indeed, by comparing Definition 2.2.13 with (2.2.12), we see that $\mathcal{L}(0) = \mathcal{L}$. The following equivalent characterization clarifies the structure of exponential tails and motivates the terminology.

Remark 2.2.18 (Equivalent characterization of exponential tails). *A random variable*

X with distribution function F belongs to $\mathcal{L}(\alpha)$ *if and only if*

$$\overline{F}(x) = W(x)e^{-\alpha x}, \tag{2.2.13}$$

where W is a long-tailed function, that is,

$$\frac{W(x+y)}{W(x)} \to 1 \qquad (x \to \infty)$$

for every fixed y.

Proof. If $X \in \mathcal{L}(\alpha)$, define

$$W(x) := \overline{F}(x)e^{\alpha x}.$$

Then

$$\frac{W(x+y)}{W(x)} = \frac{\overline{F}(x+y)}{\overline{F}(x)}e^{\alpha y} \to e^{-\alpha y}e^{\alpha y} = 1,$$

so W is long-tailed.

Conversely, if $\overline{F}(x) = W(x)e^{-\alpha x}$ with W long-tailed, then

$$\frac{\overline{F}(x+y)}{\overline{F}(x)} = \frac{W(x+y)}{W(x)}e^{-\alpha y} \to e^{-\alpha y},$$

so $X \in \mathcal{L}(\alpha)$. □

The function W does not necessarily converge to zero. However, since it is long-tailed, it cannot decay faster than any exponential function. Thus the class $\mathcal{L}(\alpha)$ consists of distributions whose tails decay essentially exponentially, up to a multiplicative long-tailed factor.

Another interpretation of this class comes from residual life distributions (see Definition 2.2.12). Condition (2.2.12) means precisely that the residual life distributions R_x converge weakly to an exponential distribution with parameter α. In other words,

$$X - x \mid X > x \xrightarrow{d} \text{Exp}(\alpha).$$

The exponential distribution is in fact the only possible non-degenerate limiting residual life distribution. To see this, suppose that there exists a non-trivial limiting residual life distribution S such that

$$\overline{S}(y) = \lim_{x\to\infty} \frac{\overline{F}(x+y)}{\overline{F}(x)}$$

exists and is positive for every $y > 0$. Then for $y, w > 0$ we can write

$$\frac{\overline{F}(x+y+w)}{\overline{F}(x)} = \frac{\overline{F}(x+y+w)}{\overline{F}(x+y)} \cdot \frac{\overline{F}(x+y)}{\overline{F}(x)}.$$

Taking limits yields

$$\overline{S}(y+w) = \overline{S}(y)\,\overline{S}(w), \qquad y, w > 0.$$

The above relationship characterize the exponential function, hence

$$\overline{S}(y) = e^{-\alpha y}.$$

Other than the exponential distribution itself, examples of distributions belonging to $\mathcal{L}(\alpha)$ can be constructed using the equivalent characterization given earlier. In particular, any distribution whose tail can be written in the form

$$\overline{F}(x) = W(x)e^{-\alpha x},$$

where W is long-tailed, belongs to $\mathcal{L}(\alpha)$. Typical examples are obtained by choosing regularly or slowly varying factors such as

$$\overline{F}(x) = x^{\ell} e^{-\alpha x}, \qquad \overline{F}(x) = (\log x)^{\ell} e^{-\alpha x}.$$

Examples of distributions that do not belong to $\mathcal{L}(\alpha)$ include bounded random variables, such as the uniform distribution. More generally, many common light-tailed distributions fail to satisfy the asymptotic memoryless property (2.2.12). For instance, the Gaussian distribution does not belong to this class.

Indeed, if $X \sim \mathcal{N}(\mu, \sigma^2)$, then

$$\overline{F}(x) = \mathbb{P}(X > x) \sim \frac{\sigma}{x-\mu}\,\phi\left(\frac{x-\mu}{\sigma}\right), \qquad x \to \infty, \tag{2.2.14}$$

where ϕ is the standard normal density. Hence

$$\frac{\overline{F}(x+y)}{\overline{F}(x)} \sim \frac{x-\mu}{x+y-\mu} \exp\left(-\frac{y(x-\mu)}{\sigma^2} - \frac{y^2}{2\sigma^2}\right) \xrightarrow{x\to\infty} 0. \tag{2.2.15}$$

Thus the limiting residual life distribution degenerates at zero. In particular, the pointwise limit of $R_x(y)$ corresponds to the Dirac measure at zero, which is not right-continuous at $y = 0$ and therefore does not define a proper residual life distribution.

Another way to see that the Gaussian distribution does not belong to $\mathcal{L}(\alpha)$ is to note that its tail decays super-exponentially, and therefore cannot be written in the form (2.2.13). The same phenomenon occurs for other light-tailed distributions with super-exponential decay. For example, the Weibull distribution with shape parameter $\alpha > 1$ does not belong to $\mathcal{L}(\alpha)$ for the same reason.

Another class that we will mention later is the following.

Definition 2.2.19 (Convolution equivalent distributions)**.** *A random variable X is said to be* convolution equivalent *with index α ($X \in \mathcal{S}(\alpha)$) if it belongs to $\mathcal{L}(\alpha)$ and*

$$\lim_{x\to\infty} \frac{\mathbb{P}(X_1 + X_2 > x)}{\mathbb{P}(X_1 > x)} = c < \infty, \tag{2.2.16}$$

where X_1, X_2 are independent copies of X.

It turns out that if $X \in \mathcal{S}(\alpha)$ then $c = 2\,\mathbb{E}\big[e^{\alpha X}\big]$; see [22]. In particular, the above and the definition of subexponential distributions (Definition 2.2.14) implies $\mathcal{S}(0) = \mathcal{S}$.

Therefore, all the subexponential distributions mentioned earlier belong to this class. An example of a light-tailed distribution in $\mathcal{S}(\alpha)$ is the distribution constructed in Example 2.2.9 (see (2.2.5)).

Note that condition (2.2.16) excludes Braverman-light-tailed random variables.

2.3 Lévy Processes

We briefly recall the main facts about Lévy processes that will be needed later. Our exposition follows Kyprianou [34].

2.3.1 Definition and Examples

Definition 2.3.1 (Lévy process). *A stochastic process* $(X_t)_{t\geq 0}$ *is called a* Lévy process *if*

1. $X_0 = 0$ *almost surely,*
2. *it has* independent increments, *i.e. for* $0 \leq t_0 < t_1 < \cdots < t_n$, *the random variables* $X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ *are independent,*
3. *it has* stationary increments, *i.e. for every* $s, t \geq 0$, *the law of* $X_{t+s} - X_t$ *depends only on* s,
4. *it has paths that are almost surely right-continuous with left limits (càdlàg).*

Lévy processes generalize both Brownian motion and the Poisson process, and can be thought of as the continuous-time analogue of sums of i.i.d. random variables as discussed in the subsection on infinite divisibility (see Subsection 2.3.2).

Example 2.3.2 (Brownian motion). *Let* $(B_t)_{t\geq 0}$ *be a standard Brownian motion with* $B_0 = 0$. *Then* $(B_t)_{t\geq 0}$ *is a Lévy process with Gaussian increments:*

$$B_t - B_s \sim \mathcal{N}(0, t - s).$$

More generally, the process

$$X_t = \sigma B_t + bt, \qquad \sigma > 0,\ b \in \mathbb{R},$$

is called a linear Brownian motion *(or Brownian motion with linear drift). It is again a Lévy process, with Gaussian increments of mean* $b(t - s)$ *and variance* $\sigma^2(t - s)$.

Example 2.3.3 (Poisson process). *A* Poisson process *with rate* $\lambda > 0$ *is a jump process* $(N_t)_{t\geq 0}$, *with unit jumps, such that the number of jumps up to time* t *follows a Poisson distribution with mean* λt, *i.e.*

$$N_t \sim \text{Poisson}(\lambda t), \qquad t \geq 0.$$

Equivalently, the waiting times between successive jumps are i.i.d. exponential random variables with mean $1/\lambda$.

Example 2.3.4 (Compound Poisson process with linear drift)**.** *A* compound Poisson process *with rate $\lambda > 0$ and jump distribution given by an i.i.d. sequence $(\xi_i)_{i\geq 1}$ is defined by*

$$X_t = \sum_{i=1}^{N_t} \xi_i,$$

where $(N_t)_{t\geq 0}$ is a Poisson process of rate λ, independent of $(\xi_i)_{i\geq 1}$. Then $(X_t)_{t\geq 0}$ is a pure-jump Lévy process with finitely many jumps on each finite interval.

Adding a linear drift gives

$$Y_t = ct + \sum_{i=1}^{N_t} \xi_i, \qquad c \in \mathbb{R},$$

which is called a compound Poisson process with (linear) drift*, and it remains a Lévy process.*

Example 2.3.5 (Sum of compound Poisson process and Brownian motion with drift)**.** *The sum of two (or any finite number of) independent Lévy processes is again a Lévy process (see [34, p. 28]). Therefore, another example of a Lévy process is a process of the form*

$$X_t = \sigma B_t + bt + Y_t, \qquad t \geq 0,$$

where B_t is a standard Brownian motion, $\sigma \geq 0$, $b \in \mathbb{R}$, and Y_t is a compound Poisson process independent of B_t.

The process X_t thus consists of three components: a continuous Gaussian part σB_t, a deterministic linear drift bt, and a jump component Y_t. In particular, the jumps of X_t coincide with the jumps of the compound Poisson process Y_t, while between jumps the process evolves continuously according to the Brownian motion with drift.

An important class of Lévy processes is given by subordinators.

Definition 2.3.6 (Subordinator)**.** *A Lévy process $(X_t)_{t\geq 0}$ is called a* subordinator *if it has almost surely non-decreasing paths.*

The Poisson process defined above is an example of a subordinator.

Another class of Lévy processes we discuss is the class of strictly stable Lévy processes. We restrict attention to strictly stable distributions and processes rather than stable distributions in full generality.

Definition 2.3.7 (Strictly stable random variable)**.** *For* $\alpha \in (0, 2]$*, a random variable* Y *is said to have a* strictly α-stable distribution *if for every* $n \geq 1$

$$\frac{Y_1 + \cdots + Y_n}{n^{1/\alpha}} \stackrel{d}{=} Y, \tag{2.3.1}$$

where $Y_1, \ldots, Y_n$ *are independent copies of* Y*.*

Stable distributions form a strict subclass of the infinitely divisible distributions that we discuss later. This follows directly from their definition. The Gaussian distribution corresponds to the case $\alpha = 2$, while the symmetric Cauchy distribution corresponds to $\alpha = 1$; see [38, p. 115]. For a detailed treatment of stable distributions we refer to [39].

Example 2.3.8 (Strictly stable Lévy processes)**.** *A* strictly α-stable Lévy process $(X_t)_{t \geq 0}$ *is a Lévy process whose increments follow a strictly* α*-stable distribution.*

Important special cases include Brownian motion (corresponding to $\alpha = 2$*) and the symmetric Cauchy process (corresponding to* $\alpha = 1$*).*

Brownian motion is distinguished among strictly α*-stable Lévy processes in that it is the only case with continuous sample paths and increments with finite moments of all orders.*

For $\alpha \neq 2$*, strictly* α*-stable processes have heavy-tailed increments. In fact, by [38, Th. 5.6], there exists a function* $(q_t)_{t \geq 0}$ *with* $q_t > 0$ *such that*

$$\mathbb{P}(|X_t| > x) = (q_t + o(1))x^{-\alpha}, \qquad x \to \infty. \tag{2.3.2}$$

Therefore, the increments are regularly varying, and as discussed in Subsection 2.2.2 for regularly varying distributions, the n*-th moment of the increments is finite only for* $n < \alpha$*.*

Moreover, for $\alpha \neq 2$ *the sample paths are discontinuous, and in fact the set of jump times is dense in every time interval. A justification of this fact is given later in Subsection 2.3.4 (see* (2.3.9) *and the discussion therein).*

Strictly stable Lévy processes also satisfy the self-similarity property

$$X_t \stackrel{d}{=} t^{1/\alpha} X_1, \qquad t > 0, \tag{2.3.3}$$

see [34, p. 12].

2.3.2 Lévy processes and infinite divisibility

Definition 2.3.9 (Infinitely divisible random variable). *A random variable Y is said to be* infinitely divisible *if for every $n \geq 1$ there exist i.i.d. random variables $Y_1, \ldots, Y_n$ such that*

$$Y \stackrel{d}{=} Y_1 + \cdots + Y_n.$$

Instead of referring to infinitely divisible random variables, one may equivalently speak about *infinitely divisible distributions* (or probability laws). We will use these two terms interchangeably.

Definition 2.3.10 (Characteristic exponent). *The* characteristic exponent *of a random variable Y is the function $\Psi : \mathbb{R} \to \mathbb{C}$ defined by*

$$\Psi(\theta) := \log \mathbb{E}\left[e^{i\theta Y}\right], \qquad \theta \in \mathbb{R}.$$

The characteristic exponent uniquely determines the distribution of a random variable since it is a monotone transformation of its characteristic function. One advantage of working with the characteristic exponent is that it behaves particularly well under sums of independent random variables. In particular, the sum of independent random variables has characteristic exponent equal to the sum of their characteristic exponents. For example, the definition of infinite divisibility can be conveniently expressed in terms of characteristic exponents.

Indeed, if Ψ is the characteristic exponent of an infinitely divisible distribution, then by the definition of infinite divisibility, for every $n \geq 1$ there exists a characteristic exponent Ψ_n such that

$$\Psi(\theta) = n\Psi_n(\theta), \qquad \theta \in \mathbb{R}.$$

It turns out that infinitely divisible distributions can be completely characterized through a triplet of parameters. This fundamental result is known as the Lévy–Khintchine formula (see [34, p. 3]). Here and throughout, we use the notation $a \wedge b := \min(a, b)$.

Theorem 2.3.11 (Lévy–Khintchine formula). *A probability law μ of a real-valued random variable is infinitely divisible with characteristic exponent Ψ if and only if*

$$\int_{\mathbb{R}} e^{i\theta x}\, \mu(dx) = e^{-\Psi(\theta)}, \qquad \theta \in \mathbb{R},$$

where

$$\Psi(\theta) = ia\theta + \frac{1}{2}\sigma^2\theta^2 + \int_{\mathbb{R}} \left(1 - e^{i\theta x} + i\theta x \mathbf{1}_{\{|x|<1\}}\right) \Pi(dx). \tag{2.3.4}$$

Here $a \in \mathbb{R}$, $\sigma \geq 0$, *and* Π *is a measure on* $\mathbb{R} \setminus \{0\}$ *satisfying*

$$\int_{\mathbb{R}} (1 \wedge x^2)\, \Pi(dx) < \infty. \tag{2.3.5}$$

Moreover, the triplet (a, σ^2, Π) *is unique.*

Remark 2.3.12 (Integrability condition). *The condition* (2.3.5) *ensures that the Lévy–Khintchine integral is well-defined. Indeed, for* $|x| \geq 1$ *it guarantees*

$$\int_{|x|\geq 1} |1 - e^{i\theta x}|\, \Pi(dx) \leq 2 \int_{|x|\geq 1} \Pi(dx) < \infty,$$

while for $|x| < 1$ *one uses the Taylor expansion* $e^{i\theta x} = 1 + i\theta x - \frac{1}{2}\theta^2 x^2 + o(x^2)$, *so that*

$$1 - e^{i\theta x} + i\theta x = \mathcal{O}(x^2) \quad \text{as } x \to 0.$$

Thus $\int_{|x|<1} x^2\, \Pi(dx) < \infty$ *ensures integrability near zero.*

Definition 2.3.13 (Lévy measure of infinite divisible random variable). *The measure* Π *is called the Lévy measure.*

Lévy processes are strongly connected with the notion of infinite divisibility. For every $t \geq 0$, the random variable X_t is infinitely divisible. Conversely, for every infinitely divisible distribution there exists a Lévy process such that X_1 has that law.

To see the former claim, note that for every $n \geq 1$

$$X_t = (X_{t/n} - X_0) + (X_{2t/n} - X_{t/n}) + \cdots + (X_t - X_{(n-1)t/n}), \tag{2.3.6}$$

which, by the independent and stationary increments of a Lévy process, is a sum of n i.i.d. random variables, each distributed as $X_{t/n}$. This observation reinforces the idea that *Lévy processes are the continuous-time analogue of sums of i.i.d. random variables*, a perspective that we hinted at earlier and will revisit later.

Let us now consider the characteristic exponent of a Lévy process. It holds that

$$\Psi_t(\theta) = t\Psi_1(\theta), \qquad t > 0, \tag{2.3.7}$$

where Ψ_t denotes the characteristic exponent of X_t, that is,

$$\Psi_t(\theta) = \log \mathbb{E}\big[e^{i\theta X_t}\big] .$$

Relationship (2.3.7) follows by applying (2.3.6) for rational values of t and then extending it to all $t > 0$ by almost sure right-continuity of $(X_t)_{t\geq 0}$ (which in turn implies right-continuity of $g(t) := \mathbb{E}\big[e^{i\theta X_t}\big]$ by the Dominated Convergence Theorem). Equation (2.3.7) shows that Ψ_1 completely characterizes the distribution of the Lévy process $(X_t)_{t\geq 0}$.

Definition 2.3.14 (Characteristic exponent of a Lévy process)**.** *Let* $(X_t)_{t\geq 0}$ *be a Lévy process. We define* $\Psi := \Psi_1$ *and call it the* characteristic exponent *of* $(X_t)_{t\geq 0}$.

Since X_1 is infinitely divisible, its characteristic exponent admits the representation given in Theorem 2.3.11. Consequently, the Lévy process $(X_t)_{t\geq 0}$ is associated with a triplet (a, σ^2, Π).

Definition 2.3.15 (Lévy measure of a Lévy process)**.** *Let* $(X_t)_{t\geq 0}$ *be a Lévy process with Lévy triplet* (a, σ^2, Π)*. The measure* Π *is called the* Lévy measure *of* $(X_t)_{t\geq 0}$.

The converse statement mentioned earlier—namely that for every infinitely divisible distribution there exists a Lévy process such that X_1 has that law—is justified by the Lévy–Khintchine formula stated above together with the following result (see [34, p. 5]).

Theorem 2.3.16 (Lévy–Khintchine formula for Lévy processes)**.** *Suppose that* $a \in \mathbb{R}$*,* $\sigma \geq 0$*, and* Π *is a measure concentrated on* $\mathbb{R} \setminus \{0\}$ *such that*

$$\int_{\mathbb{R}} (1 \wedge x^2)\, \Pi(dx) < \infty.$$

Define, for each $\theta \in \mathbb{R}$*,*

$$\Psi(\theta) = ia\theta + \frac{1}{2}\sigma^2\theta^2 + \int_{\mathbb{R}} \big(1 - e^{i\theta x} + i\theta x \mathbf{1}_{\{|x|<1\}}\big)\, \Pi(dx).$$

Then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ *on which a Lévy process is defined having characteristic exponent* Ψ*.*

2.3.3 Lévy–Itô decomposition

Let us consider a Lévy process $(X_t)_{t\geq 0}$ with Lévy triplet (a, σ^2, Π), for some $a \in \mathbb{R}$, $\sigma \geq 0$, and a Lévy measure Π satisfying (2.3.5). Then the characteristic exponent Ψ of the process, as we discussed earlier, has the form (2.3.4). With some rearranging it is easy to see that $\Psi(\theta)$ can be written as

$$\Psi(\theta) = \underbrace{\left\{ ia\theta + \frac{1}{2}\sigma^2\theta^2 \right\}}_{\Psi^{(1)}(\theta)}$$
$$+ \underbrace{\left\{ \Pi\big(\mathbb{R}\setminus(-1,1)\big) \int_{|x|\geq 1} (1 - e^{i\theta x}) \frac{\Pi(dx)}{\Pi\big(\mathbb{R}\setminus(-1,1)\big)} \right\}}_{\Psi^{(2)}(\theta)}$$
$$+ \underbrace{\left\{ \int_{0<|x|<1} \big(1 - e^{i\theta x} + i\theta x\big)\, \Pi(dx) \right\}}_{\Psi^{(3)}(\theta)}.$$

Since the sum of independent random variables has characteristic exponent equal to the sum of their characteristic exponents, the above decomposition suggests that the Lévy process $(X_t)_{t\geq 0}$ with triplet (a, σ^2, Π) can be written as the sum of three independent processes with characteristic exponents $\Psi^{(1)}(\theta)$, $\Psi^{(2)}(\theta)$ and $\Psi^{(3)}(\theta)$. The processes corresponding to $\Psi^{(1)}(\theta)$ and $\Psi^{(2)}(\theta)$ can be easily identified. The first corresponds to a Brownian motion with drift,

$$X_t^{(1)} = \sigma B_t - at,$$

while the second corresponds to a compound Poisson process with rate $\Pi(\mathbb{R}\setminus(-1,1))$ and jumps $(\xi_n)_{n\geq 0}$ with common distribution

$$\frac{\Pi(dx)}{\Pi(\mathbb{R}\setminus(-1,1))}.$$

It turns out that the last summand corresponds to another Lévy process, a square-integrable martingale with an almost surely countable number of path discontinuities (or jumps) on each finite time interval, which are of magnitude less than unity. Identifying this component rigorously is the core of proving the Lévy–Itô decomposition, which is due to Lévy and Itô [32]. A justification of this claim requires the introduction of Poisson random measures and is quite lengthy, and therefore outside the scope of this exposition. For a detailed proof see [34, Ch. 2, Sec. 1–5], and for a rigorous formulation of the theorem see [34, p. 37].

Consequently, we obtain the Lévy–Itô decomposition. That is, every Lévy process

$(X_t)_{t\geq 0}$ with triplet (a, σ^2, Π) can be written in the form

$$X_t = \sigma B_t - at + Z_t + M_t. \tag{2.3.8}$$

Here Z_t is a compound Poisson process with rate $\Pi(\mathbb{R} \setminus (-1,1))$ and jumps $(\xi_n)_{n\geq 0}$ with common distribution $\Pi(dx)/\Pi(\mathbb{R}\setminus(-1,1))$, and M_t is a square-integrable martingale with an almost surely countable number of jumps on each finite time interval, each of magnitude less than one.

The Lévy–Itô decomposition allows one to split Lévy processes into different classes by selecting components, for example Brownian motion with drift, the sum of a Brownian motion and a compound Poisson process, and so on.

2.3.4 The Lévy measure and its role

The Lévy measure is central both for understanding the path structure of a Lévy process - especially in relation to its jumps/discontinuities- and for describing the tails of its distribution.

The Lévy measure encodes the jump behaviour of the process. In particular, for a Borel set A bounded away from zero, the number of jumps with sizes in A during a time interval of length t is a Poisson random variable with mean $t\Pi(A)$ [34, p. 67]. Consequently, a Lévy process has almost surely finitely many jumps of size larger than $\epsilon > 0$ on any compact time interval. Moreover, the process has continuous paths if and only if $\Pi = 0$, that is, when there are no jumps.

On the other hand, if $\Pi((-1,1)) = \infty$, the process has infinitely many small jumps (i.e. jumps of size less than ϵ for any fixed $\epsilon > 0$) on every interval, and the jump times are dense [34, p. 67]. These infinitely many small jumps arise from the component M_t in the Lévy–Itô decomposition (2.3.8), namely the square-integrable martingale with jumps of magnitude less than one.

As an example of such a process, consider a strictly α-stable Lévy process with $\alpha \in (0,2)$ that we encountered in Example 2.3.8. The Lévy measure of such a process has the form

$$\Pi(dx) = \begin{cases} c_1 x^{-1-\alpha}\, dx, & x \in (0,\infty), \\ c_2 |x|^{-1-\alpha}\, dx, & x \in (-\infty, 0), \end{cases} \tag{2.3.9}$$

for constants $c_1, c_2 \geq 0$; see [34, p. 11]. In particular, $\Pi((-1,1)) = \infty$, and therefore the process has infinitely many small jumps on every interval $(0,t)$, with jump times that are dense.

Another important condition involving the Lévy measure and the jump behaviour is

$$\int_{|x|<1} |x| \, \Pi(dx) < \infty. \tag{2.3.10}$$

The component of small jumps, and in particular M_t, has paths of bounded variation if and only if (2.3.10) holds. Note that the compound Poisson component Z_t always has paths of bounded variation, since it is a pure jump process with almost surely finitely many jumps on every interval $(0,t]$, with the number of jumps being Poisson distributed. On the other hand, the Brownian component always has paths of unbounded variation. Consequently, a Lévy process has paths of bounded variation if and only if (2.3.10) holds and $\sigma = 0$ ([34, p. 57]).

From what we have said, the behaviour of Π near zero controls the frequency of small jumps, while its tail behaviour describes the distribution of large jumps. It is therefore not surprising that the Lévy measure is closely connected to the tail behaviour of the process itself. One immediate connection appears through exponential moments.

Theorem 2.3.17 ([34, p. 80])**.** *Let $\theta \in \mathbb{R}$. Then*

$$\mathbb{E}\big[e^{\theta X_t}\big] < \infty \quad \text{for all } t \geq 0$$

if and only if

$$\int_{|x|\geq 1} e^{\theta x} \, \Pi(dx) < \infty.$$

Recall also Definition 2.2.10 of heavy-tailed random variables. By Theorem 2.3.17, it follows that the heavy-tailedness of X_t is determined entirely by the tail of the Lévy measure.

In some cases the connection between the Lévy measure and the tail of the distribution can be made much more precise. An important example is the class of *convolution equivalent distributions*, denoted by $\mathcal{S}(\alpha)$ for $\alpha \geq 0$; see Definition 2.2.19. We state the following result as given in [1].

Theorem 2.3.18. *Let $(X_t)_{t\geq 0}$ be a Lévy process with Lévy measure Π, and let $\alpha \geq 0$ be* *a constant. Then*

$$\frac{\Pi\big([1,\infty)\cap\cdot\big)}{\Pi\big([1,\infty)\big)} \in \mathcal{S}(\alpha) \quad \Longleftrightarrow \quad X_t \in \mathcal{S}(\alpha) \;\; \textit{for some (hence all) } t>0.$$

Moreover, if these conditions hold, then

$$\lim_{x\to\infty} \frac{\mathbb{P}(X_t > x)}{\Pi\big((x,\infty)\big)} = t\,\mathbb{E}\big[e^{\alpha X_1}\big]\,, \qquad t>0.$$

Convolution equivalent distributions have the property that if $X_t \in \mathcal{S}(\alpha)$ for some $t>0$, then $X_t \in \mathcal{S}(\alpha)$ for every $t>0$; see [23] and [42]. This explains the formulation “for some (hence all) $t>0$” in the statement of Theorem 2.3.18.

Strictly stable distributions constitute an example of processes to which the above theorem applies. In Example 2.3.8, where we introduced strictly α-stable Lévy processes, we observed that for $\alpha \in (0,2)$ their distributions belong to the class $\mathcal{S}(0)$. Therefore, the theorem applies to the corresponding strictly α-stable Lévy processes and yields the following asymptotic relation:

$$\mathbb{P}(X_t > x) \sim t\,\Pi\big((x,\infty)\big), \qquad x\to\infty.$$

2.4 Compound Renewal Processes

In this section we introduce compound renewal processes. Our exposition is based on [11] and [2]. Compound renewal processes form a natural generalization of the compound Poisson processes introduced in Example 2.3.4. To this end, we first introduce renewal processes, which generalise the Poisson process.

Definition 2.4.1 (Renewal process). *Let $(W_i)_{i\geq 1}$ be a sequence of i.i.d. non-negative random variables, called* waiting times *(also called interarrival times). Define the* renewal epochs *(or simply renewals) by*

$$T_n := \sum_{i=1}^{n} W_i, \qquad n\geq 1, \quad T_0 := 0,$$

and the associated counting process

$$N_t := \max\{n\geq 0 : T_n \leq t\}, \qquad t\geq 0.$$

The process $(N_t)_{t\geq 0}$ is called a renewal counting process.

The sequence $(T_n)_{n\geq 0}$ represents the times at which events occur (e.g. machine failures), while $(W_n)_{n\geq 0}$ are the waiting times between successive events. The counting process $(N_t)_{t\geq 0}$ records the number of events that have occurred up to time t.

The terminology originates from reliability theory: after a failure at time T_n, the system is replaced (or renewed), and the time until the next failure is given by W_{n+1} [11].

Remark 2.4.2. *It is immediate from the definition that renewal processes "start afresh" at each renewal epoch. Conditioned on T_n, the shifted process $\big(N_{T_n+t} - N_{T_n}\big)_{t\geq 0}$ is independent of the past and has the same distribution as $(N_t)_{t\geq 0}$. This regeneration property also carries over to compound renewal processes and will play a key role in Chapter 4.*

Definition 2.4.3 (Compound renewal process)**.** *Let $(\xi_i)_{i\geq 1}$ be a sequence of i.i.d. random variables, called* jumps*. The process*

$$X_t := \sum_{i=1}^{N_t} \xi_i, \qquad t \geq 0,$$

is called a compound renewal process*.*

Remark 2.4.4. *In general, the waiting times (W_i) and the jumps (ξ_i) need not be independent. Throughout this thesis, however, we will always assume independence, as this is the natural analogue of the compound Poisson case and is essential for our analysis.*

Compound Poisson processes form a special case of compound renewal processes, arising when the waiting times W_i are exponentially distributed. In that case, $(N_t)_{t\geq 0}$ is a Poisson process and $(X_t)_{t\geq 0}$ reduces to a compound Poisson process, which is a Lévy process.

In general, however, a compound renewal process is *not* a Lévy process: its increments are neither stationary nor independent. Moreover, it is not even a Markov process, since the distribution of future increments depends not only on the present state but also on the time elapsed since the last renewal.

2.5 First passage time and creeping

Definition 2.5.1 (First passage time)**.** *Let* $(X_t)_{t\geq 0}$ *be a stochastic process with real-valued paths and let* $x > 0$*. The* first passage time above level x *is defined by*

$$T_x := \inf\{t > 0 : X_t > x\},$$

with the convention that $\inf \varnothing = \infty$*.*

The random variable T_x is also called the *hitting time* of the level x. Equivalently, it is the first time at which the process enters the set (x, ∞).

We now introduce the notion of creeping (see [34, p. 219]).

Definition 2.5.2 (Creeping)**.** *We say that the process* $(X_t)_{t\geq 0}$ creeps upwards over the level x *if*

$$\mathbb{P}(X_{T_x} = x) > 0.$$

We say that $(X_t)_{t\geq 0}$ creeps upwards *almost surely* if

$$\mathbb{P}(X_{T_x} = x) = 1.$$

A Lévy process with almost surely continuous sample paths creeps upwards almost surely on $\{T_x < \infty\}$ (or $\{T_x < t\}$ for fixed $t > 0$). In particular, Brownian motion is an example, since its paths are almost surely continuous and therefore the level x must be attained continuously on $\{T_x < \infty\}$[2].

Another example is a compound Poisson process with only negative jumps and strictly positive linear drift. In this case, the process can only cross a level by continuous deterministic drift, and therefore creeping occurs almost surely on $\{T_x < \infty\}$.

In contrast, a compound Poisson process with both positive and negative jumps and positive linear drift may creep upwards, but not almost surely: the level can be reached either continuously by the drift or by a jump that overshoots it.

In this thesis, we will primarily be interested in creeping on the event $\{T_x < t\}$ for a fixed $t > 0$.

[2] Actually, Brownian motion creeps upwards unconditionally since $\limsup_{t\to\infty} B_t = \infty$ (see [37, Ch. 5, Sec. 1])

2.6 Regeneration at first passage times

We introduce here a property that plays a central role in our analysis and is satisfied by Lévy and compound renewal processes.

We will make use of the notions of filtrations and stopping times; for a comprehensive treatment see [33, Ch. 1] or [20, Ch. 7]. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(X_t)_{t\geq 0}$ be a stochastic process adapted to a filtration $(\mathcal{F}_t)_{t\geq 0}$.

Definition 2.6.1 (Stopping time)**.** *A random time* $T : \Omega \to [0, \infty]$ *is called a* stopping time *with respect to the filtration* $(\mathcal{F}_t)$ *if, for every* $t \geq 0$*,*

$$\{T \leq t\} \in \mathcal{F}_t.$$

Given a stopping time T, we define the σ-algebra at time T by

$$\mathcal{F}_T := \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t \text{ for all } t \geq 0\}.$$

Intuitively, $\mathcal{F}_t$ represents the information available up to time t, while $\mathcal{F}_T$ represents the information available up to the random time T. Moreover, a stopping time is a random time whose occurrence can be determined using only the information available up to that time, without any knowledge of the future.

The first passage times defined in Section 2.5 are stopping times (see, e.g., [34, p. 73]). Indeed, at any time t, whether the event $\{T_x \leq t\}$ has occurred can be determined solely from the information available up to time t, since this event is equivalent to $\{\sup_{0\leq s\leq t} X_s > x\}$.

Lévy processes satisfy the strong Markov property; see [20, Ch. 7, Sec. 3] for a formal definition. Informally, a process satisfies the strong Markov property if, given the information up to a stopping time T, the future evolution of the process depends only on the present state X_T. More precisely, the process

$$\big(X_{T+t} - X_T\big)_{t\geq 0}$$

is independent of the past conditional on X_T, and has the same law as the original process started from 0.

Lévy processes satisfy an even stronger property: the future evolution of the process after a stopping time is independent of the entire past, not only conditionally on the present. In particular, the post-T process is independent of $\mathcal{F}_T$ and has the same law as the original process. The following result, taken from [34, Ch. 3, Sec. 1], makes this precise.

Theorem 2.6.2. *Let T be a stopping time and define, on the event $\{T < \infty\}$, the process $\tilde{X} = (\tilde{X}_t)_{t\geq 0}$ by*

$$\tilde{X}_t := X_{T+t} - X_T, \qquad t \geq 0.$$

Then, on the event $\{T < \infty\}$, the process $\tilde{X}$ is independent of $\mathcal{F}_T$ and has the same law as the original process X. In particular, $\tilde{X}$ is a Lévy process.

For our purposes, we restrict this stronger form of the Strong Markov property to a specific subset of stopping times, namely the first passage times.

Definition 2.6.3 (Regenerative at first passage times)**.** *We say that the process $(X_t)_{t\geq 0}$ is* regenerative at first passage times *if, for every x such that $T_x < \infty$ almost surely, the process $\tilde{X} = (\tilde{X}_t)_{t\geq 0}$ defined by*

$$\tilde{X}_t := X_{T_x+t} - X_{T_x}, \qquad t \geq 0,$$

is independent of $\mathcal{F}_{T_x}$ and has the same law as the original process $(X_t)_{t\geq 0}$.

This property expresses that, at each first passage time, the process $(X_t)_{t\geq 0}$ restarts afresh, in the sense that its future increments are independent of the past and distributed as a new copy of the process. Both Lévy processes and compound renewal processes satisfy this property, and it will form the basis of several arguments developed later.

Remark 2.6.4. *Lévy processes are regenerative at first passage times by Theorem 2.6.2. Compound renewal processes are also regenerative, since they renew at jump times, and every first passage time T_x necessarily coincides with a jump time.*

In particular, for both classes of processes, for any measurable set B and $t \geq s$, we have

$$\mathbb{P}(X_t - X_{T_x} \in B \mid T_x = s) = \mathbb{P}(X_{t-s} \in B)\,.$$

Remark 2.6.5. *As an example of a process that is not regenerative at first passage times, consider a compound renewal process with positive linear drift, i.e.,*

$$X_t = Y_t + bt, \qquad b > 0,$$

where $(Y_t)_{t \geq 0}$ *is a compound renewal process. In this case, the process can exceed levels continuously due to the drift, rather than only at jump times at which regeneration occurs. Consequently, the process is not regenerative at first passage times.*

Chapter 3

Suprema of Lévy processes

3.1 Processes Tail Equivalent to Their Supremum

Let $(X_t)_{t\geq 0}$ be a Lévy process and denote its supremum process by

$$\overline{X}_t := \sup_{0\leq s\leq t} X_s,$$

for $t > 0$.

The central question of this chapter is: under what assumptions does there exist a constant $c \in (0, \infty)$ such that

$$\mathbb{P}\big(\overline{X}_t > x\big) \;\sim\; c\,\mathbb{P}(X_t > x) \qquad \text{as } x \to \infty. \tag{3.1.1}$$

Motivation

In applications one often requires control of $\mathbb{P}\big(\overline{X}_t > x\big)$—for instance in risk theory (ruin probabilities), finance (barrier options), or queueing theory (buffer overflows). However, the distribution of $\overline{X}_t$ is rarely tractable.

If (3.1.1) holds, the probability $\mathbb{P}\big(\overline{X}_t > x\big)$ for large x can be estimated directly through the tail of X_t, which is typically more accessible. Even when the marginal distribution function of X_t is not available in closed form, one often has knowledge of its characteristic function. In that case, Tauberian theorems can be employed to recover tail asymptotics from Laplace transforms. (see, e.g., [29] for an overview of Esscher transforms, and [25] for an interesting variation).

Definition

Given the central role of (3.1.1), we introduce a convenient terminology.

Definition 3.1.1 (Tail equivalence to the supremum)**.** *Fix $t > 0$. We say that $(X_t)_{t\geq 0}$ is* tail equivalent to its supremum at time *t with constant $c \in (0, \infty)$ if* (3.1.1) *holds.*

Types of assumptions leading to (3.1.1)

There are two natural ways of formulating assumptions that imply (3.1.1). They are not mutually exclusive and are often used in combination.

(i) **Structural assumptions on the process.** These are imposed at the level of the Lévy–Itô decomposition (see Subsection 2.3.3). For example, one may assume that $(X_t)_{t\geq 0}$ is a compound Poisson process with linear drift (as in [14]), a linear Brownian motion, or the sum of the two (as in [15]).

(ii) **Tail assumptions.** These concern the heaviness of the tails and can be stated either:

- at the level of the marginal distribution of X_t (e.g. X_t is long-tailed as in [44]), or
- at the level of the Lévy measure Π (e.g. Π has exponential tails as in [1] or the jumps of a compound Poisson process are Braverman-light, as in [14]).

3.2 Existing Results

In this section we briefly present the most basic results on asymptotic equivalence to the supremum.

3.2.1 Brownian Motion

The first and most classical result of this type concerns standard Brownian motion. In this case the analogue of (3.1.1) holds not merely asymptotically, but as an exact identity valid for all $x > 0$.

Theorem 3.2.1 ([41] Reflection Principle). *Let $(B_t)_{t\geq 0}$ be a Brownian motion with $B_0 = 0$. Then, for $x > 0$,*

$$\mathbb{P}\big(\overline{B}_t > x\big) = 2\,\mathbb{P}(B_t > x)\,, \qquad \text{for every } t > 0.$$

A natural variant is Brownian motion with linear drift. Here the exact equality is lost, but the asymptotic analogue of (3.1.1) survives with the same constant:

Fact 3.2.2. *Let $X_t := B_t + bt$, where $b \in \mathbb{R}$. Then, as $x \to \infty$,*

$$\mathbb{P}\big(\overline{X}_t > x\big) \;\sim\; 2\,\mathbb{P}(X_t > x)\,. \tag{3.2.1}$$

This result can be established by applying the Girsanov theorem to remove the drift, then invoking the Reflection Principle, and finally reversing the change of measure. Later in this chapter we will give an alternative elementary proof that avoids Girsanov's theorem and extends to the scaled case $X_t = \sigma B_t + bt$ with $\sigma > 0$ and $b \in \mathbb{R}$ (see Fact 3.4.21).

3.2.2 Processes with Heavy Tails

The heavy-tailed setting was first studied in [7], where the author considered Lévy processes with symmetric increments and a regularly varying Lévy measure, using the analytical method of sojourn time analysis.

Theorem 3.2.3 ([7]). *Let $(X_t)_{t\geq 0}$ be a Lévy process with Lévy measure Π. Assume that $\Pi\,(x, \infty)$ is regularly varying with index $-\alpha$ for some $\alpha \in (0, 2)$ as $x \to \infty$. Then, for every $t > 0$,*

$$\mathbb{P}\big(\overline{X}_t > x\big) \;\sim\; \mathbb{P}(X_t > x)\,, \qquad x \to \infty.$$

Although regular variation is imposed on the Lévy measure, this is in fact equivalent to requiring that the distribution of X_t is regularly varying (recall that regular varying distributions are convolution equivalent and see Theorem 2.3.18).

A major simplification was later provided in [44], who gave a short and intuitive proof under substantially weaker assumptions, replacing regular variation with long-tailedness.

Theorem 3.2.4 ([44]). *Let $(X_t)_{t\geq 0}$ be a Lévy process. Fix $t > 0$. Then the following are equivalent:*

(i) X_t *is long-tailed,*

(ii) $\overline{X}_t$ *is long-tailed,*

and each of (i) and (ii) implies

$$\mathbb{P}\big(\overline{X}_t > x\big) \sim \mathbb{P}(X_t > x)\,, \qquad x \to \infty.$$

The proof in [44] is elementary: it relies only on tools like the law of total probability and the definitions of long-tailedness and Lévy processes, thereby exposing the essence of why the equivalence holds. Our own approach is closer in spirit to this line of reasoning than to the more analytical methods of [7] and [14].

3.2.3 Processes with Light Tails

Michael Braverman has written numerous papers on this problem (eg.[12, 13, 14, 15, 16]), and his work plays a big role in our approach. In particular, we will use one of his lemmas and his definition of light tails in the next chapter on compound Poisson processes.

Braverman introduced his own notion of light tails, which we refer to as *Braverman-light tails* (see Definition 2.2.3). This notion was presented earlier in Subsection 2.2.1, where we also discussed that it is strictly stronger than the standard notion of light-tailedness.

Theorem 3.2.5 ([14])**.** *Let* $(X_t)_{t\geq 0}$ *be a compound Poisson process with linear drift,*

$$X_t = \sum_{k=1}^{N_t} \xi_k + bt,$$

where b *is a constant,* $(N_t)_{t\geq 0}$ *is a Poisson process with rate* λ*, and* $(\xi_k)_{k\geq 1}$ *are i.i.d. random variables independent of* $(N_t)_{t\geq 0}$*. Let* $(\Gamma_n)_{n\geq 1}$ *be the jump times of* $(N_t)_{t\geq 0}$*, and set*

$$T := \max\{n : \Gamma_n \leq 1\}.$$

Suppose the distribution of ξ_k *has a Braverman-light right tail. Then:*

$$\mathbb{P}\big(\overline{X}_1 > x\big) \sim \begin{cases} \mathbb{P}(X_{\Gamma_T} > x)\,, & b \leq 0, \\ \\ \mathbb{P}(X_1 > x)\,, & b \geq 0, \end{cases} \qquad \text{as } x \to \infty.$$

Remark 3.2.6. *Note that Braverman states his results for* $t = 1$*. This is not a genuine restriction, since the proofs do not rely on the specific choice* $t = 1$ *and can be extended* *to arbitrary $t > 0$. However, to the best of our knowledge, such an extension cannot be deduced immediately from the above statement for the case $t = 1$.*

This theorem highlights a crucial distinction between positive and negative drift. In the negative drift case, the tail of the supremum is asymptotically equivalent not to that of X_1, but rather to that of the process evaluated at the time of the last jump before $t = 1$.

We later present an alternative, more elementary proof of this result, which in fact yields a slight generalization. Moreover, our main result, Theorem 4.1.9, extends this conclusion from the compound Poisson setting to compound renewal processes.

Braverman further showed in [14] that the distinction between positive and negative linear drift is genuine: there exist compound Poisson processes with Braverman-light jumps and negative drift for which $\overline{X}_1$ and X_1 are not asymptotically equivalent. Such examples arise, for instance, when the jump distribution is bounded and supported on a lattice.

The same paper also provides sufficient conditions on the jump distribution that ensure asymptotic equivalence even in the negative drift case. For instance, the equivalence holds when the jumps are exponentially distributed.

In subsequent work [15], Braverman introduced another such condition, which is satisfied by the Gaussian distribution. Consequently, compound Poisson processes with negative drift and Gaussian jumps also exhibit this asymptotic equivalence.

The most general light-tail result is again due to Braverman [15], who considers the Lévy process

$$X_t = \tilde{X}_t + \sigma B_t + bt, \tag{3.2.2}$$

where $\tilde{X}_t$ is a pure-jump Lévy process of bounded variation, that is,

$$\int_{-1}^{1} |x| \, \Pi(dx) < \infty. \tag{3.2.3}$$

We discussed this condition earlier in Subsection 2.3.4.

If the Lévy measure has finite total mass, then $\tilde{X}_t$ reduces to a compound Poisson process, and (3.2.2) becomes

$$X_t = Z_t + \sigma B_t + bt, \qquad Z_t = \sum_{k=1}^{N_t} \xi_k,$$

in which case the Braverman-light-tailed condition is imposed on the jump distribution of Z_t.

If, on the other hand, the Lévy measure has infinite total mass, then the pure-jump component admits the decomposition

$$\tilde{X}_t = M_t + Z_t,$$

where M_t is a square-integrable martingale with jumps of magnitude less than one, and Z_t is a compound Poisson process (see Subsection 2.3.3). In this case, the Braverman-light-tailed condition is imposed on the distribution

$$F_\Pi(x) := 1 - \frac{\Pi\big((\max\{x, \alpha\}, \infty)\big)}{\Pi\big((\alpha, \infty)\big)}, \tag{3.2.4}$$

for some $\alpha > 0$, provided that $\Pi((\alpha, \infty)) > 0$.

Theorem 3.2.7 ([15])**.** *Let $(X_t)_{t \geq 0}$ be a Lévy process with Lévy measure Π. Assume that*

(i) $\Pi((0, \infty)) > 0$,

(ii) There exists $a > 0$ such that F_Π is well defined and has Braverman-light tails,

(iii) the bounded-variation condition (3.2.3) *holds, and*

(iv) $\sigma > 0$.

Then

$$\mathbb{P}\big(\overline{X}_1 > x\big) \;\sim\; \mathbb{P}(X_1 > x)\,, \qquad x \to \infty.$$

Note that while compound Poisson processes with negative drift may fail to satisfy the asymptotic equivalence (see the discussion following Theorem 3.2.5), the addition of a Brownian component restores the equivalence under the condition of Braverman-light jumps.

Nevertheless, Braverman also constructed in [15] an example of a process given by the sum of a compound Poisson process—whose jumps are not Braverman-light—and an independent Brownian motion, for which the equivalence fails. This demonstrates that his condition is essential.

The proof of Theorem 3.2.7 is quite lengthy and technical. A closer inspection of Braverman's argument, however, reveals the underlying mechanism by which such processes attain large values.

For instance, consider $X_t = B_t + N_t$, where $(N_t)_{t\geq 0}$ is a Poisson process and $(B_t)_{t\geq 0}$ is standard Brownian motion. The most likely number of jumps when the process exceeds a large level x is approximately $x - \log x$, with the Brownian component contributing $B_1 > \log x$.

Since no single mechanism dominates—as is often the case in the heavy-tailed setting—a careful analysis of multiple competing regimes is required.

We explored whether Theorem 3.2.7 could be proved using our more elementary and intuitive approach. As we discuss later, these attempts were not successful, reflecting both the additional complexity introduced by the Brownian component and the limitations of our method in certain cases.

3.2.4 Processes with Exponential Tails

Another important contribution is due to [1], who studied Lévy processes with exponential tails, that is, processes in the class $\mathcal{L}(\alpha)$ for some $\alpha \geq 0$. We have already presented and discussed this class of distributions in Subsection 2.2.3.

In that work, the asymptotic equivalence (3.1.1) was established for several specific Lévy processes frequently used in practice (e.g. in finance). The main result of the paper is the following.

Theorem 3.2.8 ([1]). *Let $\alpha \geq 0$ and let $(X_t)_{t\geq 0}$ be a Lévy process. Suppose that for some $t > 0$ and every $s \in (0, t)$ the limit*

$$L(s) := \lim_{x\to\infty} \frac{\mathbb{P}(X_s > x)}{\mathbb{P}(X_t > x)}$$

exists. Then

$$X_t \in \mathcal{L}(\alpha) \quad \iff \quad \overline{X}_t \in \mathcal{L}(\alpha).$$

Moreover, if either condition holds, then both hold, and

$$\mathbb{P}\big(\overline{X}_t > x\big) \ \sim\ c\,\mathbb{P}(X_t > x)\,, \qquad x \to \infty.$$

In that case, we have $c = 1$, if $L(s) = 0$ for all $s \in (0,t)$.

Recall that $\mathcal{L}(0) = \mathcal{L}$ and therefore when $\alpha = 0$, the assumption on the existence of $L(s)$ is not needed: Willekens' theorem (Theorem 3.2.4) already guarantees the equivalence. Later in this chapter we will present a partial alternative proof of Theorem 3.2.8 using our method. Our argument covers precisely the case $L(s) = 0$ for all $s \in (0,t)$ (see Theorem 3.4.19).

We also regard the following as particularly interesting as it highlights the central role of the ratio $L(s)$ in understanding tail equivalence to the supremum.

Theorem 3.2.9 ([1])**.** *Let $(X_t)_{t\geq 0}$ be a Lévy process that satisfies (3.1.1), but is not a subordinator. Then one of the following holds:*

(i) $X_t \in \mathcal{L}$;

(ii) $\liminf_{x\to\infty} \dfrac{\mathbb{P}(X_s > x)}{\mathbb{P}(X_t > x)} = 0 \qquad$ *for all $s \in (0,t)$.*

Remark 3.2.10. *Theorem 3.2.9 is of particular interest to us, as it highlights the condition around which our own method for proving asymptotic equivalence is built. As we will show later, a strengthened form of condition (ii), namely*

$$\lim_{x\to\infty} \frac{\mathbb{P}(X_s > x)}{\mathbb{P}(X_t > x)} = 0, \qquad \textit{for all } s \in (0,t), \tag{3.2.5}$$

implies that the process must rely on the "synergy of all its increments" in order to exceed large levels. Viewed in this way, Theorem 3.2.9 can be seen as a manifestation of the classical dichotomy between the catastrophe principle *(extremes caused by a single large jump, typical of heavy tails) and the* conspiracy principle *(extremes caused by the cooperation of all increments, typical of light tails). Unlike the random walk case, however, these conditions are not mutually exclusive: there exist processes—such as geometric Brownian motion—that satisfy both.*

Remark 3.2.11. *An analogous statement holds for random walks. We omit the proof, as it would essentially replicate the argument of Theorem 3.2.9. In particular, let $\{S_n\}$*

be a random walk with increments of infinite support (otherwise asymptotic equivalence is not meaningful). Then, for

$$\mathbb{P}\big(\overline{S}_n > x\big) \;\sim\; \mathbb{P}(S_n > x)\,, \qquad x \to \infty, \quad n \in \mathbb{N},$$

one of the following conditions must hold:

(i') $S_n \in \mathcal{L}$*;*

(ii') $\liminf_{x\to\infty} \mathbb{P}(S_k > x)\,/\mathbb{P}(S_n > x) = 0$ *for every* $k \in \{1, \dots, n-1\}$.

Condition (i') is sufficient on its own, while (ii') becomes sufficient when the limit exists.

3.3 A Heuristic Approach

3.3.1 Conditional Probability Viewpoint

A recurring theme in this thesis is that tail equivalence with the supremum can be naturally understood via the conditional probability

$$\mathbb{P}(X_t > x \,|\, T_x \le t)\,,$$

where $T_x := \inf\{s > 0 : X_s > x\}$.

Since $\{\overline{X}_t > x\} = \{T_x \le t\}$, we have

$$\frac{\mathbb{P}(X_t > x)}{\mathbb{P}\big(\overline{X}_t > x\big)} = \frac{\mathbb{P}(X_t > x)}{\mathbb{P}(T_x \le t)}.$$

Hence, by Bayes' formula,

$$\frac{\mathbb{P}(X_t > x)}{\mathbb{P}\big(\overline{X}_t > x\big)} = \mathbb{P}(X_t > x \,|\, T_x \le t)\,.$$

Therefore, $(X_t)_{t\ge 0}$ is tail-equivalent to its supremum with constant c (see Definition 3.1.1) if and only if

$$\mathbb{P}(X_t > x \,|\, T_x \le t) \;\sim\; \frac{1}{c}, \qquad x \to \infty. \tag{3.3.1}$$

In words, the constant $1/c$ has a natural sample–path interpretation: it represents the asymptotic fraction of paths that, having crossed a very high level x before time t, are still above x at time t. Thus the problem reduces to determining the probability of remaining above a large level once it has been reached.

Seen this way, the phenomenon of tail equivalence is less about delicate analytical manipulations and more about structural features of the underlying process. These include:

- the mechanism by which the process attains large values, namely either through a single big jump or by reaching the level at the very last moment before time t using all of its increments (catastrophe versus synergy principle; recall the discussion in Subsection 2.2.2 and Remark 3.2.10),
- regeneration at first passage times, stemming from independent and stationary increments (recall Definition 2.6.3),
- the small-time behaviour near the origin (e.g. whether the process can immediately drop below zero),
- the manner in which the process crosses levels: either by creeping (recall Definition 2.5.2) continuously or by overshooting them.

Regarding the last point, to be more precise, when the process does not creep upwards we are interested not in the overshoot in its strict sense as it appears in the literature, but in

$$\lim_{x\to\infty} \mathbb{P}\big(\overline{X}_t > x + l \,\big|\, \overline{X}_t > x\big)\,,$$

for some $l > 0$. Namely, the asymptotic conditional probability that, given the process has surpassed x, it also exceeds $x + c$. This probability is determined by the residual life distribution (see Definition 2.2.12).

3.3.2 Interpretation of Existing Results

We now revisit the results of Section 3.2 through the conditional probability viewpoint, in order to obtain a simpler interpretation and to highlight the role of the factors discussed above. The discussion is primarily heuristic and intended to provide intuition.

Brownian motion.

The statement of Theorem 3.2.1 is equivalent to

$$\mathbb{P}(B_t > x \,|\, T_x \leq t) = \frac{1}{2}, \qquad x > 0.$$

That is, among all Brownian paths that cross level x before time t, half remain above x at time t. Heuristically, this follows from: regeneration of the process at T_x (recall Definition 2.6.3 and Remark 2.6.4), creeping of Brownian paths, and the symmetry $\mathbb{P}(B_t > 0) = 1/2$.

More explicitly: regeneration at T_x ensures that the post-hitting process $\tilde{B}_s := B_{s+T_x} - B_{T_x}$ has the same law as B_s and is independent of the past. Creeping upwards implies that the process stays above x at time t *if and only if* $\tilde{B}_{t-T_x} > 0$. Symmetry of Brownian motion then gives $\mathbb{P}\Big(\tilde{B}_{t-T_x} > 0\Big) = 1/2$.

This is the ideal case: we obtain a strict equality. As soon as we move away from Brownian motion, the result becomes less precise. For example, if the process does *not* creep upwards almost surely, then the above heuristic yields only an inequality. The possibility of overshoot allows for paths which remain above x at time t, even when the shifted process is negative at time $t - T_x$ (i.e., $\tilde{X}_{t-T_x} < 0$). This reflects the fact that, after exceeding level x, the process may decrease while still remaining above x.

In fact we illustrate this heuristic by formally proving the following:

Fact 3.3.1. *Let $(X_t)_{t\geq 0}$ be a Lévy process such that $\mathbb{P}(X_t > 0) = \rho$ for every $t > 0$. Then*

$$\mathbb{P}\big(\overline{X}_t > x\big) \;\leq\; \frac{1}{\rho}\,\mathbb{P}(X_t > x)\,, \qquad x > 0.$$

Proof. Since $\{\overline{X}_t > x\} = \{T_x \leq t\}$, Bayes' formula yields

$$\frac{\mathbb{P}(X_t > x)}{\mathbb{P}\big(\overline{X}_t > x\big)} = \mathbb{P}(X_t > x \mid T_x \leq t)\,.$$

Now, on the event $\{T_x \leq t\}$, if $X_t - X_{T_x} \geq 0$, then necessarily $X_t \geq X_{T_x} > x$. Thus

$$\{X_t - X_{T_x} \geq 0\} \subseteq \{X_t > x\} \qquad \text{on } \{T_x \leq t\},$$

which implies

$$\mathbb{P}(X_t > x \mid T_x \le t) \ge \mathbb{P}(X_t - X_{T_x} \ge 0 \,|\, T_x \le t)\,,$$

and by the law of total probability and conditioning on T_x, we obtain

$$\mathbb{P}(X_t > x \mid T_x \le t) \ge \int_0^t \mathbb{P}(X_t - X_{T_x} \ge 0 \mid T_x = s)\; \mathbb{P}(T_x \in ds \,|\, T_x \le t)\,.$$

By the regeneration property,

$$\mathbb{P}(X_t - X_{T_x} \ge 0 \mid T_x = s) = \mathbb{P}(X_{t-s} \ge 0) = \rho, \qquad 0 \le s \le t.$$

Therefore,

$$\mathbb{P}(X_t > x \mid T_x \le t) \ge \rho \int_0^t \mathbb{P}(T_x \in ds)\, T_x \le t = \rho.$$

Combining this with the previous identity gives

$$\mathbb{P}\big(\overline{X}_t > x\big) \le \frac{1}{\rho}\, \mathbb{P}(X_t > x)\,,$$

as claimed. □

Remark 3.3.2. *The above applies in particular to strictly α-stable Lévy processes (see Example 2.3.8), since $X_t \overset{d}{=} t^{1/\alpha} X_1$, and hence $\rho = \mathbb{P}(X_1 > 0)$.*

Linear Brownian Motion.

Fact 3.2.2 is equivalent to

$$\mathbb{P}(B_t + bt > x \,|\, T_x \le t) \;\sim\; \tfrac{1}{2}, \qquad x \to \infty.$$

We will prove this result rigorously later in Fact 3.4.21 using our method; for now, we provide a heuristic explanation.

In contrast to the ideal case of standard Brownian motion, the symmetry property no longer holds, since $\mathbb{P}(B_t + bt > 0) \neq 1/2$, and therefore we obtain only an *asymptotic equivalence* rather than an exact identity.

Why does the constant 1/2 still appear? This is due to the manner in which the process attains large levels and its behaviour near zero. A Brownian motion with drift reaches high levels in a time-concentrated manner: conditional on the event $\{T_x \le t\}$, the hitting time T_x occurs arbitrarily close to t as $x \to \infty$ (this is made precise in Proposition 3.4.9). After regeneration at T_x, the process only needs to be positive at the endpoint of the short interval $(0, t - T_x)$. Over such small time intervals, the Brownian component dominates the linear drift, and hence the probability of positivity converges to 1/2.

Processes with heavy tails

The statement of Willekens' Theorem 3.2.4 for long-tailed Lévy processes, is equivalent to

$$\mathbb{P}(X_t > x \,|\, T_x \le t) \;\sim\; 1, \qquad x \to \infty.$$

The probability that a path which has crossed a large level x before t remains above x at time t tends to one as $x \to \infty$. This behaviour stems from two features: the mechanism by which heavy-tailed processes reach large values, and the renewal property at hitting times. Long-tailedness, namely

$$\lim_{x\to\infty} \mathbb{P}(X_t > x + c \,|\, X_t > x) = 1 \quad \text{for all } c > 0,$$

implies the same for the supremum process (see the proof of Theorem 3.4.19). Thus, when a path attains a very large value x, it overshoots x by an arbitrarily large amount, leaving no chance to fall back below x in the residual time $t - T_x$.

It should be stressed that regeneration at first passage times is still crucial here. Without it, an elevation to a high level does not in itself guarantee retention above x. For example, in a mean-reverting process such as the Ornstein–Uhlenbeck process, the higher the process climbs the stronger the pull back towards the mean, making the retention above level x more complicated.

Compound Poisson with light jumps.

The statement of Theorem 3.2.5 for compound Poisson processes with positive linear drift is equivalent to

$$\mathbb{P}(X_t > x \,|\, T_x \le t) \;\sim\; 1, \qquad x \to \infty.$$

In this case the result follows from three key ingredients: (a) light jumps imply that the process reaches large values by synergy, so that when it crosses a large level x before

t it does so arbitrarily close to t; (b) regeneration at the first passage time T_x; (c) once regenerated, the process stays positive with probability 1 over very small time intervals.

Seen this way, the proof becomes much simpler and a more probabilistic than Braverman's original argument, which relied on features (such as the precise distribution of the interarrival times) that play no essential role and obscure the core mechanism behind the result.

Processes with exponential tails.

The tail equivalence in the statement of Theorem 3.2.8 is likewise equivalent

$$\mathbb{P}(X_t > x \,|\, T_x \le t) \;\sim\; 1, \qquad x \to \infty. \tag{3.3.2}$$

Here again, regeneration at the first passage time T_x is crucial, as is the manner in which the process attains large values: condition (3.2.5) ensures a synergistic behaviour, so that large levels are reached close to t. A further key ingredient is that crossings occur with a controllable overshoot, due to the existence of a limiting residual distribution. This permits the path, after crossing level x, to decrease while still remaining above x at time t. This heuristic is formalized in the proof of Theorem 3.4.19.

3.3.3 Limitations of our heuristic approach

Consider now Theorem 3.2.7, treating the process

$$X_t = \sigma B_t + Z_t + bt, \qquad Z_t \text{ compound Poisson.}$$

This case illustrates the limits of our method, and the significant role of the residual life distribution and small-time behaviour. A main goal was to provide a simplified proof of this theorem — as we managed in the compound Poisson case — but we were not able to.

Relative to the compound Poisson case, the addition of Brownian motion spoils the small-time behaviour: the Brownian component dominates both the drift and the jump part, making it likely that the renewed process will be negative at the end of the short interval $(0, t - T_x)$. Relative to the exponential-tail case, the absence of a proper limiting residual distribution prevents us from controlling the overshoot smoothly and hence from giving the path "room" to dip below x while still ending above it.

To further highlight the role of the limiting residual life distribution in controlling the overshoot, consider the example $X_t = B_t + N_t$, where N_t is a Poisson process. In this case[1],

$$\lim_{x\to\infty} \overline{R}_x(c) = \lim_{x\to\infty} \frac{\mathbb{P}(X_t > x + c)}{\mathbb{P}(X_t > x)} = 0, \qquad c \geq 1.$$

Thus $\overline{R}_x$ cannot converge to a proper distribution, since the exponential is the only possible proper limit (see Section 2.2.3). As a consequence, neither can the residual life distribution of the supremum converge pointwise to a proper distribution. The proof of the following fact is given in Section A.3.

Fact 3.3.3. *Let $(X_t)_{t\geq 0}$ be a Lévy process such that $\inf_{0\leq s\leq t} \mathbb{P}(X_s > 0) > 0$ and*

$$\lim_{x\to\infty} \overline{R}_x(c) = 0 \tag{3.3.3}$$

for some $c > 0$, where $\overline{R}_x$ denotes the residual life distribution of X_t. Then

$$\lim_{x\to\infty} \mathbb{P}\big(\overline{X}_t > x + c \,\big|\, \overline{X}_t > x\big) = 0.$$

In these cases, the best we can hope for in the pursuit of proving asymptotic equivalence through this type of heuristic is that X_t belongs to the maximum domain of attraction of the Gumbel distribution. In that case –see [24, Ch. 3, Sec. 3],

$$\lim_{x\to\infty} \frac{\mathbb{P}(X_t > x + cg(x))}{\mathbb{P}(X_t > x)} = e^{-c}, \qquad c > 0, \tag{3.3.4}$$

for some scaling function g.

This corresponds to a more delicate way of "giving room" to the path after hitting a large level. It is in fact possible to prove asymptotic equivalence for processes of this type under the additional assumption of a strengthened synergistic condition (3.2.5). However, we found the resulting statement too restrictive to be broadly applicable, and therefore chose to omit it.

[1]This follows since the residual life distribution of N_t satisfies $\lim_{x\to\infty} \overline{R}_x(c) = 0$ for $c \geq 1$, in conjunction with Lemma 3.1 of [15].

3.4 Synergistic at Extremes Lévy Processes

3.4.1 Definition and Properties

Condition (3.2.5) plays a central role in our analysis. For clarity and ease of reference, we restate it below as a formal definition. The intuition behind this terminology was already hinted at in our earlier discussion (see the paragraph "Catastrophe vs. Synergy" in Subsection 2.2.2 and Section 3.3) and will be made precise in Proposition 3.4.9.

Definition 3.4.1 (Synergistic at extremes)**.** *Let $(X_t)_{t\geq 0}$ be a stochastic process and fix $t > 0$. We say that $(X_t)_{t\geq 0}$ is* synergistic at extremes up to time t *if*

$$\lim_{x\to\infty} \frac{\mathbb{P}(X_s > x)}{\mathbb{P}(X_t > x)} = 0, \qquad \textit{for all } s \in (0,t). \tag{3.4.1}$$

As with other conditions imposed directly on X_t, such as long-tailedness, we cannot generally show that being synergistic at extremes for some t implies the property for all t. Nevertheless, for simplicity we usually suppress the explicit dependence on t and simply say that a process is synergistic at extremes.

Remark 3.4.2. *It is implicit in the above definition that processes which are synergistic at extremes have unbounded support, in the sense that $\mathbb{P}(X_t > x) > 0$ for all $x, t > 0$. Indeed, if this were not the case, the ratio appearing in the definition would not be well defined. This restriction is not significant for our purposes, as in such cases the notion of asymptotic equivalence itself would fail to be meaningful.*

Examples

According to Remark 3.4.2, examples of processes that are not synergistic at extremes include those with non- increasing paths. Another example is a compound Poisson process with only negative jumps and linear drift.

We now present some examples of synergistic Lévy processes. In the next section, we will prove that these examples also satisfy the asymptotic equivalence to the supremum.

Example 3.4.3 (Linear Brownian motion)**.** *Let $X_t := \sigma B_t + bt$ with $b \in \mathbb{R}$ and $\sigma > 0$.*

For $t > 0$ and $s < t$,

$$\frac{\mathbb{P}(X_s > x)}{\mathbb{P}(X_t > x)} = \frac{\overline{\Phi}\left(\frac{x-bs}{\sigma\sqrt{s}}\right)}{\overline{\Phi}\left(\frac{x-bt}{\sigma\sqrt{t}}\right)} \sim \exp\left(\frac{(x-bt)^2}{2\sigma^2 t} - \frac{(x-bs)^2}{2\sigma^2 s}\right) \to 0, \tag{3.4.2}$$

as $x \to \infty$, where $\overline{\Phi}$ and ϕ denote the standard normal tail and density, respectively, and we used $\overline{\Phi}(x) \sim \phi(x)/x$. Note that

$$\exp\left(\frac{(x-bt)^2}{2\sigma^2 t} - \frac{(x-bs)^2}{2\sigma^2 s}\right) = \exp\left(-\frac{t-s}{2\sigma^2 st}\, x^2 + \frac{b^2}{2\sigma^2}(t-s)\right),$$

which indeed vanishes as $x \to \infty$.

Example 3.4.4 (Compound Poisson with Braverman-light jumps)**.** *Another process we have already discussed that is synergistic at extremes is the compound Poisson process with Braverman-light jumps and linear drift. Condition* (3.2.5) *for the driftless case follows directly from Lemma 5 in [14]. The case with linear drift can be proved analogously.*

Example 3.4.5 (Perturbed ruin process)**.** *Let $X_t := B_t + Z_t + bt$, where $b > 0$, $Z_t = \sum_{i=1}^{N_t} \xi_i$ with $\xi_i \leq 0$ a.s., and $(B_t)_{t\geq 0}$ and $(N_t)_{t\geq 0}$ are independent (with $(N_t)_{t\geq 0}$ a Poisson process). Then $(X_t)_{t\geq 0}$ is synergistic at extremes.*

Indeed, observe first that

$$\mathbb{P}(B_t + Z_t + bt > x) \geq \mathbb{P}(B_t + bt > x,\, N_t = 0) = \mathbb{P}(W_1 > t)\, \mathbb{P}(B_t + bt > x)\,,$$

using independence of B_t and N_t. Thus,

$$\frac{\mathbb{P}(X_s > x)}{\mathbb{P}(X_t > x)} = \frac{\mathbb{P}(B_s + Z_s + bs > x)}{\mathbb{P}(B_t + Z_t + bt > x)} \leq \frac{\mathbb{P}(B_s + bs > x - Z_s)}{\mathbb{P}(W_1 > t)\, \mathbb{P}(B_t + bt > x)}.$$

Since $Z_s \leq 0$ a.s., $\mathbb{P}(B_s + bs > x - Z_s) \leq \mathbb{P}(B_s + bs > x)$. Therefore

$$\frac{\mathbb{P}(X_s > x)}{\mathbb{P}(X_t > x)} \leq \frac{1}{\mathbb{P}(W_1 > t)} \frac{\mathbb{P}(B_s + bs > x)}{\mathbb{P}(B_t + bt > x)} \to 0,$$

as $x \to \infty$, by Example 3.4.3.

Properties

We now work towards establishing the fundamental tools concerning synergistic processes. In the next lemma, we make use of Definition 2.6.3.

Lemma 3.4.6. *Let $(X_t)_{t\geq 0}$ be a process regenerative at first passage times. Then, for any $t > 0$ and $0 \leq x_0 < x$,*

$$\mathbb{P}\big(\overline{X}_t > x\big) \inf_{0\leq s\leq t} \mathbb{P}(X_s \geq -x_0) \;\leq\; \mathbb{P}(X_t > x - x_0)\,. \tag{3.4.3}$$

Proof. Since $\{T_x \leq t\} = \{\overline{X}_t > x\}$, we have

$$\mathbb{P}(X_t > x - x_0 \,|\, T_x \leq t) = \frac{\mathbb{P}(X_t > x - x_0,\, T_x \leq t)}{\mathbb{P}(T_x \leq t)} \leq \frac{\mathbb{P}(X_t > x - x_0)}{\mathbb{P}(T_x \leq t)} = \frac{\mathbb{P}(X_t > x - x_0)}{\mathbb{P}\big(\overline{X}_t > x\big)}.$$

Therefore,

$$\frac{\mathbb{P}(X_t > x - x_0)}{\mathbb{P}\big(\overline{X}_t > x\big)} \geq \mathbb{P}(X_t > x - x_0 \,|\, T_x \leq t)\,. \tag{3.4.4}$$

Now observe that on the event $\{T_x \leq t\}$, if

$$X_t - X_{T_x} \geq -x_0,$$

then, since $X_{T_x} \geq x$, we have

$$X_t \geq X_{T_x} - x_0 \geq x - x_0.$$

Hence

$$\{X_t - X_{T_x} \geq -x_0\} \subseteq \{X_t \geq x - x_0\} \qquad \text{on } \{T_x \leq t\},$$

which yields

$$\mathbb{P}(X_t > x - x_0 \,|\, T_x \leq t) \geq \mathbb{P}(X_t - X_{T_x} \geq -x_0 \,|\, T_x \leq t)\,. \tag{3.4.5}$$

Conditioning on the value of T_x, we obtain

$$\mathbb{P}(X_t - X_{T_x} \geq -x_0 \,|\, T_x \leq t) = \int_0^t \mathbb{P}(X_t - X_{T_x} \geq -x_0 \,|\, T_x = s)\; \mathbb{P}(T_x \in ds \,|\, T_x \leq t)\,.$$

By the regeneration property at first passage times,

$$\mathbb{P}(X_t - X_{T_x} \geq -x_0 \,|\, T_x = s) = \mathbb{P}(X_{t-s} \geq -x_0)\,, \qquad 0 \leq s \leq t.$$

Therefore

$$\begin{aligned}
\mathbb{P}(X_t - X_{T_x} \geq -x_0 \,|\, T_x \leq t) &= \int_0^t \mathbb{P}(X_{t-s} \geq -x_0)\ \mathbb{P}(T_x \in ds \,|\, T_x \leq t) \\
&\geq \inf_{0\leq s\leq t} \mathbb{P}(X_s \geq -x_0) \int_0^t \mathbb{P}(T_x \in ds \,|\, T_x \leq t) \\
&= \inf_{0\leq s\leq t} \mathbb{P}(X_s \geq -x_0)\,.
\end{aligned}$$

Combining the above inequality with (3.4.4), (3.4.5), we conclude that

$$\frac{\mathbb{P}(X_t > x - x_0)}{\mathbb{P}\big(\overline{X}_t > x\big)} \geq \inf_{0\leq s\leq t} \mathbb{P}(X_s \geq -x_0)\,.$$

□

Remark 3.4.7. *The above lemma applies in particular to Lévy processes, since they are regenerative at first passage times (see Section 2.6).*

Remark 3.4.8. *The sequence of steps used in the proof above — namely (1) applying Bayes' formula, (2) using $X_{T_x} \geq x$, (3) conditioning on the first passage time T_x and exploiting regeneration at first passage times. — will appear again in subsequent proofs. We will not carry out these steps as analytically in every instance.*

The following property provides a partial characterization of synergistic Lévy processes and motivates the terminology.

Proposition 3.4.9. *Let $(X_t)_{t\geq 0}$ be a Lévy process that is synergistic at extremes up to time t. Then*

$$\lim_{x\to\infty} \mathbb{P}(T_x > t - \epsilon \,|\, T_x \leq t) = 1, \qquad \textit{for every } \epsilon \in (0,t). \tag{3.4.6}$$

Conversely, if $(X_t)_{t\geq 0}$ satisfies (3.4.6) *for some $t > 0$ and, in addition,*

$$\inf_{0<s<t} \mathbb{P}(X_s \geq 0) > 0, \tag{3.4.7}$$

then $(X_t)_{t\geq 0}$ *is synergistic at extremes up to time* t.

Proof. (⇒) Suppose $(X_t)_{t\geq 0}$ is synergistic at extremes up to time t. Let $\epsilon > 0$. We aim to show that

$$\lim_{x\to\infty} \mathbb{P}(T_x > t-\epsilon \,|\, T_x \leq t) = 1.$$

Since $\{T_x \leq t\} = \{\overline{X}_t > x\}$ and $\{T_x \leq t-\epsilon\} = \{\overline{X}_{t-\epsilon} > x\}$, we have

$$\mathbb{P}(T_x > t-\epsilon \,|\, T_x \leq t) = 1 - \mathbb{P}(T_x \leq t-\epsilon \,|\, T_x \leq t) = 1 - \frac{\mathbb{P}\big(\overline{X}_{t-\epsilon} > x\big)}{\mathbb{P}\big(\overline{X}_t > x\big)}.$$

Therefore, it suffices to show that

$$\lim_{x\to\infty} \frac{\mathbb{P}\big(\overline{X}_{t-\epsilon} > x\big)}{\mathbb{P}\big(\overline{X}_t > x\big)} = 0. \tag{3.4.8}$$

Fix x_0 large enough so that $\mathbb{P}(\inf_{0<s<t-\epsilon} X_s \geq -x_0) > 0$, and let $x > x_0$. Then

$$\begin{aligned} \frac{\mathbb{P}\big(\overline{X}_{t-\epsilon} > x\big)}{\mathbb{P}\big(\overline{X}_t > x\big)} &\leq \frac{\mathbb{P}\big(\overline{X}_{t-\epsilon} > x\big)}{\mathbb{P}(X_t > x)} \\ &\leq \frac{\mathbb{P}(X_{t-\epsilon} > x - x_0)}{\mathbb{P}(\inf_{0\leq s\leq t-\epsilon} X_s \geq -x_0)} \cdot \frac{1}{\mathbb{P}(X_t > x)}, \end{aligned} \tag{3.4.9}$$

where the first inequality uses $\overline{X}_t \geq X_t$ almost surely, and the second inequality follows from Lemma 3.4.6 together with

$$\inf_{0<s<t} \mathbb{P}(X_s \geq -x_0) \;\geq\; \mathbb{P}\Big(\inf_{0<s<t} X_s \geq -x_0\Big).$$

Fix $l \in (t-\epsilon, t)$. By independence and stationarity of increments,

$$\mathbb{P}(X_l > x) \;\geq\; \mathbb{P}(X_{t-\epsilon} > x - x_0)\, \mathbb{P}\big(X_{l-(t-\epsilon)} > x_0\big).$$

Substituting this bound into (3.4.9) yields

$$\frac{\mathbb{P}\big(\overline{X}_{t-\epsilon} > x\big)}{\mathbb{P}\big(\overline{X}_t > x\big)} \;\leq\; \frac{1}{\mathbb{P}(\inf_{0<s<t-\epsilon} X_s \geq -x_0)} \cdot \frac{1}{\mathbb{P}\big(X_{l-(t-\epsilon)} > x_0\big)} \cdot \frac{\mathbb{P}(X_l > x)}{\mathbb{P}(X_t > x)}.$$

By the definition of synergistic-at-extremes processes, the last ratio tends to 0 as $x \to \infty$, which proves (3.4.8).

(⇐) Now assume (3.4.6) holds and $\inf_{0<s<t} \mathbb{P}(X_s \geq 0) > 0$. Let $\epsilon \in (0,t)$. We must show

$$\lim_{x\to\infty} \frac{\mathbb{P}(X_{t-\epsilon} > x)}{\mathbb{P}(X_t > x)} = 0.$$

By Lemma 3.4.6 and the inequality $\mathbb{P}(X_{t-\epsilon} > x) \leq \mathbb{P}\big(\overline{X}_{t-\epsilon} > x\big)$,

$$\frac{\mathbb{P}(X_{t-\epsilon} > x)}{\mathbb{P}(X_t > x)} \leq \frac{\mathbb{P}\big(\overline{X}_{t-\epsilon} > x\big)}{\mathbb{P}\big(\overline{X}_t > x\big)} \cdot \frac{1}{\inf_{0<s<t} \mathbb{P}(X_s \geq 0)}.$$

The first factor tends to 0 as $x \to \infty$ by assumption, completing the proof. □

We believe that the extra condition (3.4.7) is merely a proof artifact, and that condition (3.2.5) alone characterizes the entire class of Lévy processes that reach extreme values through the synergy of all their increments. Nevertheless, we present an example of a process that does not satisfy it.

Example 3.4.10. *As an example of a Lévy process that does not satisfy condition* (3.4.7), *consider a compound Poisson process with negative linear drift,*

$$X_t = Z_t + bt, \qquad b < 0,$$

where $Z_t = \sum_{k=1}^{N_t} \xi_k$, N_t *is a Poisson process with interarrival times* (W_i), *and* (ξ_i) *is an i.i.d. sequence of jumps. Then*

$$\mathbb{P}(X_s > 0) \ \leq \ \mathbb{P}(N_s > 0) = \mathbb{P}(W_1 < s)\,.$$

Since $\inf_{0<s<t} \mathbb{P}(W_1 < s) = 0$, *it follows that*

$$\inf_{0<s<t} \mathbb{P}(X_s \geq 0) = 0.$$

The exponential distribution of interarrival times played no role here: the same conclusion holds for any compound renewal process with negative drift.

Example 3.4.11. *If we add a Brownian component to the previous process, then condition* (3.4.7) *is satisfied. This is natural, since the Brownian motion, having unbounded variation, dominates the linear drift and compound Poisson part at small* *times. Specifically, for $X_t = B_t + Z_t + bt$ with $b < 0$ and Z_t as above,*

$$\begin{aligned}\mathbb{P}(B_s + Z_s + bs > 0) &\geq \mathbb{P}(B_s + bs > 0,\ N_s = 0)\\ &= \mathbb{P}(B_s > -bs)\ \mathbb{P}(N_s = 0)\\ &= \mathbb{P}(B_s > -bs)\ \mathbb{P}(W_1 > s)\,.\end{aligned}$$

Hence

$$\inf_{0<s<t} \mathbb{P}(B_s + Z_s + bs > 0)\ \geq\ \mathbb{P}(B_s > -bt)\ \mathbb{P}(W_1 > t) > 0.$$

The next Lemma will become the main tool in our proofs.

Lemma 3.4.12. *Let $(X_t)_{t\geq 0}$ be a Lévy process synergistic at extremes up to time t. Then, for every $c \geq 0$,*

$$\liminf_{x\to\infty} \mathbb{P}\big(X_t > x - c \,\big|\, \overline{X}_t > x\big)\ \geq\ \lim_{\epsilon\to 0^+} \inf_{0<s\leq\epsilon} \mathbb{P}(X_s \geq -c)\,. \tag{3.4.10}$$

Proof. We expand

$$\begin{aligned}\mathbb{P}\big(X_t > x - c \,\big|\, \overline{X}_t > x\big) &= \mathbb{P}(X_t > x - c \,|\, T_x \leq t)\\ &\geq \mathbb{P}(X_t > x - c,\ T_x > t - \epsilon \,|\, T_x \leq t)\\ &= \mathbb{P}(T_x > t - \epsilon \,|\, T_x \leq t)\ \mathbb{P}(X_t > x - c \,|\, t - \epsilon < T_x \leq t)\\ &\geq \mathbb{P}(T_x > t - \epsilon \,|\, T_x \leq t)\ \mathbb{P}(X_t - X_{T_x} \geq -c \,|\, t - \epsilon < T_x \leq t)\\ &\geq \mathbb{P}(T_x > t - \epsilon \,|\, T_x \leq t)\ \inf_{t-\epsilon<s\leq t} \mathbb{P}(X_t - X_s \geq -c)\\ &= \mathbb{P}(T_x > t - \epsilon \,|\, T_x \leq t)\ \inf_{0<s\leq\epsilon} \mathbb{P}(X_s \geq -c)\,.\end{aligned}$$

Here the steps are justified as follows:

- first equality: $\{T_x \leq t\} = \{\overline{X}_t > x\}$,
- second equality: multiplication rule of probability,
- second inequality: $X_{T_x} \geq x$ almost surely,
- third inequality: regeneration of $(X_t)_{t\geq 0}$ at first passage times,
- last equality: stationarity of increments.

Finally, by Lemma 3.4.9,

$$\liminf_{x\to\infty} \mathbb{P}\big(X_t > x - c \,\big|\, \overline{X}_t > x\big)\ \geq\ \inf_{0<s\leq\epsilon} \mathbb{P}(X_s \geq -c)\,,$$

and letting $\epsilon \to 0^+$ completes the proof. □

The next result should be viewed as the dual counterpart of Lemma 3.4.12, providing an upper bound in place of a lower bound.

Lemma 3.4.13. *Let $(X_t)_{t\geq 0}$ be a Lévy process synergistic at extremes up to time t. Then, for every $\epsilon > 0$,*

$$\limsup_{x\to\infty} \mathbb{P}\big(X_t > x \,\big|\, \overline{X}_t > x\big) \leq 1 - \inf_{0<s<\epsilon} \mathbb{P}(X_s \leq 0) \limsup_{x\to\infty} \mathbb{P}(X_{T_x} = x \,|\, t-\epsilon < T_x \leq t)\,.$$

Proof. Let $\epsilon > 0$. Using $\{T_x \leq t\} = \{\overline{X}_t > x\}$ and the law of total probability,

$$\begin{aligned} \mathbb{P}\big(X_t > x \,\big|\, \overline{X}_t > x\big) = \mathbb{P}(X_t > x \,|\, t-\epsilon < T_x \leq t)\, \mathbb{P}(T_x > t-\epsilon \,|\, T_x \leq t) \\ + \ \mathbb{P}(X_t > x,\ T_x < t-\epsilon \mid T_x \leq t)\,. \end{aligned} \tag{3.4.11}$$

Likewise,

$$\begin{aligned} \mathbb{P}(X_t > x \,|\, t-\epsilon < T_x \leq t) \leq \mathbb{P}(X_t > x \,|\, X_{T_x} = x,\ t-\epsilon < T_x \leq t) \\ \cdot\, \mathbb{P}(X_{T_x} = x \,|\, t-\epsilon < T_x \leq t) \\ + \mathbb{P}(X_{T_x} > x \,|\, t-\epsilon < T_x \leq t)\,. \end{aligned} \tag{3.4.12}$$

By stationarity and independence of increments,

$$\mathbb{P}(X_t > x \,|\, X_{T_x} = x,\ t-\epsilon < T_x \leq t) \leq \sup_{0<s<\epsilon} \mathbb{P}(X_s > 0)\,. \tag{3.4.13}$$

Combining (3.4.11)–(3.4.13) and taking $\limsup_{x\to\infty}$, the term $\mathbb{P}(X_t > x,\ T_x < t-\epsilon \mid T_x \leq t)$ vanishes and $\mathbb{P}(T_x > t-\epsilon \,|\, T_x \leq t) \to 1$ by Lemma 3.4.9, yielding

$$\limsup_{x\to\infty} \mathbb{P}\big(X_t > x \,\big|\, \overline{X}_t > x\big) \leq 1 - \inf_{0<s<\epsilon} \mathbb{P}(X_s \leq 0) \limsup_{x\to\infty} \mathbb{P}(X_{T_x} = x \,|\, t-\epsilon < T_x \leq t)\,,$$

where we also used that $\sup_{0<s<\epsilon} \mathbb{P}(X_s > 0) = 1 - \inf_{0<s<\epsilon} \mathbb{P}(X_s \leq 0)$. □

The following proposition shows that, for synergistic processes, once a large level x is crossed before time t, the process will remain arbitrarily close to x at time t, even if it does not stay strictly above x.

Proposition 3.4.14. *Let $(X_t)_{t\geq 0}$ be a Lévy process, synergistic at extremes up to some time $t > 0$. Then, for any $c > 0$,*

$$\lim_{x\to\infty} \mathbb{P}\big(X_t > x - c \,\big|\, \overline{X}_t > x\big) = 1.$$

Proof. By Lemma 3.4.12, it suffices to show that

$$\lim_{\epsilon\to 0^+} \inf_{0\leq s\leq\epsilon} \mathbb{P}(X_s \geq -c) = 1.$$

First, since

$$\inf_{0\leq s\leq\epsilon} \mathbb{P}(X_s \geq -c) \;\geq\; \mathbb{P}\left(\inf_{0\leq s\leq\epsilon} X_s \geq -c\right),$$

and both sides are non-decreasing as $\epsilon \downarrow 0$ and therefore the limits exist, we obtain

$$\lim_{\epsilon\to 0^+} \inf_{0\leq s\leq\epsilon} \mathbb{P}(X_s \geq -c) \;\geq\; \lim_{\epsilon\to 0^+} \mathbb{P}\left(\inf_{0\leq s\leq\epsilon} X_s \geq -c\right). \tag{3.4.14}$$

Again, since the limit on the right exists, and we may write

$$\lim_{\epsilon\to 0^+} \mathbb{P}\left(\inf_{0\leq s\leq\epsilon} X_s \geq -c\right) = \lim_{n\to\infty} \mathbb{P}\left(\inf_{0\leq s\leq 1/n} X_s \geq -c\right). \tag{3.4.15}$$

By continuity of measure for the increasing sequence of events

$$E_n := \bigcap_{s\in(0,\,1/n)} \{X_s \geq -c\},$$

we obtain

$$\lim_{n\to\infty} \mathbb{P}\left(\inf_{0\leq s\leq 1/n} X_s \geq -c\right) = \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \bigcap_{s\in(0,\,1/n)} \{X_s \geq -c\}\right). \tag{3.4.16}$$

With $T_{-c} := \inf\{t \geq 0 : X_t = -c\}$, it follows that

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \bigcap_{s\in(0,\,1/n)} \{X_s \geq -c\}\right) \;\geq\; \mathbb{P}(T_{-c} > 0). \tag{3.4.17}$$

Combining (3.4.14)–(3.4.17), we conclude

$$\lim_{\epsilon\to 0^+} \inf_{0\le s\le \epsilon} \mathbb{P}(X_s \ge -c) \;\ge\; \mathbb{P}(T_{-c} > 0)\,.$$

Finally, $\mathbb{P}(T_{-c} > 0) = 1$ for every process with right–continuous paths starting at $X_0 = 0$. Indeed, on $\{T_{-c} = 0\}$ there exists a sequence $t_n \downarrow 0$ such that $X_{t_n} = -c$. By right–continuity,

$$X_0 = \lim_{n\to\infty} X_{t_n} = -c.$$

Hence,

$$\{T_{-c} = 0\} \subseteq \{X_0 = -c\}$$

and since $X_0 = 0$ a.s. we conclude that $\mathbb{P}(T_{-c} = 0) = 0$. □

The following corollary of Proposition 3.4.14 provides an upper bound for $\mathbb{P}\big(\overline{X}_t > x\big)$, for large x, in terms of the distribution of X_t.

Corollary 3.4.15. *Let $(X_t)_{t\ge 0}$ be a Lévy process, synergistic at extremes up to some time $t > 0$. Then, for any $c > 0$,*

$$\limsup_{x\to\infty} \frac{\mathbb{P}\big(\overline{X}_t > x\big)}{\mathbb{P}(X_t > x - c)} \le 1.$$

3.4.2 Compound Poisson with linear drift

Here we prove asymptotic equivalence to the supremum for the compound Poisson process with linear drift. In the case of negative drift we will need the following lemma, which shows that for compound Poisson processes with non-positive drift, if they are synergistic at extremes, then conditional on $\{T_x \le t\}$ and for large x, the process reaches level x with its *last jump before time* t.

Lemma 3.4.16. *Let $X_t = Z_t + bt$, where Z_t is a compound Poisson process and $b < 0$. Let (Γ_n) be the arrival times of the Poisson process associated with Z_t. If $(X_t)_{t\ge 0}$ is synergistic at extremes up to time t, then*

$$\lim_{x\to\infty} \mathbb{P}(T_x = \Gamma_{N_t} \,|\, T_x \le t) = 1. \tag{3.4.18}$$

Proof. Let (W_i) denote the i.i.d. sequence of interarrival times of N_t. Because the drift is negative, the process X cannot creep upwards; it can cross x only at a jump. Hence $T_x = \Gamma_i$ for some $i \in \{1, \dots, N_t\}$.

We therefore write

$$\begin{aligned}
\mathbb{P}(T_x = \Gamma_{N_t} \,|\, T_x \le t) &\ge \mathbb{P}(N_t - N_{T_x} = 0 \,|\, T_x \le t) \\
&\ge \mathbb{P}(N_t - N_{T_x} = 0,\, T_x > t - \epsilon \,|\, T_x \le t) \\
&= \mathbb{P}(N_t - N_{T_x} = 0 \,|\, t - \epsilon < T_x \le t)\ \mathbb{P}(T_x > t - \epsilon \,|\, T_x \le t)\,.
\end{aligned} \tag{3.4.19}$$

Applying the law of total probability, we obtain

$$\begin{aligned}
\mathbb{P}(N_t - N_{T_x} = 0 \,|\, t - \epsilon < T_x \le t) &= \int_{t-\epsilon}^{t} \mathbb{P}(N_t - N_{T_x} = 0 \,|\, T_x = s)\ \mathbb{P}(T_x \in ds \,|\, t - \epsilon < T_x \le t) \\
&= \int_{t-\epsilon}^{t} \mathbb{P}(N_t - N_s = 0)\ \mathbb{P}(T_x \in ds \,|\, t - \epsilon < T_x \le t) \\
&= \int_{t-\epsilon}^{t} \mathbb{P}(N_{t-s} = 0)\ \mathbb{P}(T_x \in ds \,|\, t - \epsilon < T_x \le t) \\
&\ge \inf_{t-\epsilon<s<t} \mathbb{P}(W_1 > t - s) \int_{t-\epsilon}^{t} \mathbb{P}(T_x \in ds \,|\, t - \epsilon < T_x \le t) \\
&= \mathbb{P}(W_1 > \epsilon)\,.
\end{aligned} \tag{3.4.20}$$

Here we used that T_x is almost surely a jump time, and that the Poisson process regenerates at first passage times.

Combining (3.4.19) and (3.4.20), we get

$$\mathbb{P}(T_x = \Gamma_{N_t} \,|\, T_x \le t) \ge \mathbb{P}(W_1 > \epsilon)\ \mathbb{P}(T_x > t - \epsilon \,|\, T_x \le t)\,.$$

By Proposition 3.4.9,

$$\liminf_{x\to\infty} \mathbb{P}(T_x = \Gamma_{N_t} \,|\, T_x \le t) \ge \mathbb{P}(W_1 > \epsilon)\,.$$

Finally, letting $\epsilon \to 0^+$ yields the desired result (3.4.18). □

Theorem 3.4.17. *Let $(X_t)_{t\ge0}$ be a compound Poisson process with linear drift,*

$$X_t = Z_t + bt,$$

where $b \in \mathbb{R}$ is a constant. Denote by $(N_t)_{t\ge0}$ the Poisson process associated with $(Z_t)_{t\ge0}$, *and by* (Γ_n) *the sequence of its arrival times.*

Suppose $(X_t)_{t\geq 0}$ *is synergistic at extremes up to some time* $t > 0$*. Then:*

(i) If $b \geq 0$*,*

$$\mathbb{P}\big(\overline{X}_t > x\big) \;\sim\; \mathbb{P}(X_t > x)\,, \qquad x \to \infty.$$

(ii) If $b < 0$*,*

$$\mathbb{P}\big(\overline{X}_t > x\big) \;\sim\; \mathbb{P}\big(X_{\Gamma_{N_t}} > x\big)\,, \qquad x \to \infty.$$

Proof. (i) $b \geq 0$: From Lemma 3.4.12 with $c = 0$,

$$\liminf_{x\to\infty} \mathbb{P}\big(X_t > x \,\big|\, \overline{X}_t > x\big) \;\geq\; \lim_{\epsilon\to 0^+} \inf_{0<s\leq\epsilon} \mathbb{P}(X_s \geq 0)\,.$$

For a compound Poisson process with drift $b \geq 0$,

$$\mathbb{P}(X_s \geq 0) \;\geq\; \mathbb{P}(N_s = 0) = \mathbb{P}(W_1 > s)\,.$$

Taking the infimum over $0 < s \leq \epsilon$ and sending $\epsilon \to 0^+$ gives 1. Hence

$$\lim_{x\to\infty} \mathbb{P}\big(X_t > x \,\big|\, \overline{X}_t > x\big) = 1,$$

which implies the desired asymptotic equivalence.

(ii) $b < 0$: Note that

$$\{T_x = \Gamma_{N_t}\} \cap \{T_x \leq t\} \;\subseteq\; \{X_{\Gamma_{N_t}} > x\} \cap \{\overline{X}_t > x\}.$$

By Lemma 3.4.16,

$$\liminf_{x\to\infty} \mathbb{P}\big(X_{\Gamma_{N_t}} > x \,\big|\, \overline{X}_t > x\big) \;\geq\; \liminf_{x\to\infty} \mathbb{P}(T_x = \Gamma_{N_t} \,|\, T_x \leq t) = 1,$$

which yields the result.

□

Remark 3.4.18. *As we have already discussed, Theorem 3.2.5 of Braverman is implied by Theorem 3.4.17, since Lemma 5 in [14] guarantees that a compound Poisson process with Braverman-light tails is synergistic at extremes.*

3.4.3 Lévy processes with exponential tails

We now prove the following result.

Theorem 3.4.19. *Let $(X_t)_{t\geq 0}$ be a Lévy process, synergistic at extremes up to some time $t > 0$, and let $\alpha \geq 0$. Then*

$$X_t \in \mathcal{L}(\alpha) \qquad \Longleftrightarrow \qquad \overline{X}_t \in \mathcal{L}(\alpha).$$

Moreover, if either condition holds, then

$$\mathbb{P}\big(\overline{X}_t > x\big) \;\sim\; \mathbb{P}(X_t > x)\,, \qquad x \to \infty.$$

Proof. **(⇒)** Assume $X_t \in \mathcal{L}(\alpha)$. For any $c > 0$,

$$\frac{\mathbb{P}(X_t > x)}{\mathbb{P}\big(\overline{X}_t > x\big)} = \frac{\mathbb{P}(X_t > x)}{\mathbb{P}(X_t > x - c)} \cdot \frac{\mathbb{P}(X_t > x - c)}{\mathbb{P}\big(\overline{X}_t > x\big)}.$$

The first factor converges to $e^{-\alpha c}$ by the defining property of $\mathcal{L}(\alpha)$. For the second factor,

$$\frac{\mathbb{P}(X_t > x - c)}{\mathbb{P}\big(\overline{X}_t > x\big)} \;\geq\; \mathbb{P}\big(X_t > x - c \,\big|\, \overline{X}_t > x\big)\,,$$

by Bayes' formula. By Proposition 3.4.14,

$$\liminf_{x\to\infty} \mathbb{P}\big(X_t > x - c \,\big|\, \overline{X}_t > x\big) = 1.$$

Hence

$$\liminf_{x\to\infty} \frac{\mathbb{P}(X_t > x)}{\mathbb{P}\big(\overline{X}_t > x\big)} \;\geq\; e^{-\alpha c}.$$

Letting $c \to 0^+$ gives

$$\mathbb{P}(X_t > x) \;\sim\; \mathbb{P}\big(\overline{X}_t > x\big)\,, \qquad x \to \infty,$$

so $\overline{X}_t \in \mathcal{L}(\alpha)$.

(⇐) The argument is symmetric. Write

$$\frac{\mathbb{P}(X_t > x)}{\mathbb{P}\big(\overline{X}_t > x\big)} = \frac{\mathbb{P}(X_t > x)}{\mathbb{P}\big(\overline{X}_t > x + c\big)} \cdot \frac{\mathbb{P}\big(\overline{X}_t > x + c\big)}{\mathbb{P}\big(\overline{X}_t > x\big)},$$

and repeat the same reasoning as above. □

Remark 3.4.20. *This result is weaker than Theorem 3.2.8 of Albin, but stronger than Corollary 6.2 in [1]. Albin's theorem relies on a discretisation method together with Prokhorov's theorem, whereas our proof exploits the synergistic property and Proposition 3.4.14, leading to a simpler argument.*

3.4.4 Linear Brownian Motion

We now prove asymptotic equivalence to the supremum for the linear Brownian motion.

Fact 3.4.21. *Let $X_t := \sigma B_t + bt$, where $b \in \mathbb{R}$. Then, for any $t > 0$,*

$$\mathbb{P}\big(\overline{X}_t > x\big) \ \sim \ 2\,\mathbb{P}(X_t > x)\,, \qquad x \to \infty.$$

Proof. Since $(X_t)_{t\geq 0}$ is synergistic at extremes up to time t (see Example 3.4.3), the claim follows by combining Lemma 3.4.12 and Lemma 3.4.13.

In particular, for the lower bound, Lemma 3.4.12 gives

$$\liminf_{x\to\infty} \mathbb{P}\big(X_t > x \,\big|\, \overline{X}_t > x\big) \ \geq \ \lim_{\epsilon\to 0^+} \inf_{0<s\leq\epsilon} \mathbb{P}(X_s \geq 0)\,.$$

If $b < 0$, then

$$\inf_{0<s\leq\epsilon} \mathbb{P}(X_s \geq 0) = \inf_{0<s\leq\epsilon} \overline{\Phi}\Big(\tfrac{-b\sqrt{s}}{\sigma}\Big) = \overline{\Phi}\Big(\tfrac{-b\sqrt{\epsilon}}{\sigma}\Big)\,.$$

If $b \geq 0$, then

$$\inf_{0<s\leq\epsilon} \mathbb{P}(X_s \geq 0) = \inf_{0<s\leq\epsilon} \overline{\Phi}\Big(\tfrac{-b\sqrt{s}}{\sigma}\Big) = \overline{\Phi}(0).$$

In both cases,

$$\lim_{\epsilon\to 0^+} \inf_{0<s\leq\epsilon} \mathbb{P}(X_s \geq 0) = \tfrac{1}{2},$$

so

$$\liminf_{x\to\infty} \mathbb{P}\big(X_t > x \,\big|\, \overline{X}_t > x\big) \ \geq \ \tfrac{1}{2}.$$

For the upper bound, Lemma 3.4.13 yields, for any $\epsilon > 0$,

$$\limsup_{x\to\infty} \mathbb{P}\big(X_t > x \,\big|\, \overline{X}_t > x\big) \ \leq \ 1 - \inf_{0<s<\epsilon} \mathbb{P}(X_s \leq 0)\, \limsup_{x\to\infty} \mathbb{P}(X_{T_x} = x \,|\, t-\epsilon < T_x \leq t)\,.$$

Since X_t has continuous paths almost surely,

$$\mathbb{P}(X_{T_x} = x \,|\, t - \epsilon < T_x \le t) = 1.$$

Moreover, by symmetry,

$$\lim_{\epsilon \to 0^+} \inf_{0 < s \le \epsilon} \mathbb{P}(X_s < 0) = \tfrac{1}{2}.$$

Letting $\epsilon \to 0^+$ gives

$$\limsup_{x \to \infty} \mathbb{P}\big(X_t > x \,\big|\, \overline{X}_t > x\big) \;\le\; \tfrac{1}{2}.$$

Combining the lower and upper bounds, we conclude that

$$\mathbb{P}\big(\overline{X}_t > x\big) \;\sim\; 2\,\mathbb{P}(X_t > x)\,, \qquad x \to \infty.$$

□

3.4.5 Perturbed Ruin Process

We now establish asymptotic equivalence to the supremum for the process

$$X_t := \sigma B_t + Z_t + bt,$$

where $\sigma > 0$, $b \in \mathbb{R}$, and Z_t is a compound Poisson process with non-positive jumps, independent of the Brownian motion B_t. To our knowledge, this case has not been treated previously in the literature.

When $b > 0$, this model generalises the classical ruin process from risk theory (see [19, 34]). Here b represents the premium rate, while the compound Poisson process with negative jumps models the aggregate claims. The Brownian component is interpreted as additional uncertainty in either the premium income or the claim process. The infimum of this process typically attracts the most attention, as it governs the probability of default (ruin).

Proposition 3.4.22. *Let $X_t = \sigma B_t + Z_t + bt$, with $\sigma > 0$, $b \in \mathbb{R}$, and Z_t a compound Poisson process with non-positive jumps, independent of B_t. Then*

$$\mathbb{P}\big(\overline{X}_t > x\big) \;\sim\; 2\,\mathbb{P}(X_t > x)\,, \qquad x \to \infty.$$

Proof. Without loss of generality, set $\sigma = 1$.

Lower bound. By Example 3.4.5, the process $(X_t)_{t\geq 0}$ is synergistic at extremes up to time t. Hence Lemma 3.4.12 yields

$$\liminf_{x\to\infty} \mathbb{P}\big(X_t > x \,\big|\, \overline{X}_t > x\big) \;\geq\; \lim_{\epsilon\to 0^+} \inf_{0<s\leq\epsilon} \mathbb{P}(X_s \geq 0)\,.$$

Let $(N_t)_{t\geq 0}$ be the Poisson process associated with $(Z_t)_{t\geq 0}$, and (W_i) the waiting times. For $x, t > 0$,

$$\begin{aligned}\mathbb{P}(B_t + Z_t + bt > x) &\geq \mathbb{P}(B_t + Z_t + bt > x,\, N_t = 0) \\ &= \mathbb{P}(B_t + bt > x)\ \mathbb{P}(W_1 > t)\,,\end{aligned} \tag{3.4.21}$$

by independence of N_t and B_t. Thus

$$\lim_{\epsilon\to 0^+} \inf_{0<s\leq\epsilon} \mathbb{P}(X_s \geq 0) \;\geq\; \lim_{\epsilon\to 0^+} \inf_{0<s\leq\epsilon} \overline{\Phi}(-b\sqrt{s})\,\mathbb{P}(W_1 > s) = \tfrac{1}{2}.$$

Therefore

$$\liminf_{x\to\infty} \mathbb{P}\big(X_t > x \,\big|\, \overline{X}_t > x\big) \;\geq\; \tfrac{1}{2}.$$

Upper bound. By the law of total probability, for $x > 0$,

$$\begin{aligned}\mathbb{P}(B_t + bt + Z_t > x) &= \mathbb{P}(B_t + bt + Z_t > x,\, N_t = 0) + \mathbb{P}(B_t + bt + Z_t > x,\, N_t > 0) \\ &= \mathbb{P}(B_t + bt > x)\ \mathbb{P}(W_1 > t) + \mathbb{P}(B_t + bt + Z_t > x,\, Z_t < 0)\,,\end{aligned} \tag{3.4.22}$$

since jumps are a.s. non-positive. Also, for any $x > 0$,

$$\mathbb{P}\big(\overline{B_t + bt + Z_t} > x\big) \;\geq\; \mathbb{P}\big(\overline{B_t + bt} > x,\, N_t = 0\big) = \mathbb{P}\big(\overline{B_t + bt} > x\big)\ \mathbb{P}(W_1 > t)\,. \tag{3.4.23}$$

Combining (3.4.22)–(3.4.23),

$$\frac{\mathbb{P}(X_t > x)}{\mathbb{P}\big(\overline{X}_t > x\big)} \;\leq\; \frac{\mathbb{P}(B_t + bt > x)}{\mathbb{P}\big(\overline{B_t + bt} > x\big)} + \frac{\mathbb{P}(B_t + bt + Z_t > x,\, Z_t < 0)}{\mathbb{P}\big(\overline{B_t + bt} > x\big)\ \mathbb{P}(W_1 > t)}.$$

The first term converges to $1/2$ as $x \to \infty$ by Fact 3.4.21.

For the second term, note

$$\frac{\mathbb{P}(B_t + bt + Z_t > x,\, Z_t < 0)}{\mathbb{P}\big(\overline{B_t + bt} > x\big)} \;\le\; \frac{\mathbb{P}(B_t + bt + Z_t > x,\, Z_t < 0)}{\mathbb{P}(B_t + bt > x)}.$$

By independence of B_t and Z_t,

$$\frac{\mathbb{P}(B_t + bt + Z_t > x,\, Z_t < 0)}{\mathbb{P}(B_t + bt > x)} = \int_{-\infty}^{0} \frac{\mathbb{P}(B_t + bt > x - y)}{\mathbb{P}(B_t + bt > x)}\, \mathbb{P}(Z_t \in dy)\,,$$

with the integral replaced by a sum if jump distribution is discrete. Since[2] for each $y < 0$

$$\frac{\mathbb{P}(B_t + bt > x - y)}{\mathbb{P}(B_t + bt > x)} \xrightarrow{x\to\infty} 0,$$

and

$$\frac{\mathbb{P}(B_t + bt > x - y)}{\mathbb{P}(B_t + bt > x)} < 1,$$

the Bounded Convergence Theorem implies that the integral tends to 0.

Hence the second term vanishes, and

$$\limsup_{x\to\infty} \mathbb{P}\big(X_t > x \,\big|\, \overline{X}_t > x\big) \;\le\; \tfrac{1}{2}.$$

Together with the lower bound, this yields

$$\mathbb{P}\big(\overline{X}_t > x\big) \;\sim\; 2\,\mathbb{P}(X_t > x)\,, \qquad x \to \infty.$$

□

3.4.6 Recap of the Method and Possible Extensions

Roadmap of the method

We have already described the heuristic ideas behind our approach, but let us summarise the steps in a compact "roadmap." To prove tail equivalence to the supremum, we proceed as follows:

1. Show that the process $(X_t)_{t\ge 0}$ is synergistic at extremes.

[2]For the limit see also (2.2.15).

2. Apply Lemma 3.4.12 to obtain a lower bound for

$$\liminf_{x\to\infty} \frac{\mathbb{P}(X_t > x)}{\mathbb{P}\big(\overline{X}_t > x\big)},$$

in terms of the small-time behaviour of the process, i.e. $\lim_{\epsilon\to 0^+} \inf_{0<s<\epsilon} \mathbb{P}(X_s \geq 0)$.

3. Analyse this small-time behaviour: if the resulting bound equals 1, we conclude that $\mathbb{P}\big(\overline{X}_t > x\big) \sim \mathbb{P}(X_t > x)$. If not — as in the cases of linear Brownian motion or the perturbed ruin process — proceed to step 4.
4. Seek a matching upper bound, either via Lemma 3.4.13 (particularly useful when the process creeps almost surely, as with Brownian motion with drift), or through an ad hoc argument (as in the perturbed ruin process).

In what follows we present an additional tool that can also be used for establishing such upper bounds.

An additional tool

We present a simple inequality that can be useful for deriving the upper bound in step 4 above. For example, it could have been applied in the case of linear Brownian motion with negative drift, or in proving the upper bound for $X_t = \sigma B_t + Z_t + bt$ with $b < 0$, where it would have simplified the argument. It may also prove useful in the possible extensions discussed later. For completeness, we also state its symmetric counterpart. The proof is given in Section A.3.

Lemma 3.4.23. *Let $(B_t)_{t\geq 0}$ be a standard Brownian motion, $(C_t)_{t\geq 0}$ a subordinator independent of $(B_t)_{t\geq 0}$, and $\sigma > 0$. Then, for $x > 0$,*

$$\mathbb{P}\big(\overline{\sigma B_t - C_t} > x\big) \;\geq\; 2\,\mathbb{P}(\sigma B_t - C_t > x)\,, \tag{3.4.24}$$

and for all $x \in \mathbb{R}$,

$$\mathbb{P}\big(\overline{\sigma B_t + C_t} > x\big) \;\leq\; 2\,\mathbb{P}(\sigma B_t + C_t > x)\,. \tag{3.4.25}$$

Possible future directions

The method detailed above can be applied beyond Lévy processes, as we will see in the next chapter with compound renewal processes. Beyond that, it seems natural to expect it to work for processes of the form $X_t = B_t + g(t)$, where g is a deterministic drift. More generally, one could ask for which classes of processes of the type $X_t = g(B_t, t)$ or $X_t = g(t, Y_t)$, with Y_t a jump process, the method still applies. Caution is needed, however: if the resulting process is a continuous local martingale, then by the Dubins–Schwarz theorem it enjoys the reflection principle just as in the Brownian case, making the problem trivial.

Chapter 4

Suprema of Compound Renewal Processes

4.1 Introduction and Statement of Main Result

In the previous chapter we identified the structural properties responsible for tail equivalence to the supremum. In particular, while Lévy processes possess independent and stationary increments, our arguments relied only on a weaker feature: regeneration at first passage times.

Moreover, in the compound Poisson setting, the exponential distribution of the waiting times played no essential role in the proofs. These observations suggest that the method extends beyond the Lévy framework to a broader class of regenerative models.

In this chapter we prove tail equivalence to the supremum for a broad class of compound renewal processes. Specifically, we establish the result for compound renewal processes with Braverman-light jumps under a mild condition on the waiting-time distribution. This condition (stated below) includes the exponential distribution as a special case, and therefore our theorem strictly extends Braverman's Theorem 3.2.5.

Recall from Section 2.4 that a compound renewal process is constructed from two i.i.d. sequences: the waiting times $(W_n)_{n\geq 1}$, which are non-lattice, almost surely positive, and have c.d.f. F, and the jumps $(\xi_n)_{n\geq 1}$. More specifically, set $T_0 := 0$ and

$$T_n \;=\; \sum_{i=1}^{n} W_i, \qquad n \geq 1.$$

The process $(T_n)_{n\geq 0}$ denotes the sequence of renewal times, also referred to as renewals or renewal epochs. Define the associated renewal counting process

$$N_t \;:=\; \max\{\, n \in \mathbb{N}_+ : \; T_n \leq t \,\}, \qquad t \geq 0.$$

The compound renewal process is

$$Y_t \;:=\; \sum_{k=1}^{N_t} \xi_k, \qquad t \geq 0.$$

In general, Y does *not* have stationary or independent increments. However, since it regenerates at jump times T_n, and exceeds levels only via jumps, it is regenerative at first passage times; see Definition 2.6.3.

We now specify the condition we impose on the distribution of the waiting times.

Assumption 4.1.1. *Let F be the cumulative distribution function of the waiting times. We assume that F is absolutely continuous with respect to the Lebesgue measure in a neighbourhood of* 0*, and*

$$f(x) := \frac{dF(x)}{dx} = x^{a-1}\big(c_1 + c_2 x^b + o(x^b)\big), \qquad x \to 0^+,$$

for some constants $c_1 > 0$, $a > 0$, $b > 0$, and $c_2 \in \mathbb{R}$.

Remark 4.1.2. *This assumption is local: it requires absolute continuity only on some interval $(0, \rho)$ with $\rho > 0$, so that*

$$F(x) = \int_0^x f(y)\, dy, \qquad 0 < x < \rho,$$

but does not impose any restriction on the behaviour of F away from 0*. In particular, it does not imply that F is supported on $(0, \rho)$.*

The class of waiting-time distributions (non-negative, non-lattice) satisfying the above condition is broad. We now present several common examples that illustrate how Assumption 4.1.1 is satisfied both by unbounded distributions (Gamma, Weibull, Half-Normal) and by bounded-support distributions (Uniform, Beta).

Example 4.1.3 (Gamma distribution). *The* Gamma(α, λ) *law has density*

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x > 0.$$

It constitutes the prototypical example of the class of distributions satisfying Assumption 4.1.1. By Taylor expanding $e^{-\lambda x}$ *at* 0*, we identify*

$$a = \alpha, \quad b = 1, \quad c_1 = \frac{\lambda^\alpha}{\Gamma(\alpha)}, \quad c_2 = -\frac{\lambda^{\alpha+1}}{\Gamma(\alpha)}.$$

Example 4.1.4 (Uniform distribution). *If* $W \sim U(0, r)$*, then*

$$f(x) = \mathbb{1}_{(0,r)}(x), \quad x > 0,$$

so Assumption 4.1.1 holds with $a = 1$*,* $b = 1$*,* $c_1 = 1$*,* $c_2 = 0$*.*

Example 4.1.5 (Weibull distribution). *For parameters* $\alpha, \lambda > 0$ *the Weibull density is*

$$f(x) = \alpha \lambda^\alpha x^{\alpha-1} \exp\big(-(\lambda x)^\alpha\big), \quad x \geq 0.$$

Expanding $e^{-(\lambda x)^\alpha}$ *at* 0 *gives*

$$a = \alpha, \quad b = \alpha, \quad c_1 = \alpha \lambda^\alpha, \quad c_2 = -\alpha \lambda^{2\alpha}.$$

Example 4.1.6 (Half-Normal distribution). *If* $X \sim N(0, \sigma^2)$ *and* $W = |X|$*, then*

$$f(x) = \sqrt{\tfrac{2}{\pi \sigma^2}}\, e^{-x^2/(2\sigma^2)}, \quad x \geq 0.$$

Since $e^{-x^2/(2\sigma^2)} = 1 - \frac{x^2}{2\sigma^2} + o(x^2)$*, we obtain*

$$a = 1, \quad b = 2, \quad c_1 = \sqrt{\tfrac{2}{\pi\sigma^2}}, \quad c_2 = -\tfrac{1}{2\sigma^2}\sqrt{\tfrac{2}{\pi\sigma^2}}.$$

Example 4.1.7 (Beta distribution). *The* Beta(α, β) *distribution has density*

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, \qquad 0 < x < 1,$$

where $B(\alpha, \beta)$ *is the Beta function. Since* $(1-x)^{\beta-1} = 1 - (\beta - 1)x + o(x)$ *as* $x \to 0^+$*,* *we obtain*

$$a = \alpha, \quad b = 1, \quad c_1 = \tfrac{1}{B(\alpha,\beta)}, \quad c_2 = -\tfrac{\beta-1}{B(\alpha,\beta)}.$$

Remark 4.1.8. *A subclass of distributions that satisfy Assumption 4.1.1 are the ones with a density*

$$f(x) = x^{a-1} g(x) \tag{4.1.1}$$

with g right–continuous and differentiable at 0. In this case, by a Taylor expansion of g at 0, we obtain,

$$b = 1, \qquad c_1 = g(0^+), \qquad c_2 = g'_+(0).$$

This includes the Gamma, Uniform, Half-Normal, and Beta examples above.

The Weibull distribution with $\alpha \in (0,1)$ is not a member of the subclass described above. In that case, $g(x) = \alpha\lambda^\alpha e^{-(\lambda x)^\alpha}$ is not differentiable at 0 for $\alpha \in (0,1)$, but still has an asymptotic expansion in rational powers and hence satisfies the assumption.

One can build distributions that *do not* satisfy Assumption 4.1.1, by putting in place of g a function that has no limit at 0, e.g.

$$f(x) = Cx^{a-1}\big(1 + \sin(1/x)\big), \qquad 0 < x < c,$$

for $c > 0$ and a normalizing constant C. Here f oscillates too wildly near 0 to admit a power expansion.

We are now ready to state the main result of this chapter, which establishes tail equivalence to the supremum for compound renewal processes under Assumption 4.1.1.

Theorem 4.1.9. *Let $(X_t)_{t\geq 0}$ be a compound renewal process with non-negative linear drift,*

$$X_t := Y_t + bt, \qquad b \geq 0,$$

where $(Y_t)_{t\geq 0}$ is a compound renewal process. Assume that the jumps $(\xi_n)_{n\geq 1}$ are Braverman-light, the waiting times $(W_n)_{n\geq 1}$ satisfy Assumption 4.1.1, and that the jump sequence is independent of the waiting-time sequence. Then, for every $t > 0$,

$$\mathbb{P}\big(\overline{X}_t > x\big) \sim \mathbb{P}(X_t > x)\,, \qquad \text{as } x \to \infty.$$

In the sequel we work towards proving Theorem 4.1.9. In Section 4.2 we develop tools for processes that are *synergistic at extremes* beyond the Lévy setting, namely for processes that are regenerative at first passage times, in analogy with Section 3.4 of the previous chapter. To apply these tools we first need to justify that the process of interest—a compound renewal process with Braverman-light jumps (Definition 2.2.3) and waiting times satisfying Assumption 4.1.1—is synergistic at extremes (Definition 3.4.1). This requires a study of the extreme lower-tail large deviations of the renewal epochs T_n, carried out in Section 4.3. This is done by an exponential change of measure, followed by a normal approximation of the exponential integral that arises. For the latter we employ an Edgeworth expansion for a triangular array of random variables. Finally, in Section 4.4 we combine these ingredients to establish Theorem 4.1.9.

4.2 Synergistic at extremes beyond the Lévy setting

We have already introduced the property of regeneration at first passage times in Section 2.6. In this section, we develop tools analogous to those in Section 3.4 of the previous chapter, but now relying on regeneration at first passage times instead of the full Lévy property.

Recall that compound renewal processes are regenerative at first passage times by construction, as they regenerate at jump times and exceed levels only via jumps; in particular, it holds that for any measurable set B, $t \geq s$ and $x > 0$ we have

$$\mathbb{P}(X_t - X_{T_x} \in B \mid T_x = s) = \mathbb{P}(X_{t-s} \in B)\,.$$

(see also Remark 2.6.4).

The following proposition is analogous to Proposition 3.4.9.

Proposition 4.2.1. *Let $(X_t)_{t\geq 0}$ be a process that is regenerative at first passage times and synergistic at extremes up to time t, with $t > 0$ fixed. Assume in addition that*

$$\inf_{0\leq s\leq t} \mathbb{P}(X_s \geq 0) > 0. \tag{4.2.1}$$

Then,

$$\lim_{x\to\infty} \mathbb{P}(T_x > t - \epsilon \,|\, T_x \leq t) = 1, \qquad \text{for every } \epsilon \in (0, t). \tag{4.2.2}$$

Proof. Fix $\epsilon \in (0, t)$. It suffices to show that

$$\lim_{x\to\infty} \frac{\mathbb{P}\big(\overline{X}_{t-\epsilon} > x\big)}{\mathbb{P}\big(\overline{X}_t > x\big)} = 0. \tag{4.2.3}$$

By the inequality $\overline{X}_t \geq X_t$ and then Lemma 3.4.6 with $x_0 = 0$ we obtain

$$\frac{\mathbb{P}\big(\overline{X}_{t-\epsilon} > x\big)}{\mathbb{P}\big(\overline{X}_t > x\big)} \leq \frac{\mathbb{P}\big(\overline{X}_{t-\epsilon} > x\big)}{\mathbb{P}(X_t > x)} \leq \frac{\mathbb{P}(X_{t-\epsilon} > x)}{\mathbb{P}(X_t > x)} \cdot \frac{1}{\inf_{0\leq s\leq t} \mathbb{P}(X_s \geq 0)}. \tag{4.2.4}$$

By the assumption that $(X_t)_{t\geq 0}$ is synergistic at extremes up to time t, the ratio on the right-hand side tends to zero as $x \to \infty$. □

Remark 4.2.2. *The assumption* $\inf_{0\leq s\leq t} \mathbb{P}(X_s \geq 0) > 0$ *excludes compound renewal processes with negative linear drift; see Example 3.4.10.*

In the Lévy case (Proposition 3.4.14), this assumption was only required for the "only if" direction, which is not used elsewhere and serves mainly to show that the synergistic behaviour—namely, that T_x *becomes arbitrarily close to* t *as* $x \to \infty$*—holds only for processes satisfying condition 3.4.1.*

In the present setting, however, this assumption also affects the "if" direction, which is the practically relevant one, and therefore prevents us from treating compound renewal processes with negative drift. This is not a significant limitation, since it is already known that asymptotic equivalence with the supremum fails in the negative drift case; in particular, it does not hold for compound Poisson processes with negative drift and Braverman-light jumps (see Theorem 3.2.5 and Theorem 3.4.17).

It does, however, prevent us from establishing the appropriate form of equivalence that holds in this setting, namely that the tail of the supremum is asymptotically equivalent to the tail of the process evaluated at the last jump before time t*.*

We believe that this condition is an artefact of the proof. In the Lévy case, it can be removed from the "if" direction by introducing an intermediate time point l *with* $s < l < t$*; see the proof of Proposition 3.4.14 for details.*

The next lemma will be pivotal in showing that a compound renewal process with Braverman-light jumps and waiting times satisfying Assumption 4.1.1 is synergistic at extremes. Note that it also implies that a compound Poisson process with Braverman-light jumps is synergistic at extremes. It is inspired by Lemma 5 in [14]. Here we abstract away the part of Braverman's argument that relies on the exponential distribution of the waiting times.

Lemma 4.2.3. *Let $(X_t)_{t\geq 0}$ be a compound renewal process,*

$$X_t = \sum_{k=1}^{N_t} \xi_k, \qquad t \geq 0,$$

where $(\xi_n)_{n\geq 1}$ is an i.i.d. sequence of jumps and $(N_t)_{t\geq 0}$ is the associated renewal counting process. Assume that $(\xi_n)_{n\geq 1}$ is independent of $(N_t)_{t\geq 0}$ and has Braverman-light tails. Suppose that for fixed $0 < s < t$, the sequence

$$a_n(s,t) := \frac{\mathbb{P}(N_s = n)}{\mathbb{P}(N_t = n+2)}, \qquad n \geq 1,$$

is summable, i.e. $\sum_{n=1}^{\infty} a_n(s,t) < \infty$. Then, for any $c > 0$,

$$\lim_{x\to\infty} \frac{\mathbb{P}(X_s > x)}{\mathbb{P}(X_t > x+c)} = 0.$$

Proof. Let $c > 0$ and denote by $S_n := \sum_{k=1}^{n} \xi_k$ the partial sums of the jumps. Conditioning on the number of jumps up to time s gives

$$\begin{aligned} \frac{\mathbb{P}(X_s > x)}{\mathbb{P}(X_t > x+c)} &= \frac{\sum_{n=0}^{\infty} \mathbb{P}(N_s = n)\, \mathbb{P}(S_n > x)}{\mathbb{P}(X_t > x+c)} \\ &\leq \frac{1}{\mathbb{P}(\xi_1 > c)} \sum_{n=0}^{\infty} \frac{\mathbb{P}(N_s = n)}{\mathbb{P}(N_t = n+2)} \cdot \frac{\mathbb{P}(S_n > x)}{\mathbb{P}(S_{n+1} > x)}, \end{aligned} \tag{4.2.5}$$

where in the second inequality we used that, for any $n \geq 0$,

$$\mathbb{P}(X_t > x+c) \;\geq\; \mathbb{P}(N_t = n+2)\, \mathbb{P}(S_{n+1} > x)\, \mathbb{P}(\xi_1 > c)\,.$$

By Lemma 4 of [14], for each fixed n,

$$\lim_{x\to\infty} \frac{\mathbb{P}(S_n > x)}{\mathbb{P}(S_{n+1} > x)} = 0. \tag{4.2.6}$$

Hence, to conclude the proof, it suffices to justify passing the limit $x \to \infty$ inside the sum in (4.2.5).

For $n \geq 1$, note that

$$\mathbb{P}(S_{n+1} > x) \;\geq\; \mathbb{P}(S_n > x)\, \mathbb{P}(\xi_1 > 0)\,,$$

so that

$$\frac{\mathbb{P}(S_n > x)}{\mathbb{P}(S_{n+1} > x)} \;\leq\; \frac{1}{\mathbb{P}(\xi_1 > 0)}.$$

Therefore the summands in (4.2.5) are dominated by

$$\frac{\mathbb{P}(N_s = n)}{\mathbb{P}(N_t = n+2)} \cdot \frac{\mathbb{P}(S_n > x)}{\mathbb{P}(S_{n+1} > x)} \;\leq\; \frac{1}{\mathbb{P}(\xi_1 > 0)} \cdot a_n(s,t),$$

and the dominating sequence $\{a_n(s,t)\}$ is summable by assumption. Thus the Dominated Convergence Theorem applies, allowing the limit to be passed inside the sum, and the claim follows. □

4.3 Extreme Lower-Tail Large Deviations

4.3.1 Background and Statement of Main Result

Large deviations theory is concerned with computing the asymptotic probability of events in which a sum of random variables deviates from its mean on a scale larger than what is predicted by the central limit theorem.

For illustration, let $S_n := \sum_{k=1}^{n} X_k$, where $(X_n)_{n\geq 1}$ is an i.i.d. sequence with $\mathbb{E}[X_1] = \mu < \infty$ and $\mathrm{Var}(X_1) = 1$. The Strong Law of Large Numbers (SLLN) states that

$$\frac{S_n}{n} \xrightarrow{a.s.} \mu, \qquad n \to \infty,$$

so that informally $S_n \sim \mu n$ or, equivalently, $S_n = \mu n + o(n)$ almost surely. In particular, for $x \in \mathbb{R}$, deviations of the form $\{S_n > \mu n + x\sqrt{n}\}$ are called *normal deviations*, and the CLT allows us to compute their probability:

$$\mathbb{P}\big(S_n > \mu n + x\sqrt{n}\big) \;\to\; \overline{\Phi}(x), \qquad n \to \infty.$$

In contrast, for events such as $\{S_n > (\mu + a)n\}$ with $a > 0$, the CLT only tells us that

$$\mathbb{P}(S_n > (\mu + a)n) \to 0, \qquad n \to \infty,$$

which is not very informative. Such events are called *large deviations*, and large deviation theory is concerned with quantifying the rate at which the probability of these rare events decays.

Deviations from the mean may also occur from below, e.g. $\{S_n < (\mu - a)n\}$ for $a > 0$, which are called *lower-tail large deviations*. In our setting, to establish the summability of a_n in Lemma 4.2.3, we need to determine the exact rate at which

$$\mathbb{P}(T_n \leq s)\,,$$

with T_n the n-th renewal time, decays to zero for a fixed constant $s > 0$. Since this is a particularly rare event, lying far below the typical scale, we shall refer to it as an *extreme lower-tail large deviation.*

The main result of this section is the following.

Proposition 4.3.1. *Let $T_n := \sum_{i=1}^n W_i$ be the n-th renewal time, where $(W_i)_{i\geq 1}$ is an i.i.d. sequence of waiting times satisfying Assumption 4.1.1. Then, for any fixed $s > 0$,*

$$\mathbb{P}(T_n \leq s) = \frac{\big(c_1 s^a \Gamma(a)\big)^n}{\Gamma(an+1)} \exp\left(\frac{c_2}{c_1}\frac{\Gamma(a+b)}{\Gamma(a)}\left(\frac{s}{a}\right)^b n^{1-b} + o(n^{1-b})\right)(1+o(1)), \qquad n \to \infty.$$

Remark 4.3.2. *The formula in Proposition 4.3.1 is Gamma-like, in the sense that it coincides with the corresponding formula for Gamma-distributed waiting times, up to a factor of order $e^{O(n^{1-b})}\big(1 + o(1)\big)$.*

To see that, suppose $W_1 \sim \mathrm{Gamma}(a, \lambda)$ with $\lambda > 0$ and, for simplicity, $a \in \mathbb{N}$. Then $T_n \sim \mathrm{Gamma}(an, \lambda)$, that is, T_n has the same distribution as the sum of an i.i.d. exponentials with rate λ. Equivalently, if N_s denotes a Poisson process with rate λ, then

$$\mathbb{P}(T_n \leq s) = \mathbb{P}(N_s \geq an) \sim \mathbb{P}(N_s = an) = \frac{e^{-\lambda s}(\lambda s)^{an}}{(an)!}.$$

Recalling Example 4.1.3, the constants in Assumption 4.1.1 for the Gamma law are $c_1 = \frac{\lambda^a}{\Gamma(a)}$ and $c_2 = -\frac{\lambda^{a+1}}{\Gamma(a)}$, so that $\lambda = -\frac{c_2}{c_1}$. Substituting these constants and using the *definition of the Gamma function in the above expression for Gamma waiting times yields*

$$\mathbb{P}(T_n \le s) = \frac{\left(-\frac{c_1}{c_2}\right)^{an}}{\Gamma(an+1)} \exp\left(\frac{c_1}{c_2} s\right),$$

which matches the formula in Proposition 4.3.1, using in particular that $\Gamma(a) = \frac{\lambda^a}{c_1}$.

The resemblance is no surprise: just as in large deviations above the mean the tails *of the distribution drive the asymptotics, here it is the* behaviour near the origin *that matters. Assumption 4.1.1 postulates such a Gamma-like behaviour of the density near zero. In fact, it also allows for more general Weibull-type behaviour, with the added flexibility of taking* $a \neq b$. *The additional exponential correction factor, compared with the pure Gamma case, arises from this generalisation.*

Remark 4.3.3 (Sharp and extreme lower tail large deviations). *The formula in Proposition 4.3.1 can be written as*

$$\log \mathbb{P}(T_n \le s) = \begin{cases} -an\log n + n(\log A + a - a\log a) + Bn^{1-b} + o(n^{1-b}), & 0 < b < 1, \\ -an\log n + n(\log A + a - a\log a) - \dfrac{1}{2}\log(2\pi a n) + Bn^{1-b} + o(1), & b \ge 1. \end{cases} \tag{4.3.1}$$

where

$$A := c_1 s^a \Gamma(a), \qquad B := \frac{c_2}{c_1}\frac{s^b\Gamma(a+b)}{a^b\Gamma(a)}.$$

To see this, write

$$\mathbb{P}(T_n \le s) = \frac{A^n}{\Gamma(an+1)} \exp\big(Bn^{1-b} + o(n^{1-b})\big)(1 + o(1)),$$

and use Stirling's approximation

$$\log\Gamma(z) = z\log z - z + \frac{1}{2}\log\frac{2\pi}{z} + O(z^{-1}), \qquad z \to \infty,$$

with $z = an+1$. *This gives*

$$\log\Gamma(an+1) = (an+1)\log(an+1) - (an+1) + \frac{1}{2}\log\frac{2\pi}{an+1} + O(n^{-1}),$$

and therefore

$$\begin{aligned}\log \mathbb{P}(T_n \leq s) &= n \log A - \log \Gamma(an+1) + Bn^{1-b} + o(n^{1-b}) + o(1)\\ &= n \log A - (an+1)\log(an+1) + (an+1) - \frac{1}{2}\log\frac{2\pi}{an+1} + Bn^{1-b} + o(n^{1-b}) + o(1).\end{aligned}$$

After simplification this yields (4.3.1).

Hence the proposition provides sharp asymptotics. *They are exact asymptotics (see Definition 2.1.7) when* $b \geq 1$*, while for* $0 < b < 1$ *they should be interpreted as refined logarithmic asymptotics (see Definition 2.1.6). By refined we mean that the expansion reveals several successive scales in the logarithm of the probability: the leading term* $an \log n$*, which drives a superexponential decay, followed by lower–order corrections of order* n*,* $\log n$*, and* n^{1-b}.

Finally, as discussed before Proposition 4.3.1, the event $\{T_n \leq s\}$ *corresponds to an extremely large lower–tail deviation in comparison with classical large deviation events of the form* $\{S_n \geq an\}$*. This phenomenon is illustrated by Proposition 4.3.1.*

In particular, it shows that the probability of the event $\{T_n \leq s\}$ *decays superexponentially, since its leading behaviour is governed by the term* $an \log n$*, so that*

$$\frac{1}{n}\log \mathbb{P}(T_n \leq s) \to -\infty.$$

By contrast, in classical Cramér-type large deviation results (see Example 2.1.12), probabilities of events of the form $\{S_n \geq an\}$ *decay only exponentially, as*

$$\frac{1}{n}\log \mathbb{P}(S_n \geq an) \to -I(a),$$

with $I(a) \in (0, \infty)$.

4.3.2 Exponential tilting

The purpose of this section is to analyze the probability $\mathbb{P}(T_n \leq s)$ via exponential tilting. Since the event $\{T_n \leq s\}$ corresponds to an extreme lower-tail deviation, it is highly unlikely under the original measure. Exponential tilting allows us to transform this rare event into a typical one under a new probability measure, under which the tilted variables have mean s/n.

The tilted and appropriately centered and scaled variables can then be analyzed using central limit theorem–type results—in our case, Edgeworth expansions—yielding non-trivial asymptotic estimates (see the discussion in Subsection 4.3.1).

Preliminaries

For the waiting time W_1 with c.d.f. F we denote its Laplace transform by

$$\psi(\lambda) := \mathbb{E}\left(e^{-\lambda W_1}\right) = \int_0^\infty e^{-\lambda x}\, dF(x). \tag{4.3.2}$$

We define the *negatively exponentially tilted* version of W_1 under parameter $\lambda > 0$, denoted $W_1^{(\lambda)}$, as the random variable with cumulative distribution function

$$F_\lambda(x) := \frac{\int_0^x e^{-\lambda u}\, dF(u)}{\psi(\lambda)}. \tag{4.3.3}$$

For each $p \in \mathbb{N}^+$, the p-th derivative of ψ is given by

$$\psi^{(p)}(\lambda) = (-1)^p \int_0^\infty x^p e^{-\lambda x}\, dF(x), \tag{4.3.4}$$

by the dominated convergence theorem and inequality (4.3.9) below.

More generally,[1] we define the *p-th positive derivative* for any $p \geq 0$ as

$$\psi_+^{(p)}(\lambda) := \int_0^\infty x^p e^{-\lambda x}\, dF(x). \tag{4.3.5}$$

Notice that for all $p \in \mathbb{N}^+$, it holds

$$\psi_+^{(p)}(\lambda) = (-1)^p \psi^{(p)}(\lambda), \tag{4.3.6}$$

that is, the p-th positive derivative coincides with the p-th ordinary derivative up to sign.

Then the p-th moment of the tilted random variable satisfies

$$\mathbb{E}\Big[\Big(W_1^{(\lambda)}\Big)^p\Big] = \frac{\psi_+^{(p)}(\lambda)}{\psi(\lambda)}, \tag{4.3.7}$$

[1] The reason we define $\psi_+^{(p)}(\lambda)$ is to be able to prove results for the p-th moment of $W_1^{(\lambda)}$, for p not necessarily an integer.

and all moments are finite. In particular,

$$\mathbb{E}\left[\left(W_1^{(\lambda)}\right)^p\right] \leq \frac{e^{-p}(p/\lambda)^p}{\psi(\lambda)}, \tag{4.3.8}$$

since

$$\arg\max_{x\geq 0}\left(e^{-\lambda x}x^p\right) = \frac{p}{\lambda}. \tag{4.3.9}$$

Recall that we denote by $T_n = \sum_{k=1}^n W_k$ the partial sum of waiting times. It holds that

$$T_n^{(\lambda)} \stackrel{\mathrm{d}}{=} \sum_{i=1}^n W_i^{(\lambda)}. \tag{4.3.10}$$

To see this, denote as usual by F^{*n} the n-fold convolution of F, and let $s > 0$. By using the definition of the exponentially tilted c.d.f. (4.3.3) and the independence of $(W_k)_{k=1}^n$,

$$\begin{aligned}
\left(F^{*n}\right)_{(\lambda)}(s) &= \frac{\int_0^s e^{-\lambda u}\, dF^{*n}(u)}{\psi(\lambda)^n} \\
&= \frac{1}{\psi(\lambda)^n}\int_0^s dF(x_1)\int_0^{s-x_1} dF(x_2)\cdots\int_0^{s-x_1-\cdots-x_{n-1}} e^{-\lambda(x_1+\cdots+x_n)}\, dF(x_n) \\
&= \int_0^s \frac{e^{-\lambda x_1}}{\psi(\lambda)}\, dF(x_1)\int_0^{s-x_1} \frac{e^{-\lambda x_2}}{\psi(\lambda)}\, dF(x_2)\cdots\int_0^{s-x_1-\cdots-x_{n-1}} \frac{e^{-\lambda x_n}}{\psi(\lambda)}\, dF(x_n) \\
&= \int_0^s dF_\lambda(x_1)\int_0^{s-x_1} dF_\lambda(x_2)\cdots\int_0^{s-x_1-\cdots-x_{n-1}} dF_\lambda(x_n) \\
&= F_\lambda^{*n}(s).
\end{aligned} \tag{4.3.11}$$

Using the definition of the exponentially tilted c.d.f. and the above, we have

$$\begin{aligned}
\mathbb{P}(T_n \leq s) &= \int_{(-\infty,s]} dF^{*n}(u) \\
&= \psi(\lambda)^n \int_{(-\infty,s]} e^{\lambda u}\,\frac{e^{-\lambda u}}{\psi(\lambda)^n}\, dF^{*n}(u) \\
&= \psi(\lambda)^n \int_{(-\infty,s]} e^{\lambda u}\, d\big(F^{*n}\big)_{(\lambda)}(u) \\
&= \psi(\lambda)^n \int_{(-\infty,s]} e^{\lambda u}\, d\big(F_{(\lambda)}^{*n}\big)(u) \\
&= \psi(\lambda)^n\, \mathbb{E}\Big[e^{\lambda T_n^{(\lambda)}}\, \mathbb{1}_{\{T_n^{(\lambda)}\leq s\}}\Big].
\end{aligned} \tag{4.3.12}$$

We define a parameter $\theta_n = \theta_n(s)$ such that

$$\mathbb{E}\left[W_1^{(\theta_n)}\right] = \frac{s}{n}. \tag{4.3.13}$$

Throughout, we suppress the dependence on s in the notation when it is unambiguous.

We then exponentially tilt the first n waiting times $(W_i)_{i\leq n}$ by θ_n, yielding an i.i.d. sequence $(W_i^{(\theta_n)})_{i\leq n}$, forming a row-wise i.i.d. triangular array of random variables. That is, an array whose n-th row consists of n i.i.d. variables (with a distribution depending on n), while different rows may have different distributions.

Due to (4.3.14)

$$T_n^{(\theta_n)} \overset{\mathrm{d}}{=} \sum_{i=1}^{n} W_i^{(\theta_n)}, \tag{4.3.14}$$

and due to (4.3.12)

$$\mathbb{P}(T_n \leq s) = \psi(\theta_n)^n \, \mathbb{E}\left[e^{\theta_n T_n^{(\theta_n)}} \mathbf{1}_{\{T_n^{(\theta_n)} \leq s\}}\right]. \tag{4.3.15}$$

We define the standardized variables:

$$Z_{i,n} := \frac{W_i^{(\theta_n)} - s/n}{\sigma_n}, \qquad S_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{i,n}, \tag{4.3.16}$$

where $\sigma_n := \sqrt{\mathrm{Var}(W_1^{(\theta_n)})}$.

From (4.3.15), a direct rearrangement gives

$$\mathbb{P}(T_n \leq s) = \psi(\theta_n)^n e^{\theta_n s} \, \mathbb{E}\left[e^{\alpha_n S_n} \mathbf{1}_{\{S_n \leq 0\}}\right], \tag{4.3.17}$$

where

$$\alpha_n := \sqrt{n}\, \sigma_n \, \theta_n. \tag{4.3.18}$$

We prove Proposition 4.3.1 via equation (4.3.17). The following subsections, together with the next section, are devoted to analyzing the asymptotic behavior of the three components on the right-hand side:

- the Laplace transform term $\psi(\theta_n)^n$,
- the exponential factor $e^{\theta_n s}$,
- the integral $\mathbb{E}\left[e^{\alpha_n S_n} \, \mathbf{1}_{\{S_n \leq 0\}}\right]$.

Asymptotic behaviour of the Laplace transform and tilted moments

Lemma 4.3.4. *Under Assumption 4.1.1, for any $p \geq 0$, the positive derivative of the Laplace transform satisfies*

$$\psi_+^{(p)}(\lambda) = c_1\Gamma(a+p)\,\lambda^{-(a+p)} + c_2\Gamma(a+b+p)\,\lambda^{-(a+b+p)} + o\big(\lambda^{-(a+b+p)}\big)\,, \qquad \lambda \to \infty. \quad (4.3.19)$$

Proof. By Assumption 4.1.1, there exists $\rho \in (0, \infty]$ such that F is absolutely continuous on $(0, \rho)$ (see Remark 4.1.2). Fix $\delta < \rho$. We split the integral:

$$\psi_+^{(p)}(\lambda) = \int_0^\delta x^p e^{-\lambda x}\,dF(x) + \int_\delta^\infty x^p e^{-\lambda x}\,dF(x) =: I_1(\lambda) + I_2(\lambda).$$

Since the function $m_\lambda(x) := x^p e^{-\lambda x}$ attains its maximum at $x = p/\lambda$ and is decreasing on $(p/\lambda, \infty)$, it follows that for $\lambda > p/\delta$,

$$x^p e^{-\lambda x} \leq \delta^p e^{-\lambda\delta}, \qquad x \geq \delta. \quad (4.3.20)$$

Therefore, for the tail term we obtain

$$I_2(\lambda) \leq \delta^p e^{-\lambda\delta}\,\mathbb{P}(W_1 > \delta) = O(e^{-\lambda\delta}) = o\big(\lambda^{-(a+p+1)}\big)\,. \quad (4.3.21)$$

For the main term, we use Assumption 4.1.1 to write

$$I_1(\lambda) = \int_0^\delta x^p e^{-\lambda x} x^{a-1}\big(c_1 + c_2 x^b + o(x^b)\big)\,dx.$$

Denote by $\gamma(\cdot,\cdot)$ the lower incomplete gamma function. Then

$$\int_0^\delta x^k e^{-\lambda x}\,dx = \lambda^{-(k+1)}\gamma(k+1, \lambda\delta),$$

and notice that $\gamma(k, \lambda\delta) \to \Gamma(k)$ as $\lambda \to \infty$. Therefore,

$$\begin{aligned} I_1(\lambda) &= c_1\gamma(a+p, \lambda\delta)\,\lambda^{-(a+p)} + c_2\gamma(a+b+p, \lambda\delta)\,\lambda^{-(a+b+p)} + o\big(\lambda^{-(a+b+p)}\big) \\ &= c_1\Gamma(a+p)\,\lambda^{-(a+p)} + c_2\Gamma(a+b+p)\,\lambda^{-(a+b+p)} + o\big(\lambda^{-(a+b+p)}\big)\,. \end{aligned} \quad (4.3.22)$$

The claim follows by combining (4.3.21) and (4.3.22). □

Corollary 4.3.5. *Under Assumption 4.1.1, the following asymptotic expansions hold as* $\lambda \to \infty$*:*

(i)

$$\psi(\lambda) = c_1\Gamma(a)\lambda^{-a} + c_2\Gamma(a+b)\lambda^{-(a+b)} + o\big(\lambda^{-(a+b)}\big)\,.$$

(ii)

$$\begin{aligned}\mathbb{E}\Big[(W_1^{(\lambda)})^p\Big] &= \frac{\Gamma(a+p)}{\Gamma(a)}\lambda^{-p} \\ &\quad + \frac{c_2}{c_1}\left(\frac{\Gamma(a+b+p)}{\Gamma(a)} - \frac{\Gamma(a+p)\Gamma(a+b)}{\Gamma(a)^2}\right)\lambda^{-(p+b)} \\ &\quad + o\big(\lambda^{-(p+b)}\big)\,.\end{aligned}$$

(iii)

$$\mathrm{Var}(W_1^{(\lambda)}) = a\lambda^{-2} + \frac{c_2}{c_1}\cdot\frac{\Gamma(a+b)}{\Gamma(a)}\cdot b(b+1)\lambda^{-(2+b)} + o\big(\lambda^{-(2+b)}\big)\,.$$

Proof. **(i)** This follows from Lemma 4.3.4 with $p = 0$.

(ii) Recall that,

$$\mathbb{E}\Big[(W_1^{(\lambda)})^p\Big] = \frac{\psi(\lambda)_+^{(p)}}{\psi(\lambda)}.$$

Using the asymptotic expansions from Lemma 4.3.4, we get:

$$\mathbb{E}\Big[(W_1^{(\lambda)})^p\Big] = \frac{c_1\Gamma(a+p)\lambda^{-(a+p)} + c_2\Gamma(a+b+p)\lambda^{-(a+b+p)} + o\big(\lambda^{-(a+b+p)}\big)}{c_1\Gamma(a)\lambda^{-a} + c_2\Gamma(a+b)\lambda^{-(a+b)} + o\big(\lambda^{-(a+b)}\big)}.$$

Let us denote:

$$A = c_1\Gamma(a+p), \quad B = c_2\Gamma(a+b+p), \quad C = c_1\Gamma(a), \quad D = c_2\Gamma(a+b).$$

Then:

$$\mathbb{E}\Big[(W_1^{(\lambda)})^p\Big] = \lambda^{-p}\cdot\frac{A + B\lambda^{-b} + o(\lambda^{-b})}{C + D\lambda^{-b} + o(\lambda^{-b})}.$$

We apply a Taylor expansion on the function $f(x) = \frac{A+Bx}{C+Dx}$ around $x = 0$. Then:

$$f(x) = \frac{A}{C} + \frac{BC - AD}{C^2}x + o(x), \quad \text{as } x \to 0.$$

Applying this with $x = \lambda^{-b}$, we find:

$$\mathbb{E}\Big[(W_1^{(\lambda)})^p\Big] = \lambda^{-p}\left(\frac{\Gamma(a+p)}{\Gamma(a)} + \frac{c_2}{c_1}\left(\frac{\Gamma(a+b+p)}{\Gamma(a)} - \frac{\Gamma(a+p)\Gamma(a+b)}{\Gamma(a)^2}\right)\lambda^{-b} + o(\lambda^{-b})\right).$$

(iii) From (ii) with $p = 2$, we have:

$$\mathbb{E}\Big[(W_1^{(\lambda)})^2\Big] = \frac{\Gamma(a+2)}{\Gamma(a)}\lambda^{-2} + \frac{c_2}{c_1}\left(\frac{\Gamma(a+b+2)}{\Gamma(a)} - \frac{\Gamma(a+2)\Gamma(a+b)}{\Gamma(a)^2}\right)\lambda^{-(2+b)} + o(\lambda^{-(2+b)}).$$

Applying the identity $\Gamma(k+1) = k\Gamma(k)$, we simplify:

$$\mathbb{E}\left[\left(W_1^{(\lambda)}\right)^2\right] = a(a+1)\lambda^{-2} + \frac{c_2}{c_1}\cdot\frac{\Gamma(a+b)}{\Gamma(a)}(2ab + b^2 + b)\lambda^{-(2+b)} + o(\lambda^{-(2+b)}). \quad (4.3.23)$$

Now compute $\mathbb{E}\Big[W_1^{(\lambda)}\Big]^2$ using the result of part (ii) with $p = 1$:

$$\mathbb{E}\Big[W_1^{(\lambda)}\Big] = a\lambda^{-1} + \frac{c_2}{c_1}\cdot\frac{\Gamma(a+b)}{\Gamma(a)}b\lambda^{-(1+b)} + o(\lambda^{-(1+b)}). \quad (4.3.24)$$

Squaring (4.3.24), we set:

$$C := a, \quad D := \frac{c_2}{c_1}\cdot\frac{\Gamma(a+b)}{\Gamma(a)}b,$$

so that

$$\mathbb{E}\Big[W_1^{(\lambda)}\Big]^2 = \big(\lambda^{-1}(C + D\lambda^{-b} + o(\lambda^{-b}))\big)^2 = C^2\lambda^{-2} + 2CD\lambda^{-(2+b)} + o(\lambda^{-(2+b)}). \quad (4.3.25)$$

Hence, the variance is:

$$\begin{aligned}
\mathrm{Var}(W_1^{(\lambda)}) &= \mathbb{E}\Big[(W_1^{(\lambda)})^2\Big] - \mathbb{E}\Big[W_1^{(\lambda)}\Big]^2 \\
&= \big(a(a+1) - a^2\big)\lambda^{-2} \\
&\quad + \left(\frac{c_2}{c_1}\cdot\frac{\Gamma(a+b)}{\Gamma(a)}(2ab + b^2 + b) - 2ab\cdot\frac{c_2}{c_1}\cdot\frac{\Gamma(a+b)}{\Gamma(a)}\right)\lambda^{-(2+b)} \qquad (4.3.26)\\
&\quad + o\big(\lambda^{-(2+b)}\big) \\
&= a\lambda^{-2} + \frac{c_2}{c_1}\cdot\frac{\Gamma(a+b)}{\Gamma(a)}b(b+1)\lambda^{-(2+b)} + o\big(\lambda^{-(2+b)}\big).
\end{aligned}$$

which completes the proof. □

Existence and Asymptotics of the Tilting Parameter θ_n

Recall that the parameter $\theta_n(s)$ is defined as the solution to:

$$\mathbb{E}\Big[W_1^{(\theta_n(s))}\Big] = \frac{s}{n} \tag{4.3.27}$$

Lemma 4.3.6. *Under Assumption 4.1.1, the following hold:*

(i) The equation (4.3.27) *has a unique solution $\theta_n(s)$ for each $s > 0$ and $n \in \mathbb{N}^+$.*

(ii) As $n \to \infty$, we have

$$\theta_n(s) = \frac{a}{s}n + \frac{c_2}{c_1} \cdot \frac{s^{b-1} b\Gamma(a+b)}{a^b\Gamma(a)} n^{1-b} + o(n^{1-b}). \tag{4.3.28}$$

Proof. Let $h : (0, \infty) \to (0, \mathbb{E}[W_1])$ be defined as

$$h(\lambda) := \mathbb{E}\Big[W_1^{(\lambda)}\Big] = \frac{\psi'_+(\lambda)}{\psi(\lambda)} = \frac{-\psi'(\lambda)}{\psi(\lambda)}, \tag{4.3.29}$$

where the second and third equalities are the relationships (4.3.7) and (4.3.6) for $p = 1$.

(i) Existence and uniqueness of $\theta_n(s)$:

We show that h is continuous, strictly decreasing, and satisfies $\lim_{\lambda\to\infty} h(\lambda) = 0$. This implies that the inverse function h^{-1} exists and is well-defined on $(0, \mathbb{E}[W_1])$, ensuring that the equation $h(\theta_n(s)) = s/n$ has a unique solution.

Continuity: Since $W_1 \geq 0$ and $x^k e^{-\lambda x}$ is bounded on $[0, \infty)$ for fixed $\lambda > 0$, differentiation under the integral is justified by the dominated convergence theorem, ensuring $\psi(\lambda)$ is C^∞ and hence h is C^∞ (see [26, Ch. XIII.2]).

Monotonicity: Differentiating gives:

$$\begin{aligned} h'(\lambda) &= \frac{d}{d\lambda}\left(\frac{-\psi'(\lambda)}{\psi(\lambda)}\right) \\ &= \frac{\psi'(\lambda)^2 - \psi''(\lambda)\psi(\lambda)}{\psi(\lambda)^2} \\ &= \mathbb{E}\Big[W_1^{(\lambda)}\Big]^2 - \mathbb{E}\Big[\Big(W_1^{(\lambda)}\Big)^2\Big] \\ &= -\operatorname{Var}\Big(W_1^{(\lambda)}\Big) < 0. \end{aligned} \tag{4.3.30}$$

where the third equality follows from (4.3.7), and the strict inequality holds because the variance is strictly positive for a non-degenerate random variable.

Limit: By Corollary 4.3.5(ii), for $p = 1$, it follows that $h(\lambda) \to 0$.

(ii) Asymptotic expansion of $\theta_n(s)$:

By the previous part, the function h is strictly decreasing and bijective, so its inverse

$$h^{-1} : (0, \mathbb{E}[W_1]) \to (0, \infty)$$

exists and satisfies

$$\theta_n(s) = h^{-1}(s/n).$$

From Corollary 4.3.5 with $p = 1$, we have as $\lambda \to \infty$:

$$h(\lambda) = \frac{\Gamma(a+1)}{\Gamma(a)}\lambda^{-1} + \frac{c_2}{c_1}\left(\frac{\Gamma(a+b+1)}{\Gamma(a)} - \frac{\Gamma(a+1)\Gamma(a+b)}{\Gamma(a)^2}\right)\lambda^{-(1+b)} + o\left(\lambda^{-(1+b)}\right).$$

Since $\Gamma(a+1)/\Gamma(a) = a$, this simplifies to

$$h(\lambda) = A\lambda^{-1} + B\lambda^{-(1+b)} + o(\lambda^{-(1+b)}),$$

where

$$A := a, \qquad B := \frac{c_2}{c_1} \cdot \frac{b\,\Gamma(a+b)}{\Gamma(a)}.$$

Applying Lemma A.2.2- from Appendix- with these constants gives, as $x \to 0^+$:

$$h^{-1}(x) = \frac{a}{x} + \frac{c_2}{c_1} \cdot \frac{b\,\Gamma(a+b)}{a^b\,\Gamma(a)}\, x^{b-1} + o\left(x^{b-1}\right).$$

Finally, substituting $x = s/n$ yields

$$\theta_n(s) = \frac{a}{s}n + \frac{c_2}{c_1} \cdot \frac{b\,\Gamma(a+b)}{a^b\,\Gamma(a)} s^{b-1} n^{1-b} + o\left(n^{1-b}\right), \quad \text{as } n \to \infty.$$

□

Using the approximation of $\theta_n(s)$ from Lemma 4.3.6 yields:

Corollary 4.3.7. *Under Assumption 4.1.1, we have the following leading-order asymptotic expansions as $n \to \infty$:*

(i)

$$\psi(\theta_n(s)) = c_1\Gamma(a)\left(\frac{s}{a}\right)^a n^{-a} + (1-b)c_2\Gamma(a+b)\left(\frac{s}{a}\right)^{a+b} n^{-(a+b)} + o\left(n^{-(a+b)}\right), \tag{4.3.31}$$

(ii)

$$\mathbb{E}\Big[(W_1^{(\theta_n(s))})^p\Big] = \frac{\Gamma(a+p)}{\Gamma(a)}\left(\frac{s}{a}\right)^p n^{-p} + O\left(n^{-(p+b)}\right), \tag{4.3.32}$$

(iii)

$$\mathrm{Var}(W_1^{(\theta_n(s))}) = \frac{s^2}{a}n^{-2} + O\left(n^{-(2+b)}\right). \tag{4.3.33}$$

Proof. **(i)** Recall from Lemma 4.3.6(ii) that

$$\theta_n(s) = \frac{a}{s}n + \frac{c_2}{c_1}\frac{s^{b-1}b\Gamma(a+b)}{a^b\Gamma(a)}n^{1-b} + o(n^{1-b}), \qquad n\to\infty.$$

We also have from Corollary 4.3.5(i) that

$$\psi(\lambda) = c_1\Gamma(a)\lambda^{-a} + c_2\Gamma(a+b)\lambda^{-(a+b)} + o\big(\lambda^{-(a+b)}\big), \qquad \lambda\to\infty.$$

Let us define

$$L_n := \frac{a}{s}n, \qquad \delta_n := \frac{c_2}{c_1}\cdot\frac{s^b b\Gamma(a+b)}{a^{b+1}\Gamma(a)}n^{-b} + o(n^{-b}), \tag{4.3.34}$$

so that

$$\theta_n(s) = L_n(1+\delta_n).$$

Then $\delta_n = O(n^{-b})$, and for every fixed $c>0$,

$$\theta_n(s)^{-c} = L_n^{-c}(1+\delta_n)^{-c} = L_n^{-c}\left(1 - c\delta_n + o(n^{-b})\right). \tag{4.3.35}$$

Substituting this into the expansion of ψ, we obtain

$$\begin{aligned}\psi(\theta_n(s)) &= c_1\Gamma(a)L_n^{-a}\left(1 - a\delta_n + o(n^{-b})\right)\\ &\quad + c_2\Gamma(a+b)L_n^{-(a+b)}\left(1 - (a+b)\delta_n + o(n^{-b})\right) + o(n^{-(a+b)}).\end{aligned}$$

Since $L_n^{-(a+b)}\delta_n = O(n^{-(a+2b)})$, the term involving $-(a+b)\delta_n$ in the second line is

$O(n^{-(a+2b)})$, and hence is $o(n^{-(a+b)})$. Therefore,

$$\psi(\theta_n(s)) = c_1\Gamma(a)L_n^{-a} - ac_1\Gamma(a)L_n^{-a}\delta_n + c_2\Gamma(a+b)L_n^{-(a+b)} + o(n^{-(a+b)}).$$

We now compute the two terms of order $n^{-(a+b)}$. First,

$$-ac_1\Gamma(a)L_n^{-a}\delta_n = -ac_1\Gamma(a)\left(\frac{s}{a}\right)^a n^{-a}\left[\frac{c_2}{c_1}\frac{s^b b\Gamma(a+b)}{a^{b+1}\Gamma(a)}n^{-b} + o(n^{-b})\right].$$

Thus

$$-ac_1\Gamma(a)L_n^{-a}\delta_n = -bc_2\Gamma(a+b)\left(\frac{s}{a}\right)^{a+b} n^{-(a+b)} + o(n^{-(a+b)}).$$

Moreover,

$$c_2\Gamma(a+b)L_n^{-(a+b)} = c_2\Gamma(a+b)\left(\frac{s}{a}\right)^{a+b} n^{-(a+b)}.$$

Combining the two contributions gives

$$\psi(\theta_n(s)) = c_1\Gamma(a)\left(\frac{s}{a}\right)^a n^{-a} + (1-b)c_2\Gamma(a+b)\left(\frac{s}{a}\right)^{a+b} n^{-(a+b)} + o(n^{-(a+b)}).$$

(ii) The argument proceeds in the same way as in part (i). Recall from Corollary 4.3.5:

$$\mathbb{E}\Big[(W_1^{(\lambda)})^p\Big] = \frac{\Gamma(a+p)}{\Gamma(a)}\lambda^{-p} + \frac{c_2}{c_1}\left(\frac{\Gamma(a+b+p)}{\Gamma(a)} - \frac{\Gamma(a+p)\Gamma(a+b)}{\Gamma(a)^2}\right)\lambda^{-(p+b)} + o(\lambda^{-(p+b)}).$$

Apply this with $\lambda = \theta_n(s) = L_n(1+\delta_n)$, where L_n, δ_n are defined again as in (4.3.34). Using again the expansion (4.3.35), now for $c = p$, we obtain:

$$\begin{aligned}\mathbb{E}\Big[(W_1^{(\theta_n(s))})^p\Big] &= \frac{\Gamma(a+p)}{\Gamma(a)}L_n^{-p}(1-p\delta_n + o(n^{-b})) \\ &\quad + \frac{c_2}{c_1}\left(\frac{\Gamma(a+b+p)}{\Gamma(a)} - \frac{\Gamma(a+p)\Gamma(a+b)}{\Gamma(a)^2}\right)L_n^{-(p+b)}(1-(p+b)\delta_n + o(n^{-b})) \\ &= \frac{\Gamma(a+p)}{\Gamma(a)}L_n^{-p} + O(n^{-(p+b)}),\end{aligned}$$

which proves the desired result.

(iii) Follows similarly by plugging the expansion of $\theta_n(s)$ into the variance formula from Corollary 4.3.5 (iii), and using the decomposition

$$\mathrm{Var}(W_1^{(\theta_n(s))}) = \mathbb{E}\Big[(W_1^{(\theta_n(s))})^2\Big] - \mathbb{E}\Big[W_1^{(\theta_n(s))}\Big]^2.$$

□

4.3.3 Normal approximation

In this section, the goal is to approximate the exponential integral appearing in Equation (4.3.17), namely,

$$\mathbb{E}\Big[e^{\alpha_n(s)S_n(s)} \cdot \mathbf{1}_{\{S_n(s)\leq 0\}}\Big] .$$

To achieve this, we would like to take advantage of the fact that $S_n(s)$ converges weakly to a standard normal random variable $Z \sim \mathcal{N}(0,1)$.

However, since the function $f_n(x) = e^{\alpha_n(s)x} \cdot \mathbf{1}_{\{x\leq 0\}}$ depends on n, we cannot directly apply the Portmanteau theorem to justify convergence of expectations. Therefore, we require a stronger result—specifically, an Edgeworth expansion for triangular arrays of random variables.

An *Edgeworth expansion* is a refinement of the central limit theorem (CLT) which approximates the distribution function (or density) of a standardized sum of random variables by the standard normal law plus explicit correction terms involving higher-order cumulants. The first-order expansion incorporates the skewness term, improving the normal approximation to an error of order $O(n^{-1/2})$, and that is what we will need. Higher-order expansions include further terms based on cumulants of order four and above, yielding progressively more accurate approximations.

An Edgeworth expansions of arbitrary order for triangular arrays, depending on the highest bounded moment appears in [28, Theorem 2.1], which follows the i.i.d. treatment of [8]. Since the assumptions there do not fit our setting, and because we only require the first-order expansion, we instead establish our own version for triangular arrays by adapting the proof of the i.i.d. case in [6, Theorem 5.22].

The only substantive difference from the i.i.d. case is that here we must handle a *sequence* of characteristic functions, one for each row of the array, rather than a single fixed function. Apart from this, the argument is essentially identical to that in [6, Theorem 5.22]. For clarity, the proof of Theorem 4.3.8 is deferred to Appendix A A.1, so as not to interrupt the main line of exposition.

Theorem 4.3.8 (Edgeworth Expansion for Triangular Arrays)**.** *Let* $(X_{i,n})_{1\leq i\leq n}$ *be a row-*

wise i.i.d. triangular array of standardised random variables, that is,

$$\mathbb{E}[X_{1,n}] = 0, \qquad \mathbb{E}\big[X_{1,n}^2\big] = 1.$$

Denote

$$u_n(t) := \mathbb{E}\big[e^{itX_{1,n}}\big], \qquad a_{k,n} := \mathbb{E}\big[X_{1,n}^k\big], \qquad m_{k,n} := \mathbb{E}\big[|X_{1,n}|^k\big].$$

Assume the following conditions.

(a) For every $0 < \delta < K < \infty$, there exist $r_{\delta,K} \in (0,1)$ and $n_{\delta,K} \in \mathbb{N}$ such that, for all $n \geq n_{\delta,K}$,

$$\sup_{\delta \leq |t| \leq K} |u_n(t)| \leq r_{\delta,K}.$$

(b) There exists a constant $M_4 > 0$ such that, for all n,

$$m_{4,n} = \mathbb{E}\big[|X_{1,n}|^4\big] \leq M_4.$$

(c) There exist $\delta_0 > 0$, $c_{\delta_0} > 0$, and $n_0 \in \mathbb{N}$ such that, for all $n \geq n_0$,

$$\inf_{|t|<\delta_0} |u_n(t)| \geq c_{\delta_0}.$$

Let

$$S_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_{i,n},$$

and let F_n denote the c.d.f. of S_n. Then, uniformly in $x \in \mathbb{R}$,

$$F_n(x) - \Phi(x) = \frac{a_{3,n}}{6\sqrt{n}}(1 - x^2)\phi(x) + o(n^{-1/2}),$$

where Φ and ϕ denote the standard normal c.d.f. and p.d.f., respectively.

Next, we justify that the standardized random variables defined in (4.3.16) satisfy the conditions of Theorem 4.3.8. We begin by showing, via the method of moments, that these variables converge in distribution to a standardized Gamma random variable.

A distribution is said to be *determined by its moments* if no other distribution shares the same sequence of moments.

Lemma 4.3.9. *Let $s > 0$ and let $(Z_{i,n}(s))_{i \leq n}$ be the triangular array of standardized random variables defined in* (4.3.16)*:*

$$Z_{i,n} := \frac{W_i^{(\theta_n)} - s/n}{\sigma_n}, \qquad \sigma_n := \sqrt{\mathrm{Var}\Big(W_1^{(\theta_n)}\Big)}.$$

Then:

(i) *For any $p \in \mathbb{N}$, the p-th moment of $Z_{1,n}(s)$ satisfies*

$$\mathbb{E}\left[Z_{1,n}^p\right] = a^{p/2} \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} \frac{\Gamma(a+k)}{\Gamma(a)} \cdot \frac{1}{a^k} + o(1), \qquad n \to \infty.$$

(ii) *Let Z be the standardised Gamma variable*

$$Z := \frac{X - \mathbb{E}[X]}{\sqrt{\mathrm{Var}(X)}},$$

where X is a Gamma random variable with shape parameter a[2]*. Then*

$$\mathbb{E}[Z^p] = a^{p/2} \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} \frac{\Gamma(a+k)}{\Gamma(a)} \cdot \frac{1}{a^k}.$$

(iii) *The distribution of Z is uniquely determined by its moments.*

(iv) *We have*

$$Z_{1,n}(s) \xrightarrow{d} Z \quad \text{as } n \to \infty.$$

Proof. **(i)** By the binomial theorem, we have:

$$\mathbb{E}\left[Z_{1,n}^p\right] = \frac{1}{\sigma_n^p} \sum_{k=0}^{p} \binom{p}{k} \left(-\frac{s}{n}\right)^{p-k} \mathbb{E}\left[\left(W^{(\theta_n(s))}\right)^k\right].$$

From Corollary 4.3.7(ii), we know:

$$\mathbb{E}\left[\left(W^{(\theta_n(s))}\right)^k\right] = \frac{\Gamma(a+k)}{\Gamma(a)} \left(\frac{s}{a}\right)^k n^{-k} + O(n^{-(k+b)}), \quad \sigma_n^2 = \frac{s^2}{a} n^{-2} + O(n^{-(2+b)}).$$

Taking square roots and applying a first-order expansion, we obtain

$$\sigma_n = \sqrt{\sigma_n^2} = \frac{s}{\sqrt{a}} n^{-1} + O(n^{-(1+b)}). \tag{4.3.36}$$

Now, raising this expansion to the power $p \in \mathbb{N}$, we use

$$\sigma_n^p = \left(\frac{s}{\sqrt{a}} n^{-1} + O(n^{-(1+b)})\right)^p = \left(\frac{s}{\sqrt{a}}\right)^p n^{-p} \left(1 + O(n^{-b})\right)^p.$$

[2]The rate parameter λ disappears under standardisation.

Expanding the last factor using the binomial (or Taylor) expansion, we get

$$\big(1+O(n^{-b})\big)^p = 1 + p\cdot O(n^{-b}) + o(n^{-b}) = 1 + O(n^{-b}),$$

so that finally,

$$\sigma_n^p = \left(\frac{s}{\sqrt{a}}\right)^p n^{-p} + O(n^{-(p+b)}). \tag{4.3.37}$$

Therefore,

$$\sigma_n^p = \left(\frac{s^2}{a}\right)^{p/2} n^{-p} + o(n^{-p}).$$

Substituting into the expression for $\mathbb{E}\big[Z_{1,n}^p\big]$, we obtain:

$$\begin{aligned}\mathbb{E}\big[Z_{1,n}^p\big] &= \frac{1}{\left(\frac{s^2}{a}\right)^{p/2} n^{-p}} \sum_{k=0}^{p} \binom{p}{k} \left(-\frac{s}{n}\right)^{p-k} \left(\frac{\Gamma(a+k)}{\Gamma(a)} \left(\frac{s}{a}\right)^k n^{-k} + O(n^{-(k+b)})\right)\\ &= a^{p/2} \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} \frac{\Gamma(a+k)}{\Gamma(a)} \cdot \frac{1}{a^k} + o(1).\end{aligned}$$

(ii) Fix a rate parameter $\lambda > 0$ and assume that $X \sim \mathit{Gamma}(a, \lambda)$. By applying the binomial theorem exactly as in part (i), and using that the moments of a Gamma variable with shape a and rate λ are given by

$$\mathbb{E}\big[X^k\big] = \frac{\Gamma(a+k)}{\lambda^k \Gamma(a)},$$

and that

$$\mathbb{E}[X] = \frac{a}{\lambda}, \qquad \mathrm{Var}(X) = \frac{a}{\lambda^2},$$

we can write

$$\mathbb{E}[Z^p] = \left(\frac{\lambda^2}{a}\right)^{p/2} \sum_{k=0}^{p} \binom{p}{k} \left(-\frac{a}{\lambda}\right)^{p-k} \cdot \frac{\Gamma(a+k)}{\lambda^k \Gamma(a)}.$$

Combining powers of λ and simplifying:

$$\mathbb{E}[Z^p] = a^{-p/2} \sum_{k=0}^{p} \binom{p}{k} (-a)^{p-k} \cdot \frac{\Gamma(a+k)}{\Gamma(a)} \cdot \lambda^{p-k-k-p} = a^{p/2} \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} \frac{\Gamma(a+k)}{\Gamma(a)} \cdot \frac{1}{a^k}.$$

(iii) A sufficient condition for Z to be completely determined by its moments is that all moments $\mathbb{E}[Z^p]$ are finite and the power series

$$\sum_{p=0}^{\infty} \frac{\mathbb{E}[Z^p]\, r^p}{p!}$$

has a positive radius of convergence (see Theorem 30.1 in [9]). This condition is satisfied for Z due to the light-tailedness of the Gamma distribution.

In particular, Z has finite moments by part (ii). In addition, recall that $X \sim \text{Gamma}(a, \lambda)$ (for some fixed rate $\lambda > 0$). Its moment generating function is

$$M_X(t) = \mathbb{E}\big[e^{tX}\big] = (1 - \tfrac{t}{\lambda})^{-a}, \qquad t < \lambda.$$

We have

$$Z = \frac{X - \mathbb{E}[X]}{\sqrt{\mathrm{Var}(X)}} = \frac{\lambda X - a}{\sqrt{a}}$$

Then, for $t \in \mathbb{R}$ with $t < \sqrt{a}$,

$$M_Z(t) = \mathbb{E}\Big[e^{t(\lambda X - a)/\sqrt{a}}\Big] = e^{-t\sqrt{a}}\, \mathbb{E}\Big[e^{(\lambda t/\sqrt{a})X}\Big] = e^{-t\sqrt{a}}\, M_X\Big(\frac{\lambda t}{\sqrt{a}}\Big) = e^{-t\sqrt{a}}\Big(1 - \frac{t}{\sqrt{a}}\Big)^{-a}.$$

Hence, $M_Z(t)$ exists and is finite in a neighborhood of $t = 0$ (in fact, for all $t < \sqrt{a}$). In this neighborhood, the expansion

$$M_Z(t) = \sum_{p=0}^{\infty} \frac{\mathbb{E}[Z^p]\, t^p}{p!}$$

follows from the power series definition of the exponential function together with the finiteness of $M_Z(t)$ and the dominated convergence theorem. Therefore, Z is determined by its moments.

(iv) The convergence in distribution now follows from the method of moments: by (i) and (ii) the moments of $Z_{1,n}$ converge to those of Z, and by (iii) the law of Z is determined by its moments. Hence,

$$Z_{1,n} \xrightarrow{d} Z.$$

□

Corollary 4.3.10. *Let $s > 0$. The triangular array of random variables $(Z_{i,n}(s))_{i \le n}$,* *defined in* (4.3.16)*, satisfies the conditions of Theorem 4.3.8.*

Proof. We verify each condition in turn.

Condition (a). By Lemma 4.3.9 (iv), there exists a non-lattice random variable Z, such that $Z_{1,n} \xrightarrow{w} Z$. Hence $u_n(t) \to u(t)$ uniformly on compact sets, where $u(t)$ is the characteristic function of Z.

Fix $0 < \delta < K < \infty$. Since Z is a non-lattice random variable, for any fixed $\delta > 0$, there exists $\rho_{\delta,K} < 1$ such that

$$|u(t)| < \rho_{\delta,K} \quad \text{for all } t \in [-K, -\delta] \cup [\delta, K].$$

Let $\ell > 0$ such that $\ell < 1 - \rho_{\delta,K}$. By uniform convergence on compact sets, for all sufficiently large n,

$$|u_n(t) - u(t)| < \ell \quad \text{on } [-K, -\delta] \cup [\delta, K].$$

Hence, for all sufficiently large n,

$$|u_n(t)| \leq |u_n(t) - u(t)| + |u(t)| < \ell + \rho_{\delta,K} =: r_{\delta,K} < 1 \quad \text{for all } t \in [-K, -\delta] \cup [\delta, K].$$

Condition (b). This follows directly from Lemma 4.3.9 (i).

Condition (c). We use again that $u_n(t) \to u(t)$ uniformly on compact sets. Since characteristic functions are continuous with $u(0) = 1$, it follows that for any $\epsilon > 0$, there exists $\delta > 0$ such that

$$|u(t)| > 1 - \epsilon \quad \text{for all } |t| < \delta.$$

By uniform convergence on compact sets, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ and all $|t| < \delta$,

$$|u_n(t) - u(t)| < \epsilon.$$

Hence, for $n \geq n_0$,

$$|u_n(t)| \geq |u(t)| - |u_n(t) - u(t)| > 1 - 2\epsilon \quad \text{for all } |t| < \delta.$$

Setting $c_\delta := 1 - 2\epsilon$, we conclude that condition (c) holds: for δ small enough

$$\inf_{|t|<\delta} |u_n(t)| \geq c_\delta > 0 \quad \text{for every sufficiently large } n.$$

□

We conclude the section with the normal approximation of the exponential integral in formula (4.3.17).

Lemma 4.3.11. *Under Assumption* 4.1.1, *and with* $S_n(s)$ *and* $\alpha_n(s)$ *defined in* (4.3.16) *and* (4.3.18), *respectively, we have*

$$\mathbb{E}\big[e^{\alpha_n(s)S_n(s)} \cdot \mathbf{1}_{\{S_n(s)\leq 0\}}\big] = \frac{1}{\sqrt{2\pi na}}\big(1 + o(1)\big), \quad \textit{as } n \to \infty. \tag{4.3.38}$$

Proof. We proceed in two steps.

Step 1 (Gaussian benchmark). Let $Z \sim \mathcal{N}(0, 1)$. We first show that

$$\mathbb{E}\big[e^{\alpha_n(s)Z}\mathbf{1}_{\{Z\leq 0\}}\big] \sim \frac{1}{\sqrt{2\pi na}}, \quad n \to \infty. \tag{4.3.39}$$

From Corollary 4.3.7 and Lemma 4.3.6(ii),

$$\alpha_n(s) := \sqrt{n}\sigma_n\theta_n(s) = \sqrt{na}\,\big(1 + o(1)\big) \to \infty. \tag{4.3.40}$$

Direct computation gives

$$\begin{aligned}
\mathbb{E}\big[e^{\alpha_n Z}\mathbf{1}_{\{Z\leq 0\}}\big] &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0} e^{\alpha_n x}e^{-x^2/2}\,dx \\
&= \frac{1}{\sqrt{2\pi}}\int_{0}^{\infty} e^{-u^2/2-\alpha_n u}\,du \quad (u = -x) \\
&= e^{\alpha_n^2/2} \cdot \overline{\Phi}(\alpha_n),
\end{aligned}$$

and using the standard normal tail asymptotic $\overline{\Phi}(x) \sim \frac{e^{-x^2/2}}{x\sqrt{2\pi}}$ as $x \to \infty$,

$$\mathbb{E}\big[e^{\alpha_n Z}\mathbf{1}_{\{Z\leq 0\}}\big] \sim \frac{1}{\alpha_n\sqrt{2\pi}} \sim \frac{1}{\sqrt{2\pi na}}.$$

Step 2 (Comparison with $S_n(s)$). By the Step 1, it suffices to show that

$$\lim_{n\to\infty} \frac{\mathbb{E}\big[e^{\alpha_n(s)S_n(s)}\mathbf{1}_{\{S_n(s)\le 0\}}\big]}{\mathbb{E}\big[e^{\alpha_n(s)Z}\mathbf{1}_{\{Z\le 0\}}\big]} = 1.$$

Let F_n denote the c.d.f. of S_n. Then, it suffices to prove

$$\lim_{n\to\infty} \left| \frac{\int_{-\infty}^{0} e^{\alpha_n x}\,(dF_n(x) - d\Phi(x))}{\int_{-\infty}^{0} e^{\alpha_n x} d\Phi(x)} \right| = 0.$$

which is equivalent to proving

$$N(n) := \int_{-\infty}^{0} e^{\alpha_n x}\,\big(dF_n(x) - d\Phi(x)\big) = o(n^{-1/2}),$$

since from (4.3.39) we have $D(n) := \int_{-\infty}^{0} e^{\alpha_n x} d\Phi(x) \sim (2\pi n a)^{-1/2}$.

Let $\epsilon > 0$. We split $N(n) = N_1(n) + N_2(n)$, where

$$N_1(n) := \int_0^{\epsilon} e^{-\alpha_n x}(dF_n - d\Phi)(x), \quad N_2(n) := \int_{\epsilon}^{\infty} e^{-\alpha_n x}(dF_n - d\Phi)(x).$$

The tail term satisfies

$$|N_2(n)| \le \int_{\epsilon}^{\infty} e^{-\alpha_n x}\, d|F_n - \Phi|(x) \le 2e^{-\alpha_n \epsilon} = o(1/\alpha_n) = o(n^{-1/2}),$$

where $|F_n - \Phi|$ denotes the total variation measure of the signed measure $F_n - \Phi$.

For $N_1(n)$, we integrate by parts:

$$\begin{aligned}
N_1(n) &= \int_0^{\epsilon} e^{-\alpha_n x}(dF_n(x) - d\Phi(x)) \\
&= \big[e^{-\alpha_n x}(F_n(x) - \Phi(x))\big]_0^{\epsilon} - \int_0^{\epsilon} (F_n(x) - \Phi(x)) d(e^{-\alpha_n x}) \\
&= e^{-\alpha_n \epsilon}(F_n(\epsilon) - \Phi(\epsilon)) - (F_n(0) - \Phi(0)) + \alpha_n \int_0^{\epsilon} e^{-\alpha_n x}(F_n(x) - \Phi(x))\,dx.
\end{aligned}$$

Therefore,

$$N_1(n) = A_n + B_n + C_n,$$

where

$$\begin{aligned} A_n &= e^{-\alpha_n \epsilon}(F_n(\epsilon) - \Phi(\epsilon)), \\ B_n &= -(F_n(0) - \Phi(0)), \\ C_n &= \alpha_n \int_0^\epsilon e^{-\alpha_n x}(F_n(x) - \Phi(x))\, dx. \end{aligned}$$

By Theorem 4.3.8 (applicable by Corollary 4.3.10),

$$B_n = -\frac{a_{3,n}}{6\sqrt{2\pi n}} + o(n^{-1/2}).$$

Similarly, for C_n, we write:

$$C_n = \int_0^\epsilon \alpha_n e^{-\alpha_n x} \left(\frac{a_{3,n}}{6\sqrt{n}}(1-x^2)\phi(x) + r_n(x) \right) dx =: C_n^{(1)} + C_n^{(2)},$$

where $r_n(x) = o(n^{-1/2})$. Thus, $C_n^{(2)} = o(n^{-1/2})$.

For $C_n^{(1)}$, substitute $v = \alpha_n x$ and write:

$$C_n^{(1)} = \frac{a_{3,n}}{6\sqrt{n}} \int_0^{\alpha_n \epsilon} e^{-v} \left(1 - \left(\frac{v}{\alpha_n} \right)^2 \right) \phi\left(\frac{v}{\alpha_n} \right) \frac{dv}{\alpha_n}.$$

Next, we expand the standard normal density around zero:

$$\phi\left(\frac{v}{\alpha_n} \right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{v^2}{2\alpha_n^2} \right).$$

Using a second-order Taylor expansion of the exponential term, we obtain

$$\phi\left(\frac{v}{\alpha_n} \right) = \frac{1}{\sqrt{2\pi}} \left(1 - \frac{v^2}{2\alpha_n^2} + r\left(\frac{v^2}{2\alpha_n^2} \right) \right),$$

where the remainder term is

$$r\left(\frac{v^2}{2\alpha_n^2} \right) = O\left(\frac{v^4}{\alpha_n^4} \right),$$

for $\frac{v^2}{2\alpha_n^2}$ small. That is, there exists a constant $C > 0$ such that

$$\left| r\left(\frac{v^2}{2\alpha_n^2} \right) \right| \le C \frac{v^4}{\alpha_n^4}, \text{ for } \frac{v^2}{2\alpha_n^2} \le \epsilon.$$

Substituting this into the expression for $C_n^{(1)}$, we obtain

$$C_n^{(1)} = \frac{a_{3,n}}{6\sqrt{2\pi n}} \int_0^{\alpha_n \epsilon} e^{-v} \left(1 - \frac{3v^2}{2\alpha_n^2} + r'\left(\frac{v^2}{2\alpha_n^2}\right)\right) dv,$$

where $r'(v^2/2\alpha_n^2) = O(v^4/\alpha_n^4)$ for small $\frac{v^2}{2\alpha_n^2}$, with a different constant.

We now expand this integral term-by-term. Noting the identities

$$\int_0^{\alpha_n \epsilon} e^{-v} dv = \gamma(1, \alpha_n \epsilon), \quad \int_0^{\alpha_n \epsilon} v^2 e^{-v} dv = \gamma(3, \alpha_n \epsilon), \quad \int_0^{\alpha_n \epsilon} v^4 e^{-v} dv = \gamma(5, \alpha_n \epsilon),$$

we obtain:

$$C_n^{(1)} = \frac{a_{3,n}}{6\sqrt{2\pi n}} \left(\gamma(1, \alpha_n \epsilon) - \frac{3}{2\alpha_n^2}\gamma(3, \alpha_n \epsilon) + O\left(\frac{1}{\alpha_n^4}\gamma(5, \alpha_n \epsilon)\right)\right).$$

As $\alpha_n \to \infty$, we have $\gamma(k, \alpha_n \epsilon) \to \Gamma(k)$ for all fixed $\epsilon > 0$. Therefore,

$$\gamma(1, \alpha_n \epsilon) = 1 + o(1), \quad \gamma(3, \alpha_n \epsilon) = 2 + o(1), \quad \gamma(5, \alpha_n \epsilon) = 24 + o(1).$$

Substituting in, we get

$$\begin{aligned} C_n^{(1)} &= \frac{a_{3,n}}{6\sqrt{2\pi n}} \left(1 - \frac{3}{2\alpha_n^2} \cdot 2 + O\left(\frac{1}{\alpha_n^4}\right) + o(1)\right) \\ &= \frac{a_{3,n}}{6\sqrt{2\pi n}} \left(1 - \frac{3}{\alpha_n^2} + O\left(\frac{1}{\alpha_n^4}\right) + o(1)\right). \end{aligned}$$

Recalling that $\alpha_n \sim \sqrt{na}$, it follows that $1/\alpha_n^2 = O(1/n)$, so:

$$C_n^{(1)} = \frac{a_{3,n}}{6\sqrt{2\pi n}} \left(1 + O\left(\frac{1}{n}\right)\right) = \frac{a_{3,n}}{6\sqrt{2\pi n}} + O(n^{-3/2}) = \frac{a_{3,n}}{6\sqrt{2\pi n}} + o(n^{-1/2}).$$

Step 3 (Cancellation). The $O(n^{-1/2})$ terms in B_n and $C_n^{(1)}$ cancel exactly, leaving

$$N_1(n) = A_n + B_n + C_n = o(n^{-1/2}).$$

Since $N_2(n) = o(n^{-1/2})$ as well, we conclude $N(n) = o(n^{-1/2})$, proving (4.3.38). □

4.3.4 Proof of Proposition 4.3.1

We are now ready to prove Proposition 4.3.1.

Proof of Proposition 4.3.1. f Proposition 4.3.1] Recall equation (4.3.17):

$$\mathbb{P}(T_n \leq s) = \psi(\theta_n(s))^n e^{\theta_n(s)s}\mathbb{E}\left[e^{\alpha_n(s)S_n(s)}\mathbf{1}_{\{S_n(s)\leq 0\}}\right].$$

We analyze each term in the product.

By Corollary 4.3.7(i), we have

$$\psi(\theta_n(s)) = c_1\Gamma(a)\left(\frac{s}{a}\right)^a n^{-a} + (1-b)c_2\Gamma(a+b)\left(\frac{s}{a}\right)^{a+b} n^{-(a+b)} + o(n^{-(a+b)}).$$

Equivalently,

$$\psi(\theta_n(s)) = c_1\Gamma(a)\left(\frac{s}{a}\right)^a n^{-a}\left[1 + (1-b)\frac{c_2}{c_1}\frac{\Gamma(a+b)}{\Gamma(a)}\left(\frac{s}{a}\right)^b n^{-b} + o(n^{-b})\right].$$

Therefore,

$$\begin{aligned}\psi(\theta_n(s))^n = &\left[c_1\Gamma(a)\left(\frac{s}{a}\right)^a n^{-a}\right]^n\\ &\times\left[1 + (1-b)\frac{c_2}{c_1}\frac{\Gamma(a+b)}{\Gamma(a)}\left(\frac{s}{a}\right)^b n^{-b} + o(n^{-b})\right]^n.\end{aligned}$$

Since $b > 0$, applying $\log(1+x) = x + O(x^2)$ yields

$$\left(1 + \kappa n^{-b} + o(n^{-b})\right)^n = \exp\left(\kappa n^{1-b} + o(n^{1-b})\right),$$

with

$$\kappa = (1-b)\frac{c_2}{c_1}\frac{\Gamma(a+b)}{\Gamma(a)}\left(\frac{s}{a}\right)^b,$$

we obtain

$$\begin{aligned}\psi(\theta_n(s))^n = &\left[c_1\Gamma(a)\left(\frac{s}{a}\right)^a n^{-a}\right]^n\\ &\times\exp\left((1-b)\frac{c_2}{c_1}\frac{\Gamma(a+b)}{\Gamma(a)}\left(\frac{s}{a}\right)^b n^{1-b} + o(n^{1-b})\right).\end{aligned} \tag{4.3.41}$$

Next, by Lemma 4.3.6(ii),

$$\theta_n(s) = \frac{a}{s}n + \frac{c_2}{c_1}\frac{s^{b-1}b\Gamma(a+b)}{a^b\Gamma(a)}n^{1-b} + o(n^{1-b}).$$

Multiplying by s, we get

$$\theta_n(s)s = an + b\frac{c_2}{c_1}\frac{\Gamma(a+b)}{\Gamma(a)}\left(\frac{s}{a}\right)^b n^{1-b} + o(n^{1-b}).$$

Thus

$$e^{\theta_n(s)s} = \exp\left(an + b\frac{c_2}{c_1}\frac{\Gamma(a+b)}{\Gamma(a)}\left(\frac{s}{a}\right)^b n^{1-b} + o(n^{1-b})\right). \tag{4.3.42}$$

By Lemma 4.3.11,

$$\mathbb{E}\left[e^{\alpha_n(s)S_n(s)}\mathbf{1}_{\{S_n(s)\le 0\}}\right] = \frac{1}{\sqrt{2\pi an}}(1+o(1)). \tag{4.3.43}$$

Combining (4.3.41), (4.3.42), and (4.3.43), we obtain

$$\begin{aligned}
\mathbb{P}(T_n \le s) &= \left[c_1\Gamma(a)\left(\frac{s}{a}\right)^a n^{-a}\right]^n \exp(an) \\
&\quad\times \exp\left([(1-b)+b]\,\frac{c_2}{c_1}\frac{\Gamma(a+b)}{\Gamma(a)}\left(\frac{s}{a}\right)^b n^{1-b} + o(n^{1-b})\right) \\
&\quad\times \frac{1}{\sqrt{2\pi an}}(1+o(1)) \\
&= \left[c_1\Gamma(a)\left(\frac{s}{a}\right)^a n^{-a}\right]^n e^{an}\frac{1}{\sqrt{2\pi an}} \\
&\quad\times \exp\left(\frac{c_2}{c_1}\frac{\Gamma(a+b)}{\Gamma(a)}\left(\frac{s}{a}\right)^b n^{1-b} + o(n^{1-b})\right)(1+o(1)).
\end{aligned}$$

By Stirling's formula,

$$\Gamma(an+1) = \sqrt{2\pi an}\,(an)^{an}e^{-an}(1+o(1)).$$

Equivalently,

$$e^{an}n^{-an}\frac{1}{\sqrt{2\pi an}} = \frac{a^{an}}{\Gamma(an+1)}(1+o(1)).$$

Therefore,

$$\left[c_1\Gamma(a)\left(\frac{s}{a}\right)^a n^{-a}\right]^n e^{an}\frac{1}{\sqrt{2\pi an}} = \left[c_1\Gamma(a)\frac{s^a}{a^a}\right]^n n^{-an}e^{an}\frac{1}{\sqrt{2\pi an}}$$
$$= \big(c_1 s^a\Gamma(a)\big)^n\frac{1}{\Gamma(an+1)}(1+o(1)).$$

Hence

$$\mathbb{P}(T_n\le s) = \frac{\big(c_1 s^a\Gamma(a)\big)^n}{\Gamma(an+1)}\exp\left(\frac{c_2}{c_1}\frac{\Gamma(a+b)}{\Gamma(a)}\left(\frac{s}{a}\right)^b n^{1-b} + o(n^{1-b})\right)(1+o(1)).$$

□

4.4 Proof of Theorem 4.1.9

The first main step is to show that the process of interest is synergistic at extremes. We show that by applying Lemma 4.2.3. To this end, we need the following fact.

Lemma 4.4.1. *Let $(N_t)_{t\ge 0}$ be a counting process generated by the i.i.d. sequence of waiting times $(W_n)_{n\ge 1}$. Namely, let $T_n := \sum_{i=1}^n W_i$ denote the renewal epochs and*

$$N_t := \max\{n\in\mathbb{N}_+ \mid T_n\le t\}.$$

Assume that W_1 satisfies assumption (4.1.1). Let $0<s<t$. Then, for the sequence

$$a_n(s,t) := \frac{\mathbb{P}(N_s=n)}{\mathbb{P}(N_t=n+2)},$$

we have

$$\sum_{n=1}^{\infty} a_n(s,t) < \infty.$$

Proof. Since

$$\mathbb{P}(N_t=n+2) = \mathbb{P}(T_{n+2}\le t< T_{n+3})$$

and

$$\mathbb{P}(N_s=n)\le\mathbb{P}(T_n\le s)\,,$$

we have

$$a_n(s,t)\le\frac{\mathbb{P}(T_n\le s)}{\mathbb{P}(T_{n+2}\le t<T_{n+3})}.$$

To bound the denominator from below, fix $\delta > 0$ small enough that $\mathbb{P}(W_1 > \delta) > 0$. Then, by restricting T_{n+2} to the interval $(t-\delta, t)$ and using the independence of W_{n+3} from T_{n+2}, we obtain

$$\begin{aligned}\mathbb{P}(T_{n+2} \le t < T_{n+3}) &\ge \mathbb{P}(T_{n+2} \in (t-\delta, t),\, W_{n+3} > \delta) \\ &= \mathbb{P}(T_{n+2} \in (t-\delta, t))\, \mathbb{P}(W_1 > \delta)\,.\end{aligned}$$

Therefore,

$$a_n(s,t) \le \frac{1}{\mathbb{P}(W_1 > \delta)} \frac{\mathbb{P}(T_n \le s)}{\mathbb{P}(T_{n+2} \le t) - \mathbb{P}(T_{n+2} \le t-\delta)}. \tag{4.4.1}$$

It suffices to show that the ratio on the right-hand side is summable over n.

Step 1: The denominator correction is negligible. By Proposition 4.3.1,

$$\mathbb{P}(T_n \le u) = \frac{\big(c_1 u^a \Gamma(a)\big)^n}{\Gamma(an+1)} \exp\left(\kappa u^b n^{1-b} + o(n^{1-b})\right), \qquad n \to \infty, \tag{4.4.2}$$

where

$$\kappa := \frac{c_2 \Gamma(a+b)}{c_1 a^b \Gamma(a)}.$$

Applying this expansion with $u = t-\delta$ and $u = t$, we get

$$\frac{\mathbb{P}(T_n \le t-\delta)}{\mathbb{P}(T_n \le t)} = \left(\frac{t-\delta}{t}\right)^{an} \exp\left(\kappa\big((t-\delta)^b - t^b\big)n^{1-b} + o(n^{1-b})\right).$$

Since $(t-\delta)/t < 1$ and $n^{1-b} = o(n)$, the geometric factor dominates the sublinear exponential correction. Hence

$$\mathbb{P}(T_n \le t-\delta) = o(\mathbb{P}(T_n \le t))\,.$$

In particular,

$$\mathbb{P}(T_{n+2} \le t) - \mathbb{P}(T_{n+2} \le t-\delta) \sim \mathbb{P}(T_{n+2} \le t)\,. \tag{4.4.3}$$

Step 2: The main ratio is summable. By (4.4.3), it remains to show that

$$\sum_{n=1}^{\infty} \frac{\mathbb{P}(T_n \le s)}{\mathbb{P}(T_{n+2} \le t)} < \infty.$$

Applying (4.4.2) to both terms gives

$$\frac{\mathbb{P}(T_n \leq s)}{\mathbb{P}(T_{n+2} \leq t)} = \frac{\big(c_1 s^a \Gamma(a)\big)^n}{\big(c_1 t^a \Gamma(a)\big)^{n+2}} \frac{\Gamma(a(n+2)+1)}{\Gamma(an+1)} \\ \times \exp\left(\kappa\left[s^b n^{1-b} - t^b (n+2)^{1-b}\right] + o(n^{1-b})\right).$$

Using Stirling's approximation,

$$\frac{\Gamma(a(n+2)+1)}{\Gamma(an+1)} \sim (an)^{2a}.$$

Therefore, for some constant $C_{s,t} > 0$,

$$\frac{\mathbb{P}(T_n \leq s)}{\mathbb{P}(T_{n+2} \leq t)} = C_{s,t} \left(\frac{s}{t}\right)^{an} (an)^{2a} \exp\left(\kappa\left[s^b n^{1-b} - t^b (n+2)^{1-b}\right] + o(n^{1-b})\right) (1 + o(1)).$$

Since $b > 0$, the exponential correction is $\exp(o(n))$, and the polynomial factor $(an)^{2a}$ is also $\exp(o(n))$. Hence

$$\frac{\mathbb{P}(T_n \leq s)}{\mathbb{P}(T_{n+2} \leq t)} = \left(\frac{s}{t}\right)^{an} \exp(o(n)).$$

Since $s < t$, we have $(s/t)^a < 1$. Therefore,

$$\limsup_{n\to\infty} \left(\frac{\mathbb{P}(T_n \leq s)}{\mathbb{P}(T_{n+2} \leq t)}\right)^{1/n} = \left(\frac{s}{t}\right)^a < 1.$$

The root test implies

$$\sum_{n=1}^{\infty} \frac{\mathbb{P}(T_n \leq s)}{\mathbb{P}(T_{n+2} \leq t)} < \infty. \tag{4.4.4}$$

By (4.4.1), (4.4.3) and (4.4.4), it follows that

$$\sum_{n=1}^{\infty} a_n(s,t) < \infty.$$

□

Remark 4.4.2. *Interestingly, although we showed that* $\mathbb{P}(T_n \leq t - \delta) = o(\mathbb{P}(T_n \leq t))$, *the relationship*

$$\mathbb{P}(T_{n+1} \leq s) = o(\mathbb{P}(T_n \leq s))$$

does not *hold in general for every* $b > 0$. *Indeed, by the asymptotic expansion in*

Proposition 4.3.1, we have

$$\frac{\mathbb{P}(T_{n+1} \le s)}{\mathbb{P}(T_n \le s)} = n^{-a} \exp(o(n^{1-b})),$$

which is not $o(1)$ *when* $b < 1$.

We now show that Lemma 4.2.3 and Lemma 4.4.1 imply that the type of compound renewal processes considered in this chapter —both in their pure form and with an added linear drift —are synergistic at extremes. We emphasize that, in the proof of Theorem 4.1.9, we will only use the synergistic at extremes property for the *undrifted* compound renewal process.

Recall that the positive linear drift breaks the regeneration-at-first-passage-times property(see Remark 2.6.5). On the other hand, the negative drift prevents us from applying the crucial Proposition 4.2.1 (see Remark 4.2.2). The property of being synergistic at extremes for the drifted case is established here in order to justify a previous claim of ours that the compound Poisson process considered in Braverman's Theorem 3.2.5 is synergistic at extremes.

Corollary 4.4.3. *Let* $(X_t)_{t\ge 0}$ *be a compound renewal process with linear drift, i.e.,* $X_t = Y_t + bt$, *with* $b \in \mathbb{R}$ *and* $(Y_t)_{t\ge 0}$ *being a compound renewal process with Braverman-light jumps and waiting times that satisfy Assumption 4.1.1. In addition, assume that the sequence of waiting times and jumps are independent. Then,* $(X_t)_{t\ge 0}$ *is synergistic at extremes.*

Proof. First, by Lemma 4.2.3 and Lemma 4.4.1, we obtain that, for any $c > 0$ and $0 < s < t$

$$\lim_{x\to\infty} \frac{\mathbb{P}(Y_s > x)}{\mathbb{P}(Y_t > x + c)} = 0 \tag{4.4.5}$$

We treat the cases $b \ge 0$ and $b < 0$ separately.

Case 1: $b \ge 0$. Then

$$\begin{aligned}\frac{\mathbb{P}(X_s > x)}{\mathbb{P}(X_t > x)} &= \frac{\mathbb{P}(Y_s > x - bs)}{\mathbb{P}(Y_t > x - bt)} \\ &\le \frac{\mathbb{P}(Y_s > x - bs)}{\mathbb{P}(Y_t > x)}.\end{aligned}$$

Setting $u = x - bs$, and noting that $x - bt \leq x$ since $b \geq 0$, we get

$$\frac{\mathbb{P}(X_s > x)}{\mathbb{P}(X_t > x)} \leq \frac{\mathbb{P}(Y_s > u)}{\mathbb{P}(Y_t > u)} \xrightarrow{x\to\infty} 0$$

by (4.4.5).

Case 2: $b < 0$. Then we write

$$\frac{\mathbb{P}(X_s > x)}{\mathbb{P}(X_t > x)} = \frac{\mathbb{P}(Y_s > x - bs)}{\mathbb{P}(Y_t > x - bt)}.$$

Letting $u = x - bs$ and noting that $c := -b(t - s) > 0$, we rewrite:

$$\frac{\mathbb{P}(X_s > x)}{\mathbb{P}(X_t > x)} = \frac{\mathbb{P}(Y_s > u)}{\mathbb{P}(Y_t > u + c)} \xrightarrow{x\to\infty} 0$$

again by (4.4.5). □

Note that, we will not actually need for the proof that the compound renewal process with linear drift is synergistic but only that the drift-less process has that property.

Now we are ready to prove Theorem 4.1.9.

Proof of Theorem 4.1.9. We first show the tail equivalence with the supremum for the driftless compound renewal process, $(Y_t)_{t\geq 0}$. Let $\epsilon > 0$, and recall that $T_x := \inf\{t \geq 0 : Y_t \geq x\}$. In the same fashion as we have done multiple times in Chapter 3, we compute:

$$\begin{aligned}
\mathbb{P}\big(Y_t > x \mid \overline{Y}_t > x\big) &= \mathbb{P}(Y_t > x \mid T_x \leq t) \\
&\geq \mathbb{P}(Y_t > x, T_x > t - \epsilon \mid T_x \leq t) \\
&\geq \mathbb{P}(T_x > t - \epsilon \mid T_x \leq t) \cdot \mathbb{P}(Y_t > x \mid t - \epsilon < T_x \leq t) \\
&\geq \mathbb{P}(T_x > t - \epsilon \mid T_x \leq t) \cdot \mathbb{P}(Y_t - Y_{T_x} \geq 0 \mid t - \epsilon < T_x \leq t) \\
&= \mathbb{P}(T_x > t - \epsilon \mid T_x \leq t) \int_{t-\epsilon}^{t} \mathbb{P}(Y_{t-s} \geq 0)\ \mathbb{P}(T_x \in ds \mid t - \epsilon < T_x \leq t) \\
&\geq \mathbb{P}(T_x > t - \epsilon \mid T_x \leq t) \cdot \inf_{s\in(0,\epsilon)} \mathbb{P}(Y_s \geq 0)\,. \qquad (4.4.6)
\end{aligned}$$

Note that in the last equality we used the regeneration-at-first-passage-times property.

Recall that we denote the waiting times associated with the compound renewal process by $(W_i)_{i\geq 1}$. For any $s > 0$, we have

$$\mathbb{P}(Y_s \geq 0) \geq \mathbb{P}(W_1 > s)\,,$$

and therefore,

$$\inf_{s\in(0,\epsilon)} \mathbb{P}(Y_s \geq 0) \geq \mathbb{P}(W_1 > \epsilon)\,. \tag{4.4.7}$$

Combining (4.4.6) and (4.4.7), we obtain

$$\mathbb{P}\big(Y_t > x \mid \overline{Y}_t > x\big) \geq \mathbb{P}(T_x > t-\epsilon \mid T_x \leq t)\cdot \mathbb{P}(W_1 > \epsilon)\,. \tag{4.4.8}$$

By Corollary 4.4.3 $(Y_t)_{t\geq 0}$ is synergistic at extremes. In addition, it is regenerative at first passage times and by (4.4.7), it holds $\inf_{s\in(0,\epsilon)} \mathbb{P}(Y_s \geq 0) > 0$. Therefore, it satisfies the condition of proposition 4.2.1, which yields

$$\liminf_{x\to\infty} \mathbb{P}(T_x > t-\epsilon \mid T_x \leq t) = 1. \tag{4.4.9}$$

Using (4.4.8) and (4.4.9), we conclude

$$\liminf_{x\to\infty} \mathbb{P}\big(Y_t > x \mid \overline{Y}_t > x\big) \geq \mathbb{P}(W_1 > \epsilon)\,. \tag{4.4.10}$$

Finally, sending $\epsilon \to 0^+$, we obtain

$$\liminf_{x\to\infty} \mathbb{P}\big(Y_t > x \mid \overline{Y}_t > x\big) \geq 1,$$

which implies

$$\mathbb{P}\big(\overline{Y}_t > x\big) \sim \mathbb{P}(Y_t > x)\,. \tag{4.4.11}$$

Now, we prove the tail equivalence for the process $X_t = Y_t + bt$, with $b \geq 0$. It suffices to show that

$$\limsup_{x\to\infty} \frac{\mathbb{P}\big(\overline{X}_t > x\big)}{\mathbb{P}(X_t > x)} \leq 1.$$

The above is implied by

$$\mathbb{P}\big(\overline{X}_t > x\big) \leq \mathbb{P}\big(\overline{Y}_t > x - bt\big) \sim \mathbb{P}(Y_t + bt > x) = \mathbb{P}(X_t > x)\,,$$

where we used that, $\overline{X}_t = \overline{Y_t + bt} \leq \overline{Y}_t + bt$ in the inequality, and the asymptotic equivalence (4.4.11) on the next step. $\square$

Appendix A

Auxiliary Results

A.1 Edgeworth Expansion for Triangular Arrays

This chapter contains the proof of the Edgeworth expansion used in Section 4.3.3. We follow Petrov's approach for the i.i.d. case and adapt it to row-wise i.i.d. triangular arrays, highlighting the few points where uniformity in n is required.

A key tool in establishing these results is the comparison of two distribution functions F and G through their characteristic functions f and g. This is achieved via *smoothing inequalities*, which, in this context, bound the Kolmogorov distance

$$\sup_{x\in\mathbb{R}} |F(x) - G(x)|$$

in terms of an integral involving $|f(t) - g(t)|/|t|$ over a finite frequency range, plus a tail term that depends on the smoothness of G. Such bounds allow one to transfer precise local approximations of characteristic functions into global distributional approximations.

Here, we follow the proof of Theorem 5.22 in [6] for the i.i.d. case. The only substantial difference from that proof is that, in our setting, the array is row-wise i.i.d., so the distribution changes with n and we work with a *sequence* of characteristic functions rather than a single fixed one. Consequently, we require *uniform* local approximations of the characteristic functions across n — see equations (A.1.1) and (A.1.2) below. Apart from this modification, the structure of the argument remains essentially the same.

As in Theorem 5.22 in [6] for the i.i.d. case, we will use the following smoothing inequality, due to Esseen, which can be found in [6, Theorem 5.2]:

Theorem A.1.1 (Esseen's Smoothing Inequality)**.** *Let $F(x)$ be a bounded non-decreasing function, and $G(x)$ a differentiable function of bounded variation on the real line. Let $F(\infty) = G(\infty)$. Put*

$$f(t) = \int_{-\infty}^{\infty} e^{itx}\, dF(x), \qquad g(t) = \int_{-\infty}^{\infty} e^{itx}\, dG(x).$$

Suppose that $\sup_x |G'(x)| \le K$. Then, for every $T > 0$ and every $b > \frac{1}{2\pi}$,

$$\sup_{x\in\mathbb{R}} |F(x) - G(x)| \le c(b) \int_{-T}^{T} \left| \frac{f(t) - g(t)}{t} \right| dt + \frac{c(b)K}{T},$$

where $c(b) > 0$ is a constant depending only on b.

Theorem A.1.2 (Restatement of Theorem 4.3.8)**.** *Let $(X_{i,n})_{1\le i\le n}$ be a row-wise i.i.d. triangular array of standardised random variables, that is,*

$$\mathbb{E}[X_{1,n}] = 0, \qquad \mathbb{E}\big[X_{1,n}^2\big] = 1.$$

Denote

$$u_n(t) := \mathbb{E}\big[e^{itX_{1,n}}\big], \qquad a_{k,n} := \mathbb{E}\big[X_{1,n}^k\big], \qquad m_{k,n} := \mathbb{E}\big[|X_{1,n}|^k\big].$$

Assume the following conditions.

(a) For every $0 < \delta < K < \infty$, there exist $r_{\delta,K} \in (0,1)$ and $n_{\delta,K} \in \mathbb{N}$ such that, for all $n \ge n_{\delta,K}$,

$$\sup_{\delta\le|t|\le K} |u_n(t)| \le r_{\delta,K}.$$

(b) There exists a constant $M_4 > 0$ such that, for all n,

$$m_{4,n} = \mathbb{E}\big[|X_{1,n}|^4\big] \le M_4.$$

(c) There exist $\delta_0 > 0$, $c_{\delta_0} > 0$, and $n_0 \in \mathbb{N}$ such that, for all $n \ge n_0$,

$$\inf_{|t|<\delta_0} |u_n(t)| \ge c_{\delta_0}.$$

Let

$$S_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_{i,n},$$

and let F_n denote the c.d.f. of S_n. Then, uniformly in $x \in \mathbb{R}$,

$$F_n(x) - \Phi(x) = \frac{a_{3,n}}{6\sqrt{n}}(1 - x^2)\phi(x) + o(n^{-1/2}),$$

where Φ and ϕ denote the standard normal c.d.f. and p.d.f., respectively.

Proof of Theorem 4.3.8 . We apply Theorem A.1.1 with:

$$b = \frac{1}{\pi}, \quad F = F_n, \quad G(x) = \Phi(x) + \frac{1}{6\sqrt{2\pi n}} a_{3,n}(1 - x^2)e^{-x^2/2}, \quad T = K\sqrt{n},$$

where $K > 0$ is a constant.

It is straightforward to verify that the Fourier transform of G is given by

$$g(t) := \int_{-\infty}^{\infty} e^{itx} dG(x) = e^{-t^2/2} \left(1 + \frac{1}{6\sqrt{n}} a_{3,n}(it)^3\right).$$

Moreover, since the sequence of third moments $(a_{3,n})$ is uniformly bounded (say $|a_{3,n}| \leq M_3$), the derivative $G'(x)$ is uniformly bounded as well: there exists a constant $C > 0$ such that $\sup_x |G'(x)| \leq C$. Thus, by Theorem A.1.1, we obtain:

$$\sup_x |F_n(x) - G(x)| \leq \frac{1}{\pi} \int_{|t|<K\sqrt{n}} |f_n(t) - g(t)| \frac{dt}{|t|} + \frac{AC}{K} n^{-1/2},$$

where f_n is the characteristic function of F_n.

Let $\epsilon > 0$. Choose K large enough so that $\frac{AC}{K} < \epsilon$, and define

$$I := \int_{|t|<K\sqrt{n}} |f_n(t) - g(t)| \frac{dt}{|t|}.$$

Then the theorem will follow if we show that $I < \epsilon n^{-1/2}$ for all sufficiently large n. We decompose

$$I = I_1 + I_2,$$

where

$$I_1 := \int_{|t|<\delta\sqrt{n}} |f_n(t) - g(t)| \frac{dt}{|t|}, \qquad I_2 := \int_{\delta\sqrt{n}<|t|<K\sqrt{n}} |f_n(t) - g(t)| \frac{dt}{|t|},$$

and $\delta > 0$ is fixed and small.

Let $B := \{t : \delta\sqrt{n} < |t| < K\sqrt{n}\}$. We further write

$$I_2^{(1)} := \int_B |f_n(t)| \frac{dt}{|t|}, \qquad I_2^{(2)} := \int_B |g(t)| \frac{dt}{|t|},$$

so that $I_2 \leq I_2^{(1)} + I_2^{(2)}$.

By the explicit form of $g(t)$ and the uniform bound on $a_{3,n}$, we have $|g(t)| \leq$

$Ce^{-t^2/2}(1 + C'|t|^3/\sqrt{n})$, which implies that

$$I_2^{(2)} = o(n^{-p}) \quad \text{for all } p > 0.$$

Now we estimate $I_2^{(1)}$. Since $(X_{i,n})_{1\le i\le n}$ is a row-wise i.i.d. triangular array, and $X_{i,n} \stackrel{d}{=} X_{1,n}$ for all $i \le n$, it follows that

$$f_n(t) = \left(u_n\left(\frac{t}{\sqrt{n}}\right)\right)^n,$$

where u_n is the characteristic function of $X_{1,n}$. Setting $v = t/\sqrt{n}$, we write

$$I_2^{(1)} = \int_{\delta<|v|<K} |u_n(v)|^n \frac{dv}{|v|}.$$

By condition (a), applied with the present $0 < \delta < K < \infty$, there exists $r_{\delta,K} \in (0,1)$ such that, for all sufficiently large n,

$$\sup_{\delta\le|v|\le K} |u_n(v)| \le r_{\delta,K}.$$

Hence

$$I_2^{(1)} \le r_{\delta,K}^n 2\log(K/\delta) = o(n^{-1/2}).$$

Thus, $I_2 = o(n^{-1/2})$.

It remains to show that $I_1 < \epsilon n^{-1/2}$ for all sufficiently large n.

To that end, the following two estimates are (the most) crucial components of the proof, and we establish them first:

$$\sup_n \left|\log u_n(t) + \frac{t^2}{2}\right| = o(t^2) \quad \text{as } t \to 0 \tag{A.1.1}$$

$$\sup_n \left|\log u_n(t) + \frac{t^2}{2} - \frac{a_{3,n}}{6}(it)^3\right| = o(t^3) \quad \text{as } t \to 0 \tag{A.1.2}$$

We prove (A.1.2); the argument for (A.1.1) is similar.

By the bound in condition (c) and the existence of 4th moments in condition (b), we can argue that $\log u_n(t)$ is four times differentiable on $(0, \delta)$ with continuous third derivative. Therefore, we apply the fourth-order Taylor expansion of $\log u_n(t)$ around zero in Lagrange form. Denoting by $k_{3,n}$, $k_{4,n}$ the third and fourth cumulants of $X_{n,n}$, we have:

$$\log u_n(t) + \frac{t^2}{2} - \frac{k_{3,n}}{6}(it)^3 = \frac{k_{4,n}}{4!}(it)^4 + r_n(t),$$

where the remainder is

$$r_n(t) = \frac{(\log u_n)^{(4)}(\xi_n)}{4!}t^4 \quad \text{for some } \xi_n \in (0,t).$$

Since $X_{1,n}$ is standardized, we have $k_{3,n} = a_{3,n}$ and $k_{4,n} = a_{4,n} - a_{2,n}^2$. By that and condition (b), there exists a constant M_4' such that $\sup_n |k_{4,n}| \leq M_4'$. Hence,

$$\sup_n \left| \log u_n(t) + \frac{t^2}{2} - \frac{a_{3,n}}{6}(it)^3 \right| \leq \frac{M_4'}{4!}|t|^4 + \sup_n |r_n(t)|.$$

To bound $r_n(t)$, we compute the fourth derivative of the logarithm of the characteristic function:

$$(\log u_n)^{(4)}(t) = \frac{u_n^{(4)}(t)}{u_n(t)} - 4\frac{u_n^{(3)}(t)u_n'(t)}{u_n(t)^2} - 3\left(\frac{u_n''(t)}{u_n(t)}\right)^2 + 12\frac{u_n''(t)(u_n'(t))^2}{u_n(t)^3} - 6\left(\frac{u_n'(t)}{u_n(t)}\right)^4.$$

Each derivative of $u_n(t) = \mathbb{E}[e^{itX_{1,n}}]$ satisfies

$$u_n^{(j)}(t) = \mathbb{E}[(iX_{1,n})^j e^{itX_{1,n}}],$$

so that

$$|u_n^{(j)}(t)| \leq \mathbb{E}[|X_{1,n}|^j].$$

By the uniform bound on the moments of $X_{n,n}$ up to order 4 (assumption (b)) and the uniform lower bound $|u_n(t)| \geq c > 0$ on $(0,\delta)$ (assumption (c)), it follows that there exists a constant $C' > 0$ such that

$$\sup_{n\geq 1} \sup_{|t|<\delta} \left|(\log u_n)^{(4)}(t)\right| \leq C'.$$

Therefore, for all $|t| < \delta$,

$$|r_n(t)| = \left|\frac{(\log u_n)^{(4)}(\xi_n)}{4!}\right| t^4 \leq \frac{C'}{4!}t^4.$$

Putting everything together, we obtain:

$$\sup_n \left|\log u_n(t) + \frac{t^2}{2} - \frac{a_{3,n}}{6}(it)^3\right| \leq \left(\frac{M_4'}{4!} + \frac{C'}{4!}\right)|t|^4 = O(|t|^4),$$

which proves that

$$\sup_n \left|\log u_n(t) + \frac{t^2}{2} - \frac{a_{3,n}}{6}(it)^3\right| = o(|t|^3) \quad \text{as } t \to 0,$$

as desired.

Now, we define

$$L_n(t) := \log u_n(t) + \frac{t^2}{2}, \quad \text{for sufficiently small } t.$$

It is clear that

$$u_n(t) = \exp\{L_n(t) - t^2/2\},$$

and hence

$$I_1 = \int_{|t|\leq\delta\sqrt{n}} e^{-t^2/2}\left|\exp\left\{nL_n\left(\frac{t}{\sqrt{n}}\right)\right\} - 1 - \frac{a_{3,n}(it)^3}{6\sqrt{n}}\right|\frac{dt}{|t|}.$$

Making use of the power series expansion of the exponential function, we apply the identity

$$e^x - 1 - y = (e^{x-y} - 1)e^y + (e^y - 1 - y),$$

which leads to the inequality

$$|e^x - 1 - y| \leq \left(|x-y| + \tfrac{1}{2}|y|^2\right)e^{3z},$$

valid for all real or complex x, y, where $z = \max(|x|, |y|)$. Applying this to our expression with

$$x = nL_n\left(\frac{t}{\sqrt{n}}\right), \quad y = \frac{a_{3,n}(it)^3}{6\sqrt{n}},$$

we obtain

$$I_1 \leq \int_{|t|\leq\delta\sqrt{n}} e^{-t^2/8}\left\{\left|nL_n\left(\frac{t}{\sqrt{n}}\right) - \frac{a_{3,n}(it)^3}{6\sqrt{n}}\right| + \frac{a_{3,n}^2 t^6}{72n}\right\}\frac{dt}{|t|}, \tag{A.1.3}$$

provided we can verify that, for δ small enough

$$z = \max\left(\left|nL_n\left(\frac{t}{\sqrt{n}}\right)\right|, \left|\frac{a_{3,n}(it)^3}{6\sqrt{n}}\right|\right) \leq \frac{t^2}{8} \quad \text{for all } |t| < \delta\sqrt{n}. \tag{A.1.4}$$

We now justify inequality (A.1.4).

We first show that for δ small enough

$$\left|\frac{a_{3,n}(it)^3}{6\sqrt{n}}\right| \leq \frac{t^2}{8} \quad \text{for all } |t| < \delta\sqrt{n}. \tag{A.1.5}$$

Setting $t = v\sqrt{n}$, the above is equivalent to showing that, for δ small enough,

$$\frac{|a_{3,n}|}{6}|v| < \frac{1}{8} \quad \text{for all } |v| < \delta, \tag{A.1.6}$$

which holds since the sequence $(a_{3,n})$ is uniformly bounded by assumption (b).

Next, we need to show that, for δ small enough,

$$|nL_n(t/\sqrt{n})| \leq \frac{t^2}{8} \quad \text{for all } |t| < \delta\sqrt{n}. \tag{A.1.7}$$

Again, setting $t = v\sqrt{n}$, the above is equivalent to showing that, for δ small enough,

$$|L_n(v)| \leq \frac{v^2}{8} \quad \text{for all } |v| < \delta, \tag{A.1.8}$$

which follows by (A.1.1).

We now continue from (A.1.3). By (A.1.2), we have

$$\sup_n \left|L_n(t) - \frac{a_{3,n}}{6}(it)^3\right| = o(|t|^3) \quad \text{as } t \to 0,$$

which implies that for every $\epsilon_0 > 0$, there exists $\delta > 0$ such that

$$\left|L_n(t) - \frac{a_{3,n}}{6}(it)^3\right| \leq \epsilon_0|t|^3 \quad \text{for all } |t| \leq \delta \text{ and all } n.$$

Substituting this bound into (A.1.3), we get

$$I_1 \leq \int_{|t|\leq\delta\sqrt{n}} e^{-t^2/8}\left\{\epsilon_0|t|^3 n^{-1/2} + \frac{a_{3,n}^2 t^6}{72n}\right\}\frac{dt}{|t|}.$$

Evaluating the integral, and using the fact that $a_{3,n}$ is uniformly bounded, we obtain

$$I_1 \leq n^{-1/2}\left(\epsilon_0\int_{-\infty}^{\infty}|t|^2 e^{-t^2/8}dt + \frac{\sup_n a_{3,n}^2}{72\sqrt{n}}\int_{-\infty}^{\infty} t^5 e^{-t^2/8}dt\right).$$

Both integrals are finite and can be computed explicitly:

$$\int_{-\infty}^{\infty} t^2 e^{-t^2/8} dt = 2\int_0^{\infty} t^2 e^{-t^2/8} dt = 2 \cdot \left(\frac{\sqrt{\pi}}{2} \cdot 8^{3/2}\right) = 8\sqrt{2\pi},$$

$$\int_{-\infty}^{\infty} t^6 e^{-t^2/8} \frac{dt}{|t|} = 1024,$$

and hence

$$I_1 \le n^{-1/2}\left(\epsilon_0 \cdot 8\sqrt{2\pi} + \frac{M_3^2}{72\sqrt{n}} \cdot 1024\right).$$

Thus, for all sufficiently large n, we can make $I_1 < \epsilon n^{-1/2}$ by choosing ϵ_0 sufficiently small. □

A.2 Analytical and Asymptotic Identities

Fact A.2.1. *Let Y be a non-negative random variable and let $\eta > 0$. Then*

$$\mathbb{E}\big[e^{\eta Y}\big] = 1 + \eta \int_0^{\infty} e^{\eta x}\mathbb{P}(Y > x)\ dx,$$

where the right-hand side may take the value $+\infty$.

Proof. We use the well-known tail representation of the expectation: for any non-negative random variable Z,

$$\mathbb{E}[Z] = \int_0^{\infty} \mathbb{P}(Z > x)\ dx.$$

Applying this with $Z = e^{\eta Y}$ gives

$$\mathbb{E}\big[e^{\eta Y}\big] = \int_0^{\infty} \mathbb{P}\big(e^{\eta Y} > x\big)\ dx.$$

Since $Y \ge 0$ and $\eta > 0$, we have $e^{\eta Y} \ge 1$. Hence

$$\mathbb{P}\big(e^{\eta Y} > x\big) = 1, \qquad 0 \le x < 1,$$

and therefore

$$\mathbb{E}\big[e^{\eta Y}\big] = 1 + \int_1^{\infty} \mathbb{P}\big(e^{\eta Y} > x\big)\ dx.$$

For $x \geq 1$,

$$\mathbb{P}\big(e^{\eta Y} > x\big) = \mathbb{P}\Big(Y > \tfrac{1}{\eta}\log x\Big).$$

Thus

$$\mathbb{E}\big[e^{\eta Y}\big] = 1 + \int_1^\infty \mathbb{P}\Big(Y > \tfrac{1}{\eta}\log x\Big)\,dx.$$

Making the change of variables $x = e^{\eta u}$ (so $dx = \eta e^{\eta u}\,du$) yields

$$\mathbb{E}\big[e^{\eta Y}\big] = 1 + \eta \int_0^\infty e^{\eta u}\mathbb{P}(Y > u)\,du,$$

which proves the claim. □

Lemma A.2.2 (Constructive two-term inversion near zero)**.** *Let* $h : (0, \Lambda) \to (0, \varepsilon_0)$ *be strictly decreasing and bijective, so that* h *is invertible on* (Λ, ∞) *with inverse* $h^{-1} : (0, h(\Lambda)) \to (\Lambda, \infty)$*. Assume that, as* $\lambda \to \Lambda$*,*

$$h(\lambda) = A\,\lambda^{-1} + B\,\lambda^{-(1+b)} + o\big(\lambda^{-(1+b)}\big), \qquad A > 0, \quad b > 0, \quad B \in \mathbb{R}. \tag{A.2.1}$$

Then, as $x \to 0^+$*,*

$$h^{-1}(x) = \frac{A}{x} + \frac{B}{A^b}\,x^{\,b-1} + o\big(x^{\,b-1}\big). \tag{A.2.2}$$

Fact A.2.3. *For* $b > 0$

$$(1 + O(n^{-2b}))^n = \exp\big(O(n^{1-2b})\big).$$

Proof. Let $\varepsilon_n = O(n^{-2b})$. By Taylor expansion

$$\log\big(1 + \varepsilon_n\big) = \varepsilon_n + O(\varepsilon_n^2) = O(n^{-2b}) + O(n^{-4b}).$$

Hence

$$n\log\big(1 + \varepsilon_n\big) = O(n^{1-2b}) + O(n^{1-4b}) = O(n^{1-2b}),$$

and exponentiating

$$(1 + O(n^{-2b}))^n = \exp\big(O(n^{1-2b})\big).$$

□

A.3 Auxiliary Technical Results

Proof of Fact 3.3.3. We apply Lemma 3.4.6 with $x_0 = 0$. This yields

$$\mathbb{P}\big(\overline{X}_t > x + c\big) \leq \frac{\mathbb{P}(X_t > x + c)}{\inf_{0\leq s\leq t} \mathbb{P}(X_s > 0)}.$$

Since

$$\inf_{0\leq s\leq t} \mathbb{P}(X_s > 0) > 0$$

it follows that

$$\frac{\mathbb{P}\big(\overline{X}_t > x + c\big)}{\mathbb{P}\big(\overline{X}_t > x\big)} \leq C\,\frac{\mathbb{P}(X_t > x + c)}{\mathbb{P}(X_t > x)},$$

where

$$C := \frac{1}{\inf_{0\leq s\leq t} \mathbb{P}(X_s > 0)} > 0.$$

Now taking limits and using (3.3.3) proves the claim. □

Proof of Lemma 3.4.23. For the first inequality, note that

$$\overline{\sigma B_t - C_t} \;\geq\; \sigma\overline{B}_t - \overline{C}_t \;=\; \sigma\overline{B}_t - C_t,$$

since C_t has non-decreasing paths. Thus,

$$\mathbb{P}\big(\overline{\sigma B_t - C_t} > x\big) \;\geq\; \mathbb{P}\big(\sigma\overline{B}_t > x + C_t\big) \;=\; 2\mathbb{P}(\sigma B_t > x + C_t)\,,$$

where the last step uses the reflection principle (Theorem 3.2.1), valid for $x + C_t \geq 0$, which holds since $C_t \geq 0$ almost surely.

For the second inequality, note that

$$\overline{\sigma B_t + C_t} \;\leq\; \sigma\overline{B}_t + \overline{C}_t \;=\; \sigma\overline{B}_t + C_t,$$

so

$$\mathbb{P}\big(\overline{\sigma B_t + C_t} > x\big) \;\leq\; \mathbb{P}\big(\sigma\overline{B}_t > x - C_t\big)\,.$$

Applying the reflection principle, $\mathbb{P}\big(\overline{B}_t > y\big) \leq 2\mathbb{P}(B_t > y)$ for all $y \in \mathbb{R}$, with equality when $y > 0$, gives the claim. □

Appendix B

Logarithmic Asymptotics for Powered Random Walk Excursions

B.1 Introduction

In this chapter, we develop logarithmic asymptotics for the area under a polynomial functional of a random walk with semi-exponential increments, observed during its first busy period. Such random walks arise naturally in the modeling of waiting times in single-server queues operating under a FIFO discipline, and are thus of significant interest in Operational Research. The case where the increments are light-tailed has been treated in [5]. In the present work, we extend this methodology to the case where the increments have a semi-exponential right tail and an arbitrary left tail. The problem of determining exact asymptotics for the area under a random walk, without any polynomial transformation, has already received significant attention. Local asymptotics for increments with light-tails were derived in [40], while [10] treated the case of regularly varying increments (with index greater than 1). A particularly relevant contribution is found in [18], where the authors obtained exact asymptotics for the area under the random walk during its first excursion for a broad class of increments, including lognormal and semi-exponential distributions with parameter $\alpha \leq 1/2$. While they conjecture that exact asymptotics may not be attainable for the full class of semi-exponential distributions, they do provide logarithmic asymptotics—of the type we consider here—for the regime $\alpha \in (0, 1)$. In this work, we derive analogous logarithmic asymptotics under the assumption that the increments are semi-exponential with $\alpha \in (0, 1)$. Our contribution differs from [18] in the following ways: (a) We employ a Large Deviation Principle (LDP) as our main methodological tool; (b) We extend the analysis by allowing the random walk to be raised to an arbitrary positive power $p > 0$; (c) We do not assume that the increments have finite variance[1].

Our approach builds on the methodology developed in [5]. There a sample-path Large Deviation Principle (LDP) for the scaled random walk with light-tailed increments was employed. Here, we proceed in a similar spirit, making use of an LDP for scaled random walks with semi-exponential tails, developed in [4].

At first sight, one might expect that applying the contraction principle directly to this sample-path LDP would yield the desired asymptotics for the area under the first excursion (defined later via the functional ϕ_T). The difficulty is that ϕ_T is not continuous under the M_1' topology in which the LDP is formulated, and so the contraction principle does not apply directly.

Nevertheless, the sample-path LDP for semi-exponential increments remains a crucial ingredient in our analysis. The first issue is that it is stated only for random walks with non-negative increments. We overcome this by extending it to two-sided increments bounded from below. Later, in the proof of the main theorem, we further combine this with a truncation argument.

A central step in the proof is then the solution of the associated variational problem, which identifies the most likely path leading to the rare event of a very large area under the powered random walk during its first excursion. This optimal path turns out to be a manifestation of the "single big jump" phenomenon: the rare event is most likely realised by one large jump occurring at the origin, followed by a roughly linear decrease governed by the Law of Large Numbers.

Specifically, we prove the following:

Let $\{X_k,\, k \geq 0\}$ be a stochastic process defined by Lindley's recursion:

$$X_{k+1} = [X_k + U_{k+1}]^+, \quad k \geq 0, \quad X_0 = 0,$$

where $\{U_i\}_{i\geq 1}$ are i.i.d. random variables with mean $\mathbb{E}[U_1] = \mu < 0$. Define the first busy

[1] Semi-exponential distributions have finite variance, but since we do not impose any assumption on the left tail, the infinite variance might be due to the left tail. For example, we could assume regularly varying left tail with index $\alpha \in (1, 2]$.

period by

$$T_1 := \min\{k > 0 : X_k = 0\},$$

and consider the random variable

$$W_1(p) := \sum_{k=0}^{T_1} X_k^p,$$

for some $p > 0$.

Assume further that U_1 has a semi-exponential right tail. Then, as $t \to \infty$, the tail probability of W_1 satisfies the asymptotic equivalence

$$\log \mathbb{P}(W_1(p) \geq t) = \log \mathbb{P}\Big(U_1 > (|\mu|(p+1)t)^{1/(p+1)}\Big)\,(1 + o(1))\,.$$

We remark that this result was previously established in [18] in the special case $p = 1$, under the additional assumption that U_1 has finite variance.

B.2 Notation and Preliminaries

We begin by presenting the definition of a Large Deviation Principle (LDP). Let $(\mathcal{S}, d)$ denote a metric space, and let $\mathcal{T}$ denote the topology induced by the metric d. The interior and closure of a set A are denoted by A° and $\overline{A}$, respectively. Let (X_n) be a sequence of $\mathcal{S}$-valued random variables. Let $I : \mathcal{S} \to [0, \infty]$ be a lower semicontinuous function. Such a function is called a *rate function*, and it is said to be a *good* rate function if the level set $\Psi_I(\alpha) := \{x \in \mathcal{S} : I(x) \leq \alpha\}$ is compact for every $\alpha > 0$. Suppose that (a_n) is a sequence of positive real numbers such that $a_n \to \infty$ as $n \to \infty$.

Definition B.2.1. *We say that the sequence (X_n) satisfies a Large Deviation Principle (LDP) in $(\mathcal{S}, d)$ with speed a_n and rate function I if, for every measurable set $A \subseteq \mathcal{S}$,*

$$-\inf_{x \in A^{\circ}} I(x) \leq \liminf_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}(X_n \in A) \leq \limsup_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}(X_n \in A) \leq -\inf_{x \in \overline{A}} I(x).$$

The following definition, taken from Amir Dembo's Large Deviation Techniques and Applications, though more general than needed in our setting, will be useful in transferring a known LDP to a related sequence of measures.

Definition B.2.2. *Let $(\mathcal{S}, d)$ be a metric space. The probability measures $\{\mu_\epsilon\}$ and $\{\tilde{\mu}_\epsilon\}$ on $\mathcal{S}$ are said to be* exponentially equivalent *if there exist probability spaces $\{(\Omega, \mathcal{B}_\epsilon, P_\epsilon)\}$ and two families of $\mathcal{S}$-valued random variables $\{Z_\epsilon\}$ and $\{\tilde{Z}_\epsilon\}$, with joint laws $\{P_\epsilon\}$ and marginals $\{\mu_\epsilon\}$ and $\{\tilde{\mu}_\epsilon\}$, respectively, such that the following holds:*

For each $\delta > 0$, the set $\{\omega : (\tilde{Z}_\epsilon, Z_\epsilon) \in \Gamma_\delta\}$ is $\mathcal{B}_\epsilon$-measurable, and

$$\limsup_{\epsilon \to 0} \epsilon \log P_\epsilon(\Gamma_\delta) = -\infty,$$

where

$$\Gamma_\delta := \{(\tilde{s}, s) : d(\tilde{s}, s) > \delta\} \subset \mathcal{S} \times \mathcal{S}.$$

Let $\mathbb{D}[0,T]$ denote the Skorokhod space — the space of càdlàg paths from $[0,T]$ to $\mathbb{R}$. Define the subclass

$$\mathbb{D}^\mu[0,T] := \{\xi \in \mathbb{D}[0,T] : \xi \text{ is piecewise linear with slope } \mu \text{ and has positive jumps}\}.$$

Note that any càdlàg path has at most countably many discontinuities. Therefore, any $\xi \in \mathbb{D}^\mu[0,T]$ can be expressed as

$$\xi(s) = \mu s + \sum_{l \in [0,s]} (\xi(l) - \xi(l-)).$$

We assume that $\mathbb{D}[0,T]$ is equipped with the modified Skorokhod topology known as M_1'. For a formal definition of this topology, we proceed as follows.

Definition B.2.3. *For any $\xi \in \mathbb{D}[0,T]$, define the* extended completed graph $\Gamma'(\xi)$ *of ξ as*

$$\Gamma'(\xi) := \{(u,t) \in \mathbb{R} \times [0,T] : u \in [\xi(t-) \wedge \xi(t),\, \xi(t-) \vee \xi(t)]\},$$

*where **we set** $\xi(0-) := 0$. Define an order on $\Gamma'(\xi)$ by declaring $(u_1, t_1) < (u_2, t_2)$ if either $t_1 < t_2$, or $t_1 = t_2$ and $|\xi(t_1-) - u_1| < |\xi(t_2-) - u_2|$.*

A continuous non-decreasing function $(u(s), t(s)) : [0,T] \to \Gamma'(\xi)$ is called a parametrization *of $\Gamma'(\xi)$ if*

$$\Gamma'(\xi) = \{(u(s), t(s)) : s \in [0,T]\}.$$

We also call such a pair (u, t) a parametrization of the path ξ. The set of all such parametrizations of ξ is denoted by $\Pi'(\xi)$.

Definition B.2.4. *Define the M_1' metric $d_{M_1'}$ on $\mathbb{D}[0, T]$ as*

$$d_{M_1'}(\xi, \zeta) := \inf_{\substack{(u,t)\in\Pi'(\xi) \\ (v,r)\in\Pi'(\zeta)}} \left\{ \|u - v\|_\infty + \|t - r\|_\infty \right\}.$$

Let Ψ denote the reflection (Skorokhod) map, defined for any $\xi \in \mathbb{D}[0, T]$ by

$$\Psi(\xi)(t) := \xi(t) - \inf_{s\in[0,t]} \{\xi(s) \wedge 0\}.$$

We also define the following functional, which will be central to our analysis:

$$\phi_T(\xi) := \int_0^T \Psi(\xi)(s)^p \, ds.$$

We consider the time-homogeneous Markov chain $\{X_k,\ k \geq 0\}$ defined by Lindley's recursion:

$$X_{k+1} = [X_k + U_{k+1}]^+, \quad X_0 = 0,$$

where $\{U_i\}_{i\geq1}$ are i.i.d. random variables with $\mathbb{E}[U_1] = \mu < 0$. Define the first excursion period by

$$T_1 := \min\{k > 0 : X_k = 0\},$$

and the associated area under the random walk (raised to the power $p > 0$) by

$$W_1(p) := \sum_{k=0}^{T_1} X_k^p.$$

Furthermore, let $K_n := \sum_{i=1}^n U_i$ denote the unrestricted random walk. Notice that

$$K_n = X_n \quad \text{for } n < T_1. \tag{B.2.1}$$

Define the scaled version of K_n by

$$\bar{K}_x(t) := \frac{1}{x} K_{\lfloor xt \rfloor}, \quad t \in [0, T], \quad T \in \mathbb{R}.$$

We now introduce a standing assumption on the distribution of the increments:

Assumption B.2.5. *The right tail of U_1 satisfies*

$$\mathbb{P}(U_1 \geq x) = e^{-x^\alpha L(x)},$$

where $\alpha \in (0,1)$, L is a slowly varying function, and $x^{\alpha-1}L(x)$ is eventually decreasing.

B.3 Statement of Main Result

Based on the notation introduced above, we now present the main result of the chapter.

Theorem B.3.1. *Let $W_1(p) = \sum_{k=0}^{T_1} X_k^p$, where X is defined by Lindley's recursion, T_1 is the first busy period and $p > 0$. Let U_1 satisfy Assumption B.2.5, with $\mathbb{E}[U_1] = \mu < 0$. Then, the following logarithmic asymptotics hold:*

$$\lim_{t\to\infty} \frac{1}{t^{\frac{\alpha}{p+1}} L\left(t^{\frac{1}{p+1}}\right)} \log \mathbb{P}(W_1(p) \geq t) = -\left(|\mu|(p+1)\right)^{\frac{\alpha}{p+1}}.$$

Since L is slowly varying, the above result admits the equivalent formulation:

$$\log \mathbb{P}(W_1(p) \geq t) \sim \log \mathbb{P}\Big(U_1 > (|\mu|(p+1)t)^{1/(p+1)}\Big).$$

We note that for $p = 1$, this reproduces the same logarithmic asymptotics as in the regularly varying case studied in [10] and the semi-exponential case studied in [18].

The reformulation above suggests that the event $\{W_1(p) \geq t\}$ is dominated by the its most likely realization/path leading to such a large value. Specifically, the dominant contribution comes from the path which has a single large jump of size approximately $(|\mu|(p+1)t)^{1/(p+1)}$ occurring early, followed by a roughly linear decrease governed by the Law of Large Numbers.

The proof formalizes this intuition by showing that this is indeed the optimal (most likely) way in which the rare event $\{W_1(p) \geq t\}$ occurs. In this way, we not only establish the correct logarithmic scale for the tail probability, but also provide insight into the mechanism that drives the rare event.

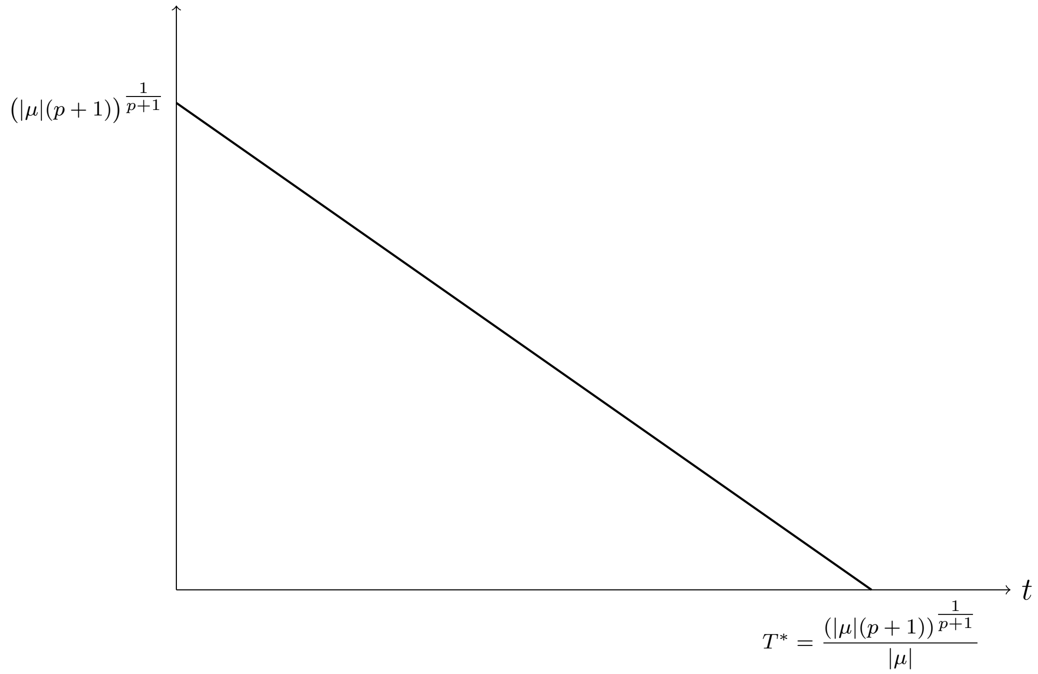

Figure B.1: Most likely realization: a single large jump at the origin followed by linear decay with slope μ.

B.4 Auxiliary Results

The following is a classical tool known as the Principle of the Maximum Term. We include it here for completeness. The version stated below can be found as Lemma 1.2.15 in Dembo and Zeitouni's Large Deviations Techniques and Applications.

Lemma B.4.1 (Principle of the Maximum Term)**.** *Let $N \in \mathbb{N}$ be fixed. Then, for any family of non-negative real numbers $\{a_\epsilon^i\}_{i=1}^N$, we have:*

$$\limsup_{\epsilon \to 0} \epsilon \log \left(\sum_{i=1}^N a_\epsilon^i \right) = \max_{1 \le i \le N} \limsup_{\epsilon \to 0} \epsilon \log a_\epsilon^i.$$

We will make use of the following result, which can be found in [3].

Theorem B.4.2. *Let $(U_i)_{i \ge 1}$ be a sequence of non-negative i.i.d. random variables satisfying Assumption B.2.5, with $\mathbb{E}[U_1] = \mu > 0$. Define*

$$\bar{S}_n(t) := \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} (U_i - \mu), \quad t \in [0, T],$$

for some fixed $T > 0$. Then, the process $\bar{S}_n$ satisfies a Large Deviation Principle in

$(\mathbb{D}[0,T], \mathcal{T}_{M_1'})$ *with speed* $L(n)n^\alpha$ *and good rate function* $I_{M_1'}^{(0)} : \mathbb{D}[0,T] \to [0,\infty]$ *given by*

$$I_{M_1'}^{(0)}(\xi) = \begin{cases} \sum_{t\in[0,T]} \big(\xi(t) - \xi(t-)\big)^\alpha, & \text{if } \xi \in \mathbb{D}^0[0,T] \text{ and } \xi(0) \geq 0, \\ \infty, & \text{otherwise.} \end{cases}$$

Remark B.4.3. *Recall that every path* $\xi \in (\mathbb{D}[0,T], M_1')$ *is assumed to satisfy* $\xi(0^-) = 0$*. Therefore, if* $\xi(0) > 0$*, this is treated as a jump at time* 0 *and contributes to the rate function* $I_{M_1'}^{(0)}$ *as the first summand.*

The following extension of Theorem B.4.2 to the case where U_1 may take negative values of bounded magnitude is almost immediate.[2]

Theorem B.4.4. *Let* $(U_i)_{i\geq 1}$ *be a sequence of i.i.d. random variables satisfying Assumption B.2.5, with* $\mathbb{E}[U_1] = \mu < 0$*. Additionally, assume there exists* $K > 0$ *such that* $U_1 \geq -K$ *almost surely. Then, the scaled random walk*

$$\bar{K}_n(t) := \frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor} U_i, \quad t \in [0,T],$$

satisfies a Large Deviation Principle in $(\mathbb{D}[0,T], \mathcal{T}_{M_1'})$*, with speed* $L(n)n^\alpha$ *and good rate function* $I_{M_1'}^{(\mu)} : \mathbb{D}[0,T] \to [0,\infty]$ *defined by*

$$I_{M_1'}^{(\mu)}(\xi) = \begin{cases} \sum_{t\in[0,T]} \big(\xi(t) - \xi(t-)\big)^\alpha, & \text{if } \xi \in \mathbb{D}^\mu[0,T] \text{ and } \xi(0) \geq 0, \\ \infty, & \text{otherwise.} \end{cases}$$

Proof. The result follows by applying Theorem B.4.2 to the random walk with non-negative i.i.d. increments $(V_n)_{n\geq 1}$, with $V_1 := U_1 + K$. We establish the upper bound; the lower bound follows similarly.

Note that V_1 satisfies Assumption B.2.5, having a semi-exponential right tail with the same index α and a slowly varying function

$$L'(x) := L(x-K) \cdot \frac{(x-K)^\alpha}{x^\alpha}. \tag{B.4.1}$$

Moreover, $\mathbb{E}[V_1] = \mu + K > 0$ and $V_1 \geq 0$.

[2] In [4], it is remarked that the result holds for two-sided jumps if the left tail is light. However, the argument there contains an error; in discussion with one of the authors this was confirmed.

Therefore, for any measurable set $A \subseteq \mathbb{D}[0,T]$, Theorem B.4.2 yields

$$\limsup_{n\to\infty} \frac{1}{L'(n)n^\alpha} \log \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor}(V_i - \mathbb{E}[V_i]) \in A \right) \le -\inf_{\xi\in\bar{A}} I^{(0)}_{M_1'}(\xi). \tag{B.4.2}$$

Using the definition of V_i, this becomes

$$\limsup_{n\to\infty} \frac{1}{L'(n)n^\alpha} \log \mathbb{P}\Big(\bar{K}_n(t) - \mu\cdot \tfrac{\lfloor nt \rfloor}{n} \in A \Big) \le -\inf_{\xi\in\bar{A}} I^{(0)}_{M_1'}(\xi). \tag{B.4.3}$$

Now, observe that

$$\Big\| \mu\cdot \tfrac{\lfloor nt \rfloor}{n} - \mu t \Big\|_\infty \le |\mu|\frac{1}{n},$$

which implies that the sequences of distribution measures of $\bar{K}_n(t) - \mu\cdot \frac{\lfloor nt \rfloor}{n}$ and $\bar{K}_n(t) - \mu t$ are exponentially equivalent. Therefore, by Theorem 4.2.13 in *Large Deviations Techniques and Applications* by Dembo and Zeitouni, we can replace $\bar{K}_n(t) - \mu\cdot \frac{\lfloor nt \rfloor}{n}$ with $\bar{K}_n(t) - \mu t$ in (B.4.3), to obtain

$$\limsup_{n\to\infty} \frac{1}{L'(n)n^\alpha} \log \mathbb{P}\big(\bar{K}_n(t) - \mu t \in A \big) \le -\inf_{\xi\in\bar{A}} I^{(0)}_{M_1'}(\xi). \tag{B.4.4}$$

The result follows because the translation $\xi \mapsto \xi + \mu t$ shifts the domain of $I^{(0)}_{M_1'}$ to that of $I^{(\mu)}_{M_1'}$. Formally, for any measurable $A \subseteq \mathbb{D}[0,T]$,

$$\inf_{\xi\in\overline{A+\mu t}} I^{(0)}_{M_1'}(\xi) = \inf_{\xi\in\bar{A}} I^{(\mu)}_{M_1'}(\xi),$$

where $A + \mu t := \{\xi + \mu t : \xi \in A\}$. Finally, since $L'(x) \sim L(x)$ as $x \to \infty$ by L being slowly varying, the claimed LDP with speed $L(n)n^\alpha$ follows. □

To apply the above LDP, we must solve a variational problem. Specifically, we consider the following question:

Among all paths in $\mathbb{D}^\mu[0,T]$ with the same total jump magnitude, which one maximizes the area under its graph after applying the reflection map and raising it to a positive power?

The answer is that the optimal path is the one that concentrates all of its jumps at the origin.

To explain this more precisely, recall the reflection map $\Psi(\xi)(t) := \xi(t) - \inf_{s\in[0,t]}\{\xi(s)\wedge 0\}$, and the functional $\phi_T(\xi) := \int_0^T \Psi(\xi)(s)^p\, ds$. Let $\xi \in \mathbb{D}^\mu[0,T]$, which, by definition, takes the form

$$\xi(t) = \mu t + \sum_{l\in[0,t]} \big(\xi(l) - \xi(l-)\big).$$

Now consider the path $\tilde{\xi}$ obtained by shifting all the jumps of ξ to the origin:

$$\tilde{\xi}(t) := \mu t + \sum_{l\in[0,T]} \big(\xi(l) - \xi(l-)\big), \quad t \in [0,T].$$

Then, the inequality

$$\phi_T(\xi) \le \phi_T(\tilde{\xi})$$

holds. That is, concentrating the jumps at the beginning maximizes the area under the reflected and powered path. This claim is established in Lemma B.4.6.

To facilitate our analysis, we introduce additional notation and then prove an auxiliary result, Lemma B.4.5. Although this lemma could have been stated as a list of straightforward facts, and its claims can be readily deduced from the definitions below—especially with the aid of a graphical illustration like Figure B.2—we provide detailed proofs for the sake of completeness and rigor.

Let $\xi \in \mathbb{D}^\mu[0,T]$ be a path such that $\xi \ge 0$ and

$$\operatorname{supp}\Psi(\xi) := \{t \in [0,T] : \Psi(\xi)(t) > 0\} \ne \emptyset.$$

For each $t \in \operatorname{supp}\Psi(\xi)$, define the entrance and exit time of the excursion containing t,

$$l(t) := \inf\{s \in [0,T] : \forall u \in [s,t],\ \Psi(\xi)(u) > 0\},$$

$$r(t) := \sup\{s \in [0,T] : \forall u \in [t,s],\ \Psi(\xi)(u) > 0\},$$

and let $\sigma(t) := [l(t), r(t)]$ denote the excursion interval containing t or else the maximal interval containing t on which $\Psi(\xi)$ remains strictly positive.

Now define the set of such intervals:

$$C^+ := \{\sigma(t) \subseteq [0,T] : t \in \operatorname{supp}\Psi(\xi)\}.$$

By construction, the elements of C^+ are disjoint intervals, each corresponding to a maximal positive excursion of $\Psi(\xi)$, and hence C^+ is at most countable. We enumerate them as

$$C^+ = \{[l_i, r_i] : i \geq 1,\ l_i < l_j \text{ for } i < j\}.$$

Additionally, define $r_0 := 0$, and let

$$C^0 := \{(r_{i-1}, l_i) : i \geq 1\},$$

the collection of intervals on which $\Psi(\xi)$ vanishes identically between successive positive excursions.

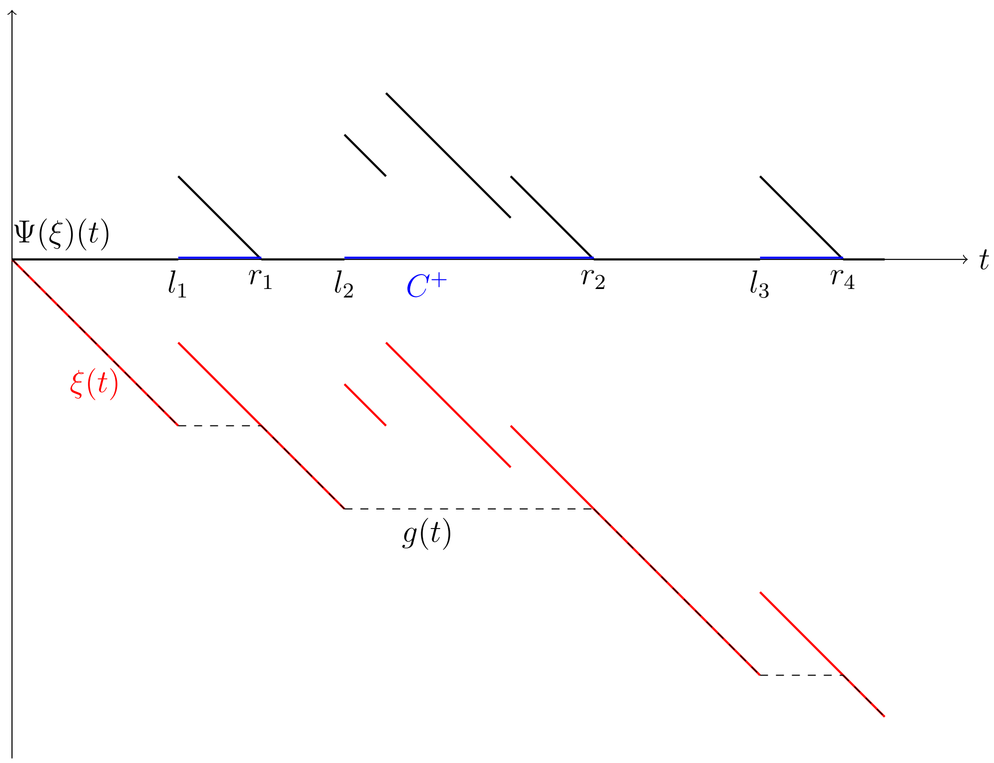

Figure B.2: Example paths: $\xi(t)$ (red), and reflected $\Psi(\xi)(t)$ (black), $g(t)$(dashed). Blue segments indicate elements of C^+.

Lemma B.4.5. *Let $\xi \in \mathbb{D}^\mu[0,T]$ with $\xi(0) \geq 0$. Denote the running infimum of $\xi \wedge 0$ by*

$$g(t) := \inf_{s\in[0,t]} \{\xi(s) \wedge 0\}.$$

Recall that $\Psi(\xi)(t) := \xi(t) - g(t)$, and let C^+ and C^0 be as defined above. Then, the following properties hold:

1. ***Properties of the running infimum*** g***:***

   (i) ξ *does not jump on any interval in* C^0*.*

   (ii) g *is continuous on* $[0, T]$*.*

   (iii) g *is constant on each interval in* C^+*, i.e.,* $g(t) = g(l_i)$ *for all* $t \in [l_i, r_i]$ *and* $i \leq |C^+|$*.*

   (iv) g *is linear with slope* μ *on every interval in* C^0*. In particular, for every* $i \leq |C^+|$*,*

   $$g(r_{i-1}) - g(l_i) = \mu(l_i - r_{i-1}).$$

   (v) For every $i \leq |C^+|$ *and all* $t \in [l_i, r_i)$*, it holds that*

   $$g(t) = \mu \sum_{j=1}^{i} (l_j - r_{j-1}).$$

2. ***Properties of the reflected path*** $\Psi(\xi)$***:***

   (i) $\Psi(\xi)(t) = 0$ *for every* $t \in [r_{i-1}, l_i)$ *and every* $i \leq |C^+|$*. In particular,* $\Psi(\xi)(r_i) = 0$ *for every* $r_i < T$*.*

   (ii) For every $i \leq |C^+|$ *and* $t \in [l_i, r_i)$*,*

   $$\Psi(\xi)(t) = \mu t + \sum_{l \in [0,t]} (\xi(l) - \xi(l-)) - \mu \sum_{j=1}^{i} (l_j - r_{j-1}).$$

   (iii) For every $i \leq |C^+|$*, the following holds:*

   $$r_i \mu = \begin{cases} -\left(\sum_{l \in [0, r_i]} \xi(l) - \xi(l-) - \mu \sum_{j=1}^{i} (l_j - r_{j-1}) \right), & \text{if } r_i < T, \\ \geq -\left(\sum_{l \in [0, r_i]} \xi(l) - \xi(l-) - \mu \sum_{j=1}^{i} (l_j - r_{j-1}) \right), & \text{if } r_i = T. \end{cases}$$

   (iv) For every $i \leq |C^+|$*, the integral of* $\Psi(\xi)^p$ *over the excursion satisfies*

   $$\int_{l_i}^{r_i} \Psi(\xi)(s)^p \, ds \leq \frac{\left(\sum_{l \in (r_{i-1}, r_i]} \xi(l) - \xi(l-) \right)^{p+1}}{-\mu(p+1)}.$$

Proof. 1. (i) Let $i \leq |C^+|$ and $l \in (r_{i-1}, l_i)$. By the construction of C^+, we have $\Psi(\xi)(l) = 0$. Combined with the definition of Ψ, this implies that ξ attains a running infimum at l, i.e., $\xi(l) \leq \xi(l-)$. On the other hand, since ξ has only positive jumps, we also have $\xi(l) \geq \xi(l-)$. Therefore, it must be that $\xi(l) = \xi(l-)$, and hence ξ has no jump at l (and is continuous at l).

(ii) The function $g(t) := \inf_{s\in[0,t]}\{\xi(s) \wedge 0\}$ is non-increasing by definition. If g were discontinuous at some $l \in [0, T]$, it could only have a jump discontinuity, i.e., $g(l^-) > g(l)$ or $g(l^+) < g(l)$. However, the former contradicts the fact that $\xi \wedge 0$ only has positive jumps, while the latter contradicts the right-continuity of $\xi \wedge 0$. Hence, g is continuous on $[0, T]$.

(iii) Let $t \in [l_i, r_i)$ and suppose for contradiction that $g(t) < g(l_i)$. Then $\xi \wedge 0$ would hit a new running infimum at some $u \in [l_i, t)$, which would imply $\Psi(\xi)(u) = 0$. But by construction of C^+, $\Psi(\xi)(u) > 0$ for all $u \in [l_i, r_i)$. This contradiction shows that $g(t) = g(l_i)$ for all $t \in [l_i, r_i)$. Since g is continuous by part (ii), we also get $g(r_i) = g(l_i)$.

(iv) Let $i \leq |C^+|$. From part (i), we know that ξ is continuous on (r_{i-1}, l_i), so for any $s, u \in (r_{i-1}, l_i)$ with $s < u$, we have

$$\xi(u) - \xi(s) = \mu(u - s).$$

Furthermore, since $\Psi(\xi)(t) = 0$ for all $t \in (r_{i-1}, l_i)$ by construction of C^+, it follows that $g(t) = \xi(t)$ on this interval. Hence, $g(u) - g(s) = \mu(u - s)$, i.e., g is linear with slope μ on (r_{i-1}, l_i). By continuity (part (ii)), this extends to the endpoints, so:

$$g(l_i) - g(r_{i-1}) = \mu(l_i - r_{i-1}).$$

(v) The claim follows by g being constant on the intervals in C^+ (part (iii)) and decreasing linearly with slope μ on the intervals in C^0 (part (iv)). More analytically, for $t \in [l_i, r_i)$, by summing the increments of g over the partition induced by the alternating intervals in C^+ and C^0:

$$g(t) = g(r_0) + \sum_{j=1}^{i} (g(l_j) - g(r_{j-1})) + \sum_{j=1}^{i-1} (g(r_j) - g(l_j)) + g(t) - g(r_i). \tag{B.4.5}$$

Since $\xi \geq 0$, we have $g(r_0) = g(0) = 0$. By part (iii), the second sum and the final term vanish, and by part (iv), we have $g(l_j) - g(r_{j-1}) = \mu(l_j - r_{j-1})$. Substituting into (B.4.5) yields:

$$g(t) = \mu \sum_{j=1}^{i} (l_j - r_{j-1}),$$

as claimed.

2. (i) By the construction of C^+, we have $\Psi(\xi)(t) = 0$ for every $t \in (r_{i-1}, l_i)$. At the same time, $\Psi(\xi)$ is right-continuous, as it is the difference of two right-continuous functions. Hence, $\Psi(\xi)(r_{i-1}) = 0$.

   (ii) The claim follows directly from the definition of $\Psi(\xi)$, the form of ξ, and part 1(v).

   (iii) Fix $i \leq |C^+|$. If $r_i < T$, then $\Psi(\xi)(r_i) = 0$ by (i), which implies $\xi(r_i) = g(r_i)$. Using part 1(v), we get

$$\xi(r_i) = \mu \sum_{j=1}^{i} (l_j - r_{j-1}).$$

   On the other hand, by the definition of ξ, we have

$$\xi(r_i) = \mu r_i + \sum_{l \in [0, r_i]} (\xi(l) - \xi(l-)) .$$

   Equating the two expressions for $\xi(r_i)$ and solving for μr_i gives the desired identity.

   If $r_i = T$ and $\Psi(\xi)(r_i) = 0$, the same argument applies. If instead $\Psi(\xi)(r_i) > 0$, then $g(r_i) < \xi(r_i)$, and combining this with the second expression above yields an inequality for μr_i, completing the proof in both cases.

   (iv) From 2(ii), we have

$$\Psi(\xi)(t)\mathbb{1}_{[l_i, r_i)}(t) = \left(\mu t + \sum_{l \in [0,t]} (\xi(l) - \xi(l-)) - \mu \sum_{j=1}^{i} (l_j - r_{j-1}) \right) \mathbb{1}_{[l_i, r_i)}(t).$$

   Since the jumps are positive, for $t < r_i$, we can bound this above by replacing

t with r_i in the jump sum:

$$\Psi(\xi)(t)^p \mathbb{1}_{[l_i,r_i)}(t) \leq \left(\mu t + \sum_{l\in[0,r_i]} (\xi(l) - \xi(l-)) - \mu \sum_{j=1}^{i} (l_j - r_{j-1}) \right)^p \mathbb{1}_{[l_i,r_i)}(t).$$

Therefore,

$$\int_{l_i}^{r_i} \Psi(\xi)(s)^p ds \leq \int_{l_i}^{r_i} \left(\mu s + \sum_{l\in[0,r_i]} (\xi(l) - \xi(l-)) - \mu \sum_{j=1}^{i} (l_j - r_{j-1}) \right)^p ds. \tag{B.4.6}$$

Set

$$u(s) := \mu s + \sum_{l\in[0,r_i]} (\xi(l) - \xi(l-)) - \mu \sum_{j=1}^{i} (l_j - r_{j-1}),$$

so that $du = \mu\, ds$. Then, the limits of integration become:

$$u(l_i) = \sum_{l\in(r_{i-1},r_i]} (\xi(l) - \xi(l-)),$$
$$u(r_i) = 0 \quad \text{if } r_i < T, \quad \text{and } u(r_i) \geq 0 \text{ if } r_i = T,$$

by 2(iii). Plugging into (B.4.6):

$$\begin{aligned}\int_{l_i}^{r_i} \Psi(\xi)(s)^p ds &\leq \frac{1}{-\mu} \int_{u(r_i)}^{u(l_i)} u^p du \\ &\leq \frac{1}{-\mu} \int_{0}^{\sum_{l\in(r_{i-1},r_i]}(\xi(l)-\xi(l-))} u^p du \\ &= \frac{\left(\sum_{l\in(r_{i-1},r_i]} (\xi(l) - \xi(l-))\right)^{p+1}}{-\mu(p+1)}.\end{aligned}$$

□

Lemma B.4.6. *Let $p > 0$, $T > 0$, and let $\xi \in \mathbb{D}^\mu[0,T]$ be such that $\xi(0) \geq 0$. Then, the following inequality holds:*

$$\phi_T(\xi) \leq \frac{\left(\sum_{s\in[0,T]} \xi(s) - \xi(s-)\right)^{p+1}}{-\mu(p+1)}.$$

Proof. We have

$$\phi_T(\xi) = \sum_{i=1}^{|C^+|} \int_{l_i}^{r_i} \Psi(\xi)(s)^p \, ds \le \frac{1}{-\mu(p+1)} \sum_{i=1}^{|C^+|} \left(\sum_{l \in (r_{i-1}, r_i]} \xi(l) - \xi(l-) \right)^{p+1}$$
$$\le \frac{1}{-\mu(p+1)} \left(\sum_{s \in [0,T]} \xi(s) - \xi(s-) \right)^{p+1} .$$

The equality follows from the fact that $\Psi(\xi)$ is positive only on the intervals in C^+. The first inequality is due to Lemma B.4.5(2.iv), and the second inequality uses the convexity of the function $x \mapsto x^{p+1}$ for $p+1 > 1$. □

Proposition B.4.7. *Let $T > 0$. Let U_1 satisfy $\mathbb{E}[U_1] = \mu < 0$ and Assumption B.2.5. Assume further that there exists $K > 0$ such that $U_1 \ge -K$. Then, for the scaled random walk $\bar{K}_x(t) := \frac{1}{x} \sum_{i=1}^{[xt]} U_i$, it holds that*

$$\limsup_{x \to \infty} \frac{1}{x^\alpha L(x)} \log \mathbb{P}\big(\phi_T(\bar{K}_x) \ge 1\big) \le - \left(|\mu|(p+1)\right)^{\frac{\alpha}{p+1}} .$$

Proof. Note that ϕ_T is continuous with respect to the M_1 topology (see Theorems 11.5.1 and 13.4.1 in [43]). Hence, the set

$$\mathcal{V}^T := \{\xi \in \mathbb{D}[0,T] : \phi_T(\xi) \ge 1\}$$

is closed in $(\mathbb{D}[0,T], \mathcal{T}_{M_1'})$. Therefore, by the LDP in Theorem B.4.4,

$$\limsup_{x \to \infty} \frac{1}{x^\alpha L(x)} \log \mathbb{P}\big(\phi_T(\bar{K}_x) \ge 1\big) = \limsup_{x \to \infty} \frac{1}{x^\alpha L(x)} \log \mathbb{P}\big(\bar{K}_x \in \mathcal{V}^T\big)$$
$$\le - \inf_{\xi \in \mathcal{V}^T} I_{M_1'}^{(\mu)}(\xi).$$

Thus, it suffices to show that

$$I_{M_1'}^{(\mu)}(\xi) \ge \left(-\mu(p+1)\right)^{\frac{\alpha}{p+1}} \quad \text{for all } \xi \in \mathcal{V}^T.$$

Let $\xi \in \mathcal{V}^T$. If $\xi \notin \mathbb{D}^\mu[0,T]$ or $\xi(0) < 0$, then $I_{M_1'}^{(\mu)}(\xi) = \infty$ and the claim follows. Hence, we assume $\xi \in \mathbb{D}^\mu[0,T]$, $\xi(0) \ge 0$, and $\phi_T(\xi) \ge 1$.

By Lemma B.4.6,

$$1 \leq \frac{\left(\sum_{s\in[0,T]} \xi(s) - \xi(s-)\right)^{p+1}}{-\mu(p+1)}.$$

Rewriting the above inequality and raising both sides to the power $\alpha \in (0,1)$, we obtain

$$\left(\sum_{s\in[0,T]} \xi(s) - \xi(s-)\right)^{\alpha} \geq (-\mu(p+1))^{\frac{\alpha}{p+1}}. \tag{B.4.7}$$

Now, since $\alpha \in (0,1)$, we use the concavity of $x \mapsto x^\alpha$ and Jensen's inequality in reverse form to get

$$\sum_{s\in[0,T]} (\xi(s) - \xi(s-))^\alpha \geq \left(\sum_{s\in[0,T]} \xi(s) - \xi(s-)\right)^{\alpha}. \tag{B.4.8}$$

Combining (B.4.7) and (B.4.8) yields

$$I^{(\mu)}_{M_1'}(\xi) := \sum_{s\in[0,T]} (\xi(s) - \xi(s-))^\alpha \geq (-\mu(p+1))^{\frac{\alpha}{p+1}}.$$

□

Recall that T_1 denotes the first excursion of random walk.

Proposition B.4.8. *Let U_1 satisfy $\mathbb{E}[U_1] = \mu < 0$ and Assumption B.2.5. Assume further that there exists $K > 0$ such that $U_1 \geq -K$. Define*

$$T^* := \frac{(|\mu|(p+1))^{\frac{1}{p+1}}}{|\mu|},$$

and assume $T \geq T^*$. Then,

$$\limsup_{x\to\infty} \frac{1}{x^\alpha L(x)} \log \mathbb{P}(T_1 > xT) \leq -(|\mu|(p+1))^{\frac{\alpha}{p+1}}.$$

Proof. Define the set

$$\mathcal{B} := \left\{\xi \in \mathbb{D}[0,T] : \inf_{s\in[0,T]} \xi(s) \geq 0\right\}.$$

If $T_1 > xT$, then $X_n = K_n \geq 0$ for all $n \leq \lfloor xT \rfloor$ (by $X_n = K_n$ before the first return), hence $\bar{K}_x \in \mathcal{B}$. Therefore,

$$\begin{aligned}\limsup_{x\to\infty} \frac{1}{x^\alpha L(x)} \log \mathbb{P}(T_1 > xT) &\le \limsup_{x\to\infty} \frac{1}{x^\alpha L(x)} \log \mathbb{P}\big(\bar{K}_x \in \mathcal{B}\big) \\ &\le -\inf_{\xi\in\mathcal{B}} I^{(\mu)}_{M_1'}(\xi),\end{aligned} \tag{B.4.9}$$

where the first inequality holds due to $\mathbb{P}(T_1 > xT) \le \mathbb{P}\big(\bar{K}_x \in \mathcal{B}\big)$, and the second follows from Theorem B.4.4, because $\mathcal{B}$ is closed (as the pointwise infimum is continuous under the M_1 topology — see Theorem 13.4.1 in [43]).

To conclude, it suffices to show that for every $\xi \in \mathcal{B}$,

$$I^{(\mu)}_{M_1'}(\xi) \ge (-\mu(p+1))^{\frac{\alpha}{p+1}}.$$

Let $\xi \in \mathcal{B}$. If $\xi \notin \mathbb{D}^\mu[0,T]$, then $I^{(\mu)}_{M_1'}(\xi) = \infty$, so we may assume $\xi \in \mathbb{D}^\mu[0,T]$. Then,

$$\xi(s) = \mu s + \sum_{l\in[0,s]} \left(\xi(l) - \xi(l-)\right).$$

Evaluating at time $T^* \le T$, we get:

$$\begin{aligned}0 \le \xi(T^*) &= \mu T^* + \sum_{l\in[0,T^*]} (\xi(l) - \xi(l-)) \\ &\le -(-\mu(p+1))^{\frac{1}{p+1}} + \sum_{l\in[0,T]} \left(\xi(l) - \xi(l-)\right),\end{aligned}$$

where the first inequality follows from $\xi \in \mathcal{B}$, and the second from the definition of T^* and the assumption $T \ge T^*$.

Rearranging gives

$$\sum_{l\in[0,T]} (\xi(l) - \xi(l-)) \ge (-\mu(p+1))^{\frac{1}{p+1}}.$$

Raising both sides to the power $\alpha \in (0,1)$ and using inequality (B.4.8) from Proposition B.4.7, we conclude:

$$I^{(\mu)}_{M_1'}(\xi) := \sum_{l\in[0,T]} (\xi(l) - \xi(l-))^\alpha \ge \left(\sum_{l\in[0,T]} (\xi(l) - \xi(l-))\right)^\alpha \ge (-\mu(p+1))^{\frac{\alpha}{p+1}}.$$

□

B.5 Proof of Main Result

Proof of Theorem B.3.1. Let $K > 0$ and define the truncated increments $U_i^{(K)} := \max(U_i, -K)$, $i \geq 1$. Let

$$X_{i+1}^{(K)} := [X_i^{(K)} + U_{i+1}^{(K)}]^+, \qquad X_0^{(K)} := 0,$$

with first return time $T_1^{(K)}$, and set

$$W_1^{(K)}(p) := \sum_{i=0}^{T_1^{(K)}} \big(X_i^{(K)}\big)^p, \qquad \mu_K := \mathbb{E}\Big[U_1^{(K)}\Big], \qquad T_K^* := \frac{(|\mu_K|(p+1))^{1/(p+1)}}{|\mu_K|}.$$

Note that $U_1^{(K)}$ and U_1 have the same right tail; hence $U_1^{(K)}$ is also semi-exponential with index α and slowly varying function L.

Upper bound.

$$\begin{aligned}
&\limsup_{t\to\infty} \frac{1}{t^{\frac{\alpha}{p+1}} L(t^{\frac{1}{p+1}})} \log \mathbb{P}(W_1(p) \geq t) \\
\leq &\limsup_{t\to\infty} \frac{1}{t^{\frac{\alpha}{p+1}} L(t^{\frac{1}{p+1}})} \log \mathbb{P}\Big(W_1^{(K)}(p) \geq t\Big) \\
= &\limsup_{x\to\infty} \frac{1}{x^\alpha L(x)} \log \mathbb{P}\left(\sum_{i=0}^{T_1^{(K)}} (X_i^{(K)})^p \geq x^{1+p}\right) \\
= &\limsup_{x\to\infty} \frac{1}{x^\alpha L(x)} \log \mathbb{P}\left(\int_0^{T_1^{(K)}/x} \left(\frac{X_{\lfloor xs \rfloor}^{(K)}}{x}\right)^p ds \geq 1\right) \\
\leq &\limsup_{x\to\infty} \frac{1}{x^\alpha L(x)} \log \left(\mathbb{P}\Big(\textstyle\int_0^{T_1^{(K)}/x} (\frac{X_{\lfloor xs \rfloor}^{(K)}}{x})^p ds \geq 1,\, T_1^{(K)}/x \leq T_K^*\Big) + \mathbb{P}\Big(T_1^{(K)}/x > T_K^*\Big)\right) \\
\leq &\left[\limsup_{x\to\infty} \frac{1}{x^\alpha L(x)} \log \mathbb{P}\left(\int_0^{T_K^*} \left(\frac{X_{\lfloor xs \rfloor}^{(K)}}{x}\right)^p ds \geq 1\right)\right] \vee \left[\limsup_{x\to\infty} \frac{1}{x^\alpha L(x)} \log \mathbb{P}\Big(T_1^{(K)} > xT_K^*\Big)\right] \\
= &\left[\limsup_{x\to\infty} \frac{1}{x^\alpha L(x)} \log \mathbb{P}\big(\phi_{T_K^*}(\bar{K}_x^{(K)}) \geq 1\big)\right] \vee \left[\limsup_{x\to\infty} \frac{1}{x^\alpha L(x)} \log \mathbb{P}\Big(T_1^{(K)} > xT_K^*\Big)\right] \\
\leq &- \big(|\mu_K|(p+1)\big)^{\frac{\alpha}{p+1}} \vee -\big(|\mu_K|(p+1)\big)^{\frac{\alpha}{p+1}} \\
= &- \big(|\mu_K|(p+1)\big)^{\frac{\alpha}{p+1}},
\end{aligned}$$

where the first inequality uses $W_1^{(K)}(p) \geq W_1(p)$ almost surely, the second equality uses $t = x^{1+p}$, the third equality uses

$$\frac{1}{x^{p+1}} \sum_{i=0}^{m-1} \big(X_i^{(K)}\big)^p = \frac{1}{x^{p+1}} \int_0^m \big(X_{\lfloor u \rfloor}^{(K)}\big)^p \, du = \int_0^{m/x} \left(\frac{X_{\lfloor xs \rfloor}^{(K)}}{x} \right)^p ds, \tag{B.5.1}$$

the second inequality is the union bound, the third inequality due to the principle of the maximum term (Lemma B.4.1), and the last inequality follows from Propositions B.4.7 and B.4.8. Here

$$\bar{K}_x^{(K)}(t) \ := \ \frac{1}{x} \sum_{i=1}^{\lfloor xt \rfloor} U_i^{(K)}, \qquad t \in [0, T_K^*].$$

Lower bound.

Fix $\epsilon > 0$ and define

$$A_\epsilon := \Big\{ \xi \in \mathbb{D}[0, T_K^*] : \ \xi(0) = \epsilon, \ \int_0^{T_K^*} \Psi(\xi)(s)^p ds > 1, \ \inf_{s \in [0, T_K^*]} \xi(s) > 0 \Big\},$$
$$\bar{A}_\epsilon := \Big\{ \xi \in \mathbb{D}[0, T_K^*] : \ \xi(0) = \epsilon, \ \int_0^{T_K^*} \Psi(\xi)(s)^p ds > 1, \ \inf_{s \in [0, T_K^*]} \xi(s) > \epsilon/2 \Big\}.$$

Let $b := (|\mu_K|(p+1))^{1/(p+1)}$ and set

$$\xi(s) := \epsilon + b \, \mathbb{1}_{[\epsilon/(4|\mu_K|), T_K^*]}(s) + \mu_K s.$$

Then $\xi(0) = \epsilon$, $\inf_{[0, T_K^*]} \xi = 3\epsilon/4$, and using $\mu_K T_K^* + b = 0$,

$$\begin{aligned} \int_0^{T_K^*} \Psi(\xi)(s)^p ds &= \int_0^{\frac{\epsilon}{4|\mu_K|}} (\mu_K s + \epsilon)^p ds + \int_{\frac{\epsilon}{4|\mu_K|}}^{T_K^*} (\mu_K s + b + \epsilon)^p ds \\ &= \frac{\epsilon^{p+1} - (\frac{3}{4}\epsilon)^{p+1}}{|\mu_K|(p+1)} + \frac{(\frac{3}{4}\epsilon + b)^{p+1} - \epsilon^{p+1}}{|\mu_K|(p+1)} \\ &= \frac{(\frac{3}{4}\epsilon + b)^{p+1} - (\frac{3}{4}\epsilon)^{p+1}}{|\mu_K|(p+1)} > \frac{b^{p+1}}{|\mu_K|(p+1)} = 1, \end{aligned}$$

where we used the definition of b in the last equality. Hence $\xi \in \bar{A}_\epsilon$, and

$$I_{M_1'}^{(\mu_K)}(\xi) = \epsilon^\alpha + b^\alpha = \epsilon^\alpha + \big(|\mu_K|(p+1)\big)^{\frac{\alpha}{p+1}},$$

so

$$\inf_{\xi \in \bar{A}_\epsilon} I_{M_1'}^{(\mu_K)}(\xi) \ \leq \ \epsilon^\alpha + \big(|\mu_K|(p+1)\big)^{\frac{\alpha}{p+1}}. \tag{B.5.2}$$

Define $B_{x,\epsilon} := \{U_i^{(K)} > \sqrt{\epsilon},\ \forall i \leq \lceil x\sqrt{\epsilon}\rceil\}$ and $k^* := \lceil x\sqrt{\epsilon}\rceil + 1$. Then

$$
\begin{aligned}
&\liminf_{t\to\infty} \frac{1}{t^{\frac{\alpha}{p+1}} L\left(t^{\frac{1}{p+1}}\right)} \log \mathbb{P}(W_1(p) \geq t) \\
\geq\ &\liminf_{t\to\infty} \frac{1}{t^{\frac{\alpha}{p+1}} L\left(t^{\frac{1}{p+1}}\right)} \log \mathbb{P}(W_1(p) \geq t,\ U_i \geq -K\ \forall i \leq T_1) \\
\geq\ &\liminf_{t\to\infty} \frac{1}{t^{\frac{\alpha}{p+1}} L(t^{\frac{1}{p+1}})} \log \mathbb{P}(W_1(p) \geq t \mid U_i \geq -K\ \forall i \leq T_1) \\
&+ \liminf_{t\to\infty} \frac{1}{t^{\frac{\alpha}{p+1}} L(t^{\frac{1}{p+1}})} \log \mathbb{P}(U_i \geq -K\ \forall i \leq T_1) \\
=\ &\liminf_{t\to\infty} \frac{1}{t^{\frac{\alpha}{p+1}} L\left(t^{\frac{1}{p+1}}\right)} \log \mathbb{P}\left(W_1^{(K)}(p) \geq t\right) + 0 \\
=\ &\liminf_{x\to\infty} \frac{1}{x^\alpha L(x)} \log \mathbb{P}\left(\sum_{i=0}^{T_1^{(K)}} (X_i^{(K)})^p \geq x^{1+p}\right) \\
\geq\ &\liminf_{x\to\infty} \frac{1}{x^\alpha L(x)} \log \mathbb{P}\left(\sum_{i=k^*}^{T_1^{(K)}} (X_i^{(K)})^p \geq x^{1+p},\ B_{x,\epsilon}\right) \\
=\ &\liminf_{x\to\infty} \frac{1}{x^\alpha L(x)} \log\left[\mathbb{P}\left(\sum_{i=k^*}^{T_1^{(K)}} (X_i^{(K)})^p \geq x^{1+p} \,\middle|\, B_{x,\epsilon}\right) \mathbb{P}(B_{x,\epsilon})\right] \\
\geq\ &\liminf_{x\to\infty} \frac{1}{x^\alpha L(x)} \log\left[\mathbb{P}\left(\sum_{i=0}^{T_1^{(K)}} (X_i^{(K)})^p \geq x^{1+p} \,\middle|\, X_0 = x\epsilon\right) \mathbb{P}(B_{x,\epsilon})\right] \\
=\ &\liminf_{x\to\infty} \frac{1}{x^\alpha L(x)} \log\left[\mathbb{P}\left(\int_0^{T_1^{(K)}/x} \left(\frac{X_{\lfloor xs\rfloor}^{(K)}}{x}\right)^p ds \geq 1 \mid X_0 = x\epsilon\right) \cdot \mathbb{P}(B_{x,\epsilon})\right] \\
\geq\ &\liminf_{x\to\infty} \frac{1}{x^\alpha L(x)} \log\left[\mathbb{P}_{x\epsilon}\left(\int_0^{T_K^*} \left(\frac{X_{\lfloor xs\rfloor}^{(K)}}{x}\right)^p ds \geq 1,\ \inf_{s\in[0,T_K^*]} \frac{X_{\lfloor xs\rfloor}^{(K)}}{x} > 0\right) \cdot \mathbb{P}(B_{x,\epsilon})\right] \\
\geq\ &\liminf_{x\to\infty} \frac{1}{x^\alpha L(x)} \log\left[\mathbb{P}_\epsilon\left(\bar{K}_x^{(K)} \in A_\epsilon,\ \inf_{s\in[0,T_K^*]} \frac{X_{\lfloor xs\rfloor}^{(K)}}{x} > 0\right) \mathbb{P}(B_{x\epsilon})\right] \\
\geq\ &-\inf_{\xi\in A_\epsilon^\circ} I_{M_1'}^{(\mu_K)}(\xi) + \sqrt{\epsilon} \log \mathbb{P}\left(U_1^{(K)} > \sqrt{\epsilon}\right) \\
\geq\ &-\inf_{\xi\in \bar{A}_\epsilon} I_{M_1'}^{(\mu')}(\xi) + \sqrt{\epsilon} \log \mathbb{P}\left(U_1^{(K)} > \sqrt{\epsilon}\right) \\
\geq\ &\epsilon^\alpha + (|\mu_K|(p+1))^{\frac{\alpha}{p+1}} + \sqrt{\epsilon} \log \mathbb{P}\left(U_1^{(K)} > \sqrt{\epsilon}\right),
\end{aligned}
$$

where the third to last inequality is due to Theorem B.4.4 and the last one is due to (B.5.2). In addition, we used again (B.5.1) on the forth equiality.

Sending $\epsilon \downarrow 0$ for both the upper and lower bounds and then $K \to \infty$ for the lower bound proves the theorem. □

B.6 A Direct Upper Bound without LDP

In this section, we provide a self-contained proof of the upper bound in Theorem B.3.1 that does not rely on a Large Deviation Principle.

Proposition B.6.1. *Let $(U_i)_{i\geq 1}$ be a sequence of i.i.d. random variables with $\mathbb{E}[U_1] = \mu < 0$, and suppose that U_1 is semi-exponential with parameter $\alpha \in (0,1)$. Then,*

$$\limsup_{t\to\infty} \frac{\log \mathbb{P}(W_1(p) > t)}{\log \mathbb{P}\Big(U_1 > (|\mu|(p+1)t)^{\frac{1}{p+1}}\Big)} \leq 1.$$

Proof. Let $g(t) := (|\mu|(p+1)t)^{\frac{1}{p+1}}$ to simplify notation, and fix $\epsilon > 0$.

Consider the shifted random walk $K'_n := U_2 + \cdots + U_{n+1}$, for $n \geq 2$, and define the event

$$A_1 := \{K'_n \geq -n|\mu - \epsilon| - A, \ \forall n \geq 1\},$$

for some constant $A = A(\epsilon) < \infty$, such that, by the SLLN,

$$\mathbb{P}(A_1) \geq 1 - \epsilon. \tag{B.6.1}$$

Next, define

$$\gamma_\epsilon := \left(\frac{|\mu - \epsilon|}{|\mu|}\right)^{\frac{1}{p+1}}, \quad \text{and} \quad A_2(t) := \{U_1 > \gamma_\epsilon g(t) + A\}.$$

Observe that on $A_1 \cap A_2(t)$ we have

$$T_1 > \frac{\gamma_\epsilon g(t)}{|\mu - \epsilon|} + 1.$$

Hence, recalling $X_k = [X_{k-1} + U_k]^+$ and $X_1 = U_1$, we obtain

$$\begin{aligned}
W_1(p) &= \sum_{n=1}^{T_1} X_n^p \\
&= U_1^p + \sum_{n=2}^{T_1-1} (U_1 + K'_{n-1})^p \\
&\geq (\gamma_\epsilon g(t) + A)^p + \sum_{n=2}^{\lfloor \frac{\gamma_\epsilon g(t)}{|\mu-\epsilon|} \rfloor + 1} (\gamma_\epsilon g(t) - (n-1)|\mu - \epsilon|)^p \\
&> \sum_{n=0}^{\lfloor \frac{\gamma_\epsilon g(t)}{|\mu-\epsilon|} \rfloor} (\gamma_\epsilon g(t) - n|\mu - \epsilon|)^p \\
&> \int_0^{\frac{\gamma_\epsilon g(t)}{|\mu-\epsilon|}} (\gamma_\epsilon g(t) - |\mu - \epsilon| y)^p \, dy = t,
\end{aligned} \tag{B.6.2}$$

where the last inequality follows because the left Riemann sum of a decreasing function overestimates the integral.

Thus,

$$\mathbb{P}(W_1(p) > t) \geq \mathbb{P}(A_1 \cap A_2(t)) = \mathbb{P}(A_1)\, \mathbb{P}(A_2(t)), \tag{B.6.3}$$

since A_1 and $A_2(t)$ are independent.

It follows that

$$\limsup_{t\to\infty} \frac{\log \mathbb{P}(W_1(p) > t)}{\log \mathbb{P}(U_1 > \gamma_\epsilon g(t) + A)} \leq 1 + \limsup_{t\to\infty} \frac{\log \mathbb{P}(A_1)}{\log \mathbb{P}(A_2(t))} = 1. \tag{B.6.4}$$

Moreover, since semi-exponential random variables are long tailed, they are log-long-tailed as well, and we have:

$$\lim_{t\to\infty} \frac{\log \mathbb{P}(U_1 > \gamma_\epsilon g(t) + A)}{\log \mathbb{P}(U_1 > \gamma_\epsilon g(t))} = 1. \tag{B.6.5}$$

Moreover, semi-exponential random variables are log-regularly varying, i.e.,

$$\lim_{t\to\infty} \frac{\log \mathbb{P}(U_1 > \gamma_\epsilon g(t))}{\log \mathbb{P}(U_1 > g(t))} = \gamma_\epsilon^\alpha. \tag{B.6.6}$$

Hence,

$$\frac{\log \mathbb{P}(W_1(p) > t)}{\log \mathbb{P}(U_1 > g(t))} = \frac{\log \mathbb{P}(W_1(p) > t)}{\log \mathbb{P}(U_1 > \gamma_\epsilon g(t) + A)} \cdot \frac{\log \mathbb{P}(U_1 > \gamma_\epsilon g(t) + A)}{\log \mathbb{P}(U_1 > \gamma_\epsilon g(t))} \cdot \frac{\log \mathbb{P}(U_1 > \gamma_\epsilon g(t))}{\log \mathbb{P}(U_1 > g(t))}$$
$$\leq 1 \cdot 1 \cdot \gamma_\epsilon^\alpha = \gamma_\epsilon^\alpha.$$

Sending $\epsilon \to 0$ yields $\gamma_\epsilon \to 1$ and concludes the proof:

$$\limsup_{t\to\infty} \frac{\log \mathbb{P}(W_1(p) > t)}{\log \mathbb{P}(U_1 > g(t))} \leq 1.$$

□

Remark B.6.2. *While we defined A_1 to hold with high probability (via the SLLN), we did not explicitly use that $\mathbb{P}(A_1)$ is large — only that $\mathbb{P}(A_1) > 0$ and that the event A_1 is independent of $A_2(t)$. That hints in that we can induce something stronger using that argument. Actually, if U_1 is regularly varying with index $\alpha > 1$ (so that $\mathbb{E}[U_1] < \infty$)), we may strengthen this argument and obtain a estimate outside of the logarithmic scale.*

Proposition B.6.3. *Let $(U_i)_{i\geq 1}$ be i.i.d. with $\mathbb{E}[U_1] = \mu < \infty$. Assume that U_1 has a regularly varying tail with index $\alpha > 0$. Then*

$$\limsup_{t\to\infty} \frac{\mathbb{P}(W_1(p) > t)}{\mathbb{P}(U_1 > (|\mu|(p+1)t)^{1/(p+1)})} \;\geq\; 1.$$

Proof. Set $g(t) := \big(|\mu|(p+1)t\big)^{\frac{1}{p+1}}$ and $\gamma_\epsilon := \big(\frac{|\mu-\epsilon|}{|\mu|}\big)^{\frac{1}{p+1}}$. By (B.6.3),

$$\frac{\mathbb{P}(W_1(p) > t)}{\mathbb{P}(U_1 > g(t))} \;\geq\; \mathbb{P}(A_1)\, \frac{\mathbb{P}(U_1 > \gamma_\epsilon g(t) + A)}{\mathbb{P}(U_1 > g(t))}.$$

Because U_1 is long-tailed and regularly varying with index $\alpha > 0$,

$$\lim_{t\to\infty} \frac{\mathbb{P}(U_1 > \gamma_\epsilon g(t) + A)}{\mathbb{P}(U_1 > g(t))} = \lim_{t\to\infty} \frac{\mathbb{P}(U_1 > \gamma_\epsilon g(t) + A)}{\mathbb{P}(U_1 > \gamma_\epsilon g(t))} \lim_{t\to\infty} \frac{\mathbb{P}(U_1 > \gamma_\epsilon g(t))}{\mathbb{P}(U_1 > g(t))} = 1 \cdot \gamma_\epsilon^\alpha.$$

By the SLLN we have $\mathbb{P}(A_1) \geq 1 - \epsilon$ (see (B.6.1)). Hence

$$\limsup_{t\to\infty} \frac{\mathbb{P}(W_1(p) > t)}{\mathbb{P}(U_1 > g(t))} \;\geq\; (1-\epsilon)\,\gamma_\epsilon^\alpha.$$

Letting $\epsilon \downarrow 0$ yields the claim. □

Bibliography

[1] JMP Albin and Mattias Sundén. "On the asymptotic behaviour of Lévy processes, Part I: Subexponential and exponential processes". In: *Stochastic processes and their applications* 119.1 (2009), pp. 281–304.

[2] Søren Asmussen. *Applied probability and queues*. Springer, 2003.

[3] Mihail Bazhba. "Large Deviations for Semi-exponential Distributions: Theory and Applications". In: (2021).

[4] Mihail Bazhba et al. "Sample path large deviations for Lévy processes and random walks with Weibull increments". In: (2020).

[5] Mihail Bazhba et al. "Sample-path large deviations for unbounded additive functionals of the reflected random walk". In: *arXiv preprint arXiv:2003.14381* (2020).

[6] V Bentkus et al. *Limit theorems of probability theory*. Springer Science & Business Media, 2013.

[7] Simeon M Berman. "The supremum of a process with stationary independent and symmetric increments". In: *Stochastic processes and their applications* 23.2 (1986), pp. 281–290.

[8] Rabi N Bhattacharya and R Ranga Rao. *Normal approximation and asymptotic expansions*. SIAM, 2010.

[9] Patrick Billingsley. *Probability and Measure*. 3rd. Wiley Series in Probability and Statistics. Hoboken, NJ: Wiley, 1995. ISBN: 0471007102.

[10] Aleksandr Alekseevič Borovkov, Onno Johan Boxma, and Zbigniew Palmowski. "On the integral of the workload process of the single server queue". In: *Journal of applied probability* 40.1 (2003), pp. 200–225.

[11] Aleksandr Alekseevich Borovkov. *Compound renewal processes*. Vol. 184. Cambridge University Press, 2022.

[12] Michael Braverman. "Suprema and sojourn times of Lévy processes with exponential tails". In: *Stochastic processes and their applications* 68.2 (1997), pp. 265–283.

[13] Michael Braverman. "Remarks on suprema of Lévy processes with light tailes". In: *Statistics & probability letters* 43.1 (1999), pp. 41–48.

[14] Michael Braverman. "Suprema of compound Poisson processes with light tails". In: *Stochastic processes and their applications* 90.1 (2000), pp. 145–156.

[15] Michael Braverman. "On suprema of Lévy processes with light tails". In: *Stochastic processes and their applications* 120.4 (2010), pp. 541–573.

[16] Michael Braverman and Gennady Samorodnitsky. "Functionals of infinitely divisible stochastic processes with exponential tails". In: *Stochastic processes and their applications* 56.2 (1995), pp. 207–231.

[17] Daren BH Cline. "Convolution tails, product tails and domains of attraction". In: *Probability Theory and Related Fields* 72.4 (1986), pp. 529–557.

[18] Denis Denisov, Elena Perfilev, and Vitali Wachtel. "Tail asymptotics for the area under the excursion of a random walk with heavy-tailed increments". In: *Journal of Applied Probability* 58.1 (2021), pp. 217–237.

[19] Francois Dufresne and Hans U Gerber. "Risk theory for the compound Poisson process that is perturbed by diffusion". In: *Insurance: mathematics and economics* 10.1 (1991), pp. 51–59.

[20] Rick Durrett. *Probability: theory and examples*. Vol. 49. Cambridge university press, 2019.

[21] Paul Embrechts and Charles M Goldie. "On closure and factorization properties of subexponential and related distributions". In: *Journal of the Australian Mathematical Society* 29.2 (1980), pp. 243–256.

[22] Paul Embrechts and Charles M Goldie. "On convolution tails". In: *Stochastic Processes and their Applications* 13.3 (1982), pp. 263–278.

[23] Paul Embrechts, Charles M Goldie, and Noël Veraverbeke. "Subexponentiality and infinite divisibility". In: *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 49.3 (1979), pp. 335–347.

[24] Paul Embrechts, Claudia Klüppelberg, and Thomas Mikosch. *Modelling extremal events: for insurance and finance*. Vol. 33. Springer Science & Business Media, 2013.

[25] Paul D Feigin and Emmanuel Yashchin. "On a strong Tauberian result". In: *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 65.1 (1983), pp. 35–48.

[26] William Feller. *An introduction to probability theory and its applications, Volume 2*. Vol. 2. John Wiley & Sons, 1991.

[27] Sergey Foss, Dmitry Korshunov, Stan Zachary, et al. *An introduction to heavy-tailed and subexponential distributions*. Vol. 6. Springer, 2011.

[28] Pilar H García-Soidán. "Edgeworth expansions for triangular arrays". In: *Communications in Statistics-Theory and Methods* 27.3 (1998), pp. 705–722.

[29] Jan Grandell and Sven-Åeke Widaeus. "The Esscher approximation method". In: *Scandinavian Actuarial Journal* 1969.sup3 (1969), pp. 34–50.

[30] Geoffrey Grimmett and David Stirzaker. *Probability and random processes*. Oxford university press, 2020.

[31] Frank Hollander. *Large deviations*. Vol. 14. American Mathematical Soc., 2000.

[32] Kiyosi Itô. "On stochastic processes (I) Infinitely divisible laws of probability". In: *Japanese journal of mathematics: transactions and abstracts*. Vol. 18. The Mathematical Society of Japan. 1941, pp. 261–301.

[33] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*. springer, 2014.

[34] Andreas E Kyprianou. *Fluctuations of Lévy processes with applications: Introductory Lectures*. Springer Science & Business Media, 2014.

[35] Paul Lévy. "Sur certains processus stochastiques homogènes". In: *Compositio mathematica* 7 (1940), pp. 283–339.

[36] Peter David Miller. *Applied asymptotic analysis*. Vol. 75. American Mathematical Soc., 2006.

[37] Peter Mörters and Yuval Peres. *Brownian motion*. Vol. 30. Cambridge University Press, 2010.

[38] Jayakrishnan Nair, Adam Wierman, and Bert Zwart. *The fundamentals of heavy tails: Properties, emergence, and estimation*. Vol. 53. Cambridge University Press, 2022.

[39] John P Nolan. "Financial modeling with heavy-tailed stable distributions". In: *Wiley Interdisciplinary Reviews: Computational Statistics* 6.1 (2014), pp. 45–55.

[40] Elena Perfilev and Vitali Wachtel. "Local asymptotics for the area under the random walk excursion". In: *Advances in Applied Probability* 50.2 (2018), pp. 600–620.

[41] Sidney I Resnick. *Adventures in stochastic processes*. Springer Science & Business Media, 2013.

[42] Mikhail Sergeyevich Sgibnev. "Asymptotics of infinitely divisible distributions on R". In: *Siberian Mathematical Journal* 31.1 (1990), pp. 115–119.

[43] Ward Whitt. "Stochastic-process limits: an introduction to stochastic-process limits and their application to queues". In: *Space* 500 (2002), pp. 391–426.

[44] Eric Willekens. "On the supremum of an infinitely divisible process". In: *Stochastic processes and their applications* 26 (1987), pp. 173–175.