Heavy tails

Heavy tails: asymptotics, algorithms, applications

Bohan Chen

A catalogue record is available from the Eindhoven University of Technology Library
ISBN: 1234567890

Typeset by LaTeX

Printing: Printer
Cover: Designer

Heavy tails: asymptotics, algorithms, applications

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus prof.dr.ir. F.P.T. Baaijens, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op 11 december 2019 om 16:00 uur

door

Bohan Chen

geboren te Jiangsu, China

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter:	prof. dr. J.J. Lukkien
1^e promotor:	prof. dr. A.P. Zwart
copromotor:	dr. C.H. Rhee (Northwestern University)
leden:	prof. dr. ir. S.C. Borst
	prof. dr. ir. O.J. Boxma
	prof. dr. R.W. van der Hofstad
	prof. dr. C. Klüppelberg (Technische Universität München)
	prof. dr. M.R.H. Mandjes (Universiteit van Amsterdam)

Het onderzoek of ontwerp dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

Acknowledgements

This thesis would not have been possible without the help and encouragement of many people. I take this opportunity to express my gratitude to them.

First of all, I would like to thank my advisor Bert Zwart and co-advisor Chang-Han Rhee for all the support they have provided during the last four years. I am deeply grateful for their devoted guidance, their patience with my working style, and their genuine caring attitude. I could not have imagined having a better group of advisors and mentors for my Ph.D. study.

Besides my promoters, I would like to thank my thesis committee: Sem Borst, Onno Boxma, Remco van der Hofstad, Claudia Klüppelberg, and Michel Mandjes for their time, effort and insightful feedback on the thesis. Special thanks go to my former supervisor Claudia, who made this amazing journey possible by connecting me with Bert in the first place.

I gratefully acknowledge the financial support for my Ph.D. project provided by The Netherlands Organization for Scientific Research (NWO). Also, I want to thank my institute, the Centrum Wiskunde & Informatica (CWI), for making my relocation to Amsterdam silky smooth, and for the countless opportunities I have been offered there as a Ph.D. student: from the career-boosting courses to the ping-pong tables.

Thanks to my colleagues at CWI, the last four years there have always been so enjoyable. I definitely will miss the stimulating discussions and the exciting conversations about research, mathematics, and things in which we share a common interest.

I am happy to have met many friends during these years. For practical reasons, I cannot list the names of all of them. But I am sure that each one of them knows how much I valued those moments we have shared. They mean a lot to me.

To Xin, my wife, thank you for being there with me, as well as taking care of our daughter Amelia while I was writing the dissertation. Speaking of whom, to Amelia, thank you for giving me the extra motivation and for making my life colorful.

Last but not least, I would like to thank my father Xin Chen and my mother Xia Yuan for supporting me in every sense throughout writing this thesis and my life in general. I hope this booklet could be some joy to you.

Bohan Chen
Amsterdam, October 2019

Contents

Chapter 1

Introduction

Heavy-tailed distributions play an important role, especially in many human-made stochastic systems. For example, they accurately model inputs to computer and communication networks (see e.g. [60]), and they are an essential component of the description of many financial risk processes (see e.g. [54]). Much empirical evidence shows that the populations of cities, the intensities of earthquakes, and the sizes of power outages (see [33] and the references therein) follow power-law distributions. Intuitively, heavy tails occur in systems whose behavior mainly is determined by large values that occasionally shock the system.

Large deviations theory provides a rigorous mathematical foundation to analyze rare events such as those described in the previous paragraph. While classical large deviations theory has been successful in explaining rare events in light-tailed systems, the corresponding theory for heavy tails is much less developed. A well-known result is the so-called principle of a single big jump, which states that the asymptotic behavior of the sum of random variables is determined by a single large summand. Unfortunately, the principle of a single big jump cannot deal with all the essential applications, see e.g. Sections 2.4–2.6. To remedy this, Rhee, Blanchet, and Zwart present in [105] large deviations results for processes with independent increments, which go beyond the framework of the principle of a single big jump. As one of the main focus points of this thesis, we extend their theory to processes whose increments are dependent.

Another way to understand extreme behaviors in a heavy-tailed system is to perform numerical simulations. Many methods for performing simulations exist, each with their advantages and disadvantages. The most popular method is crude Monte Carlo sampling, thanks to its simplicity. However, as will be demonstrated below, it is not well suited to estimate events that have a small probability, due to the high number of runs that are required. A sampling method that tries to address this is importance sampling. Roughly speaking, importance sampling is a method that simulates the model under a different probability measure. The most technical issue is to find a new measure that performs well, in some yet to be defined sense. Often, large deviations theory can provide a basis for finding such a new measure. As the second goal of this thesis, we aim to provide simulation algorithms for rare events in heavy-tailed systems, based on importance sampling.

As mentioned above, heavy tails occur in applications that have a significant impact on society. Hence, as the last theme of this thesis, we apply the developed methods to specific problems that arise in finance, actuarial science, queueing theory, etc.

As the reader of this thesis might not be familiar with the mathematical results on which this thesis is based, we give an introduction to the required knowledge below. More specifically, we can roughly divide the required base knowledge into four topics: heavy tails, large deviations theory, rare-event simulation, and applications. References to more in-depth literature will be provided as well. We conclude this chapter with an overview of the contributions made in this thesis.

1.1 Heavy tails

In this section, we introduce the basic definition of heavy-tailed distributions. Moreover, we consider two common subclasses that often appear in stochastic modeling with heavy tails. Let X, X_i, $i \geq 1$, be independent non-negative random variables with common distribution function F. Let $\bar{F}$ denote the tail distribution of X, i.e., $\bar{F}(x) = \mathbf{P}(X > x)$.

Definition 1.1.1 (Definition 2.1.1 of [119]). F is heavy-tailed if, for all $\epsilon > 0$,

$$\mathbf{E}[e^{\epsilon X}] = \infty,$$

or equivalently, if for all $\epsilon > 0$, $e^{\epsilon x}\mathbf{P}(X > x) \to \infty$, as $x \to \infty$.

Basically, F is heavy-tailed if its tail decreases more slowly than exponentially. Otherwise, F is said to be light-tailed. Examples of heavy-tailed distributions are Pareto distributions, Cauchy distributions, lognormal distributions, Weibull distributions, etc. Next, we consider two important subclasses of heavy-tailed distributions. The following definition of subexponentiality can be found in [52].

Definition 1.1.2. F is subexponential if for *some* $n \geq 2$,

$$\mathbf{P}(X_1 + \cdots + X_n > x) \sim n\mathbf{P}(X_1 > x), \quad \text{as } x \to \infty.$$

It turns out that (see e.g. [60]) if F is subexponential then this relation holds for *every* $n \geq 2$. Another characterization of subexponential distributions (see e.g. [54]) is the following: F is subexponential if for every $n \geq 2$,

$$\mathbf{P}(X_1 + \cdots + X_n > x) \sim \mathbf{P}(\max(X_1, \ldots, X_n) > x), \quad \text{as } x \to \infty.$$

The interpretation of this characterization is that the only significant way in which the sum $X_1 + \cdots + X_n$ can exceed some large value x is that one of the individual random variables $X_1, \ldots, X_n$ also exceeds x. This is the principle of a single big jump which describes the most likely behaviour of sums of independent and identically distributed (i.i.d.) subexponential random variables, see e.g. page 44 of [60].

Subexponential distributions were introduced independently by [31] and [32]. In these references, the framework of subexponential distributions was used to derive asymptotic properties of branching processes, see also the textbook [11]. One of the first papers that recognized the usefulness of the class of subexponential distributions is [112]. In Chapter 5, we consider a rare-event simulation problem where the logarithm of the probability of interest is subexponential.

Another interesting subclass of heavy-tailed distributions is the class of regularly varying distributions, which can be considered as a generalization of the Pareto distribution.

Definition 1.1.3. A non-negative random varaible X is said to be regularly varying of index $-\alpha \leq 0$, if

$$\mathbf{P}(X > x) = x^{-\alpha} L(x),$$

with L being a slowly varying function, that is, for all $t > 0$, $L(tx)/L(x) \to 1$, as $x \to \infty$. Equivalently, we write $\bar{F} \in \mathcal{R}_{-\alpha}$.

Note that regularly varying distributions are all subexponential, see e.g. [56]. On the other hand, the reverse is not true since lognormal and Weibull distributions are subexponential but not regularly varying. We now list some useful properties of regularly varying distributions.

Lemma 1.1.1. *Let X be a regularly varying random varaible of index $-\alpha$, i.e., $\bar{F}(x) = L(x)x^{-\alpha}$ for some slowly varying function L. The following holds.*

1. *Let L be locally bounded in $[x_0, \infty)$ for some $x_0 \geq 0$. Then*

   (a) *for $\alpha < 1$,*

$$\int_{x_0}^{x} t^{-\alpha} L(t)dt \sim (1-\alpha)^{-1} x^{1-\alpha} L(x), \quad x \to \infty,$$

   (b) *for $\alpha > 1$,*

$$\int_{x}^{\infty} t^{-\alpha} L(t)dt \sim (\alpha-1)^{-1} x^{1-\alpha} L(x), \quad x \to \infty.$$

2. $\mathbf{E}[X^r] < \infty$ *if* $r < \alpha$, $\mathbf{E}[X^r] = \infty$ *if* $r > \alpha$.

3. *If Y is non-negative and independent of X such that $\mathbf{P}(Y > y) = L_1(y)y^{-\beta}$ for some slowly varying function L_1, then $X + Y$ is regularly varying of index $-\min\{\alpha, \beta\}$ and*

$$\mathbf{P}(X + Y > x) \sim \mathbf{P}(X > x) + \mathbf{P}(Y > x), \quad \textit{as } x \to \infty.$$

4. *If Y is non-negative and independent of X such that $\mathbf{E}[Y^{\alpha+\epsilon}] < \infty$ for some $\epsilon > 0$, then XY is regularly varying of index $-\alpha$ and*

$$\mathbf{P}(XY > x) \sim \mathbf{E}[Y^\alpha]\mathbf{P}(X > x), \quad \textit{as } x \to \infty.$$

Proof. The first statement is the so-called Karamata theorem, which can be found, for example, on page 567 of [54]. For the proof for the other statements, we refer to Lemma 2.1.8 of [119]. □

The concept of regular variation was introduced in [80]. In the scope of probability theory, the great potential of regular variation was first realized in [40] and [56]. Other key references are [15], [102, 103, 104], [54] and [119].

1.2 Large deviations theory

In this section, we give a short introduction to large deviations theory. Large deviations theory is the collection of mathematical theories and tools that aim to quantify and analyze rare events such as those mentioned in the very beginning of the chapter. As an illustration, consider first a sequence of i.i.d. random variables X_i, $i \geq 1$, having a normal distribution with zero mean and unit variance. Note that the sample mean $\hat{S}_n = 1/n \sum_{i=1}^n X_i$ is again normally distributed with zero mean and variance $1/n$. Thus, for any $\delta > 0$,

$$\lim_{n\to\infty} \mathbf{P}(|\hat{S}_n| \geq \delta) = 0, \tag{1.2.1}$$

and for any interval A,

$$\lim_{n\to\infty} \mathbf{P}(\sqrt{n}\hat{S}_n \in A) = \frac{1}{\sqrt{2\pi}} \int_A e^{-x^2/2} dx. \tag{1.2.2}$$

Moreover, we have that

$$\mathbf{P}(|\hat{S}_n| \geq \delta) = \frac{2}{\sqrt{2\pi}} \int_{\delta\sqrt{n}}^{\infty} e^{-x^2/2} dx.$$

From this expression it can be shown that

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbf{P}(|\hat{S}_n| \geq \delta) = -\frac{\delta^2}{2}. \tag{1.2.3}$$

By the weak law of large numbers and the central limit theorem, equations (1.2.1) and (1.2.2) hold for any i.i.d. sequence with finite second moment. On the other hand, equation (1.2.3) is an example of a large deviations statement: with probability of the order $e^{-\delta^2 n/2}$, the sample mean $\hat{S}_n$ deviates at least δ away from its "typical" value. One could ask whether equation (1.2.3) also holds more generally.

The answer to this question is given in Cramér's theorem, which was first derived in [39] for distributions possessing densities and then extended to general distributions in [30]. Cramér's theorem states that there exists a function I such that

$$-\inf_{x\in A^\circ} I(x) \leq \varliminf_{n\to\infty} \frac{\log \mathbf{P}(\hat{S}_n \in A)}{n} \leq \varlimsup_{n\to\infty} \frac{\log \mathbf{P}(\hat{S}_n \in A)}{n} \leq -\inf_{x\in A^-} I(x), \tag{1.2.4}$$

for all measurable A, where in the infima, the interior A° and the closure A^- of A are meant respectively. We say a sequence of probability measures ν_n satisfies a *large deviations principle* (LDP) with linear speed (cf. Chapter 1.2 of [42]), if (1.2.4) holds by replacing ν_n with $\mathbf{P}(\hat{S}_n \in \cdot)$. Roughly speaking, if $\mathbf{P}(\hat{S}_n \in \cdot)$, $n \geq 1$, satisfies a LDP with rate function I, then we have that

$$\mathbf{P}(\hat{S}_n \in A) \approx \exp\left\{-n \inf_{x\in A} I(x)\right\},$$

that is, the probability of interest decays exponentially in n with decay rate $\inf_{x\in A} I(x)$. Note that the approximation sign in the previous equation is only an informal expression, for its precise meaning we refer to (1.2.4).

As a natural extension of Cramér's theorem, a multivariate version exists, i.e., in the case where X_i, $i \geq 1$, is a sequence of i.i.d. random vectors in $\mathbb{R}^d$. Another generalization is the Gärtner-Ellis theorem (see [62] and [51]), where a LDP is derived for dependent sequences. Note that all large deviations results that have been discussed so far in this chapter are stated in terms of the empirical mean of a sequence of random variables. However, sometimes one is interested in obtaining information about the sample path of such a sequence. For example, the supremum of the random walk can be of interest, or one might want to know the probability of the (one-dimensional) random walk staying between two curves. These types of problems were studied by Mogulskii in [94]. Below we give a simpler version of Mogulskii's theorem. To begin with, let $\bar{S}_n = \{\bar{S}_n(t), t \in [0,1]\}$ be such that

$$\bar{S}_n(t) = \frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor} X_i, \quad 0 \leq t \leq 1, \tag{1.2.5}$$

let $L_\infty([0,1])$ be the space of measurable functions on $[0,1]$ equiped with the essential supremum norm, and let

$$\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}}\{\lambda x - \log \mathbf{E}[e^{\lambda X_1}]\}. \tag{1.2.6}$$

Theorem 1.2.1 (Mogulskii). *Assume* $\log \mathbf{E}[e^{\lambda X_1}] < \infty$ *for all* $\lambda \in \mathbb{R}$. *The sequence of probability measures* $\mathbf{P}(\bar{S}_n \in \cdot\,)$ *satisfies in* $L_\infty([0,1])$ *a LDP with rate function*

$$I(\phi) = \begin{cases} \int_0^1 \Lambda^*(\phi'(t))dt, & \textit{if } \phi \in \mathcal{AC},\ \phi(0) = 0, \\ \infty, & \textit{otherwise,} \end{cases} \tag{1.2.7}$$

where $\mathcal{AC}$ *is the set of absolutely continuous functions.*

Next we consider an example where Theorem 1.2.1 is applicable.

Example 1.2.1. Let $\bar{S}_n$ be as in (1.2.5). Assume $\log \mathbf{E}[e^{\lambda X_1}] < \infty$ for all $\lambda \in \mathbb{R}$. We are interested in the rare event probability of the sample path of a random walk ever crossing some level. To be precise, we are interested in $\mathbf{P}(\bar{S}_n \in A)$, where

$$A = \{\xi \in \mathbb{D}\colon \sup_{t\in[0,1]} \xi(t) \geq a\}.$$

Using Jensen's inequality, it is not difficult to show that ϕ with $\phi(t) = at$ for $t \in [0, 1]$ minimizes the rate function I as in (1.2.7) over all sample paths in A. Hence,

$$\lim_{n\to\infty} n^{-1} \log \mathbf{P}(\bar{S}_n \in A) = \Lambda^*(a),$$

where Λ^* is as in (1.2.6). Basically, the "most likely" sample path in the light-tailed setting is given by the linear function with drift a, since it solves the variational problem given by

$$\inf_{\phi\in A} \int_0^1 \Lambda^*(\phi'(t))dt \quad \text{subject to} \quad \phi \in \mathcal{AC},\ \phi(0) = 0.$$

In other words, conditional on the rare event, the limit of the scaled sample path is the same one that would be obtained if someone made a systematic change to the random walk. Thus, this phenomenon is referred to as the "conspiracy theory".

So far, all large deviations results that have been discussed in this chapter are stated in the light-tailed setting. Next we give a brief introduction to the heavy-tailed counterpart, which is also the focus of Chapter 4. To begin with, we consider the tail estimates $\mathbf{P}(\hat{S}_n > x)$ with $\hat{S}_n$ being the sample mean of i.i.d. heavy-tailed, but not necessarily regularly varying, random variables. The investigation of $\mathbf{P}(\hat{S}_n > x)$ was initiated in [95, 96], where the author studied the sequences x_n for which

$$\mathbf{P}(\hat{S}_n > x_n) = n\mathbf{P}(X_1 > x_n)(1 + o(1)), \quad \text{as } n \to \infty, \tag{1.2.8}$$

holds. For a detailed description of the large deviations regime we refer to e.g. [24, 43, 60]. If (1.2.8) is valid, the so-called *principle of a single big jump* is said to hold.

As mentioned above, the principle of a single big jump holds for a general class of heavy-tailed distributions. For simplicity, we focus here on the case where X_i, $i \geq 1$, is a sequence of i.i.d. random variables such that $X_1 \in \mathcal{R}_{-\alpha}$ is non-negative, and $\alpha > 1$. One could ask whether it is possible to derive a similar principle as in (1.2.4) in the regularly varying setup. The answer is yes. Indeed, using the result in [43], it is not difficult to show that $\mathbf{P}(\hat{S}_n \in \cdot)$ satisfies an LDP, however, with a sublinear speed. To be precise, for all measurable A, we have that

$$-\inf_{x\in A^\circ} I(x) \leq \varliminf_{n\to\infty} \frac{\log \mathbf{P}(\hat{S}_n \in A)}{\log n} \leq \varlimsup_{n\to\infty} \frac{\log \mathbf{P}(\hat{S}_n \in A)}{\log n} \leq -\inf_{x\in A^-} I(x), \tag{1.2.9}$$

where $I(x)$ equals $(\alpha-1)\mathbb{1}_{\{x>\mathbf{E}[X_1]\}}$ if $x \geq \mathbf{E}[X_1]$ and infinity (by the law of large numbers) otherwise.

Similar to the light-tailed setting, one could consider the functional version of (1.2.8). In fact, [75] establishes large deviations results for the sample path of heavy-tailed multi-dimensional random walks, which deal with rare events caused by a single big jump. However, the principle of a single big jump is not enough in the sense that many rare events—for concrete examples we refer to Chapters 2 and 3—cannot be caused by a single big jump. Instead, they are caused by multiple big jumps. Such an issue has been addressed in [105]. Let $\bar{S}_n$ be as in (1.2.5) such that $X_1 \in \mathcal{R}_{-\alpha}$ and $\alpha > 1$. Rhee, Blanchet, and Zwart develop in [105] asymptotic estimates of $\mathbf{P}(\bar{S}_n \in A)$ for a general set A in the sample path space, so that it is possible to study continuous functionals of $\bar{S}_n$ in a systematic manner. Note that the theory can especially deal with rare events that are caused by multiple big jumps.

We conclude this section with a comparison between the LDP under the light-tailed and the heavy-tailed settings. For this purpose, we introduce a simpler version of Theorem 4.1 of [105]. Let $\mathbb{D} = \mathbb{D}([0,1],\mathbb{R})$ denote the space of real-valued càdlàg functions on $[0,1]$, and let d_{J_1} denote the usual Skorokhod J_1 metric, where the precise definition can be found in Chapter 2 below. Finally, for $b \in \mathbb{R}$, define

$$\mathbb{D}_j^b = \left\{ \xi \in \mathbb{D} \colon \xi(t) = bt + \sum_{i=1}^{j} x_i \mathbb{1}_{[u_i,1]}(t), \text{ for } x_i > 0,\ 0 < u_1 < \cdots < u_j < 1 \right\}$$

and $\mathbb{D}_{<j}^b = \bigcup_{i=0}^{j-1} \mathbb{D}_i^b$.

Theorem 1.2.2 (Rhee, Blanchet, and Zwart [105])**.** *Let X_1 be a non-negative random variable such that $\mathbf{P}(X_1 > x) = x^{-\alpha}L(x)$ for some slowly varying L. Let $\mu = \mathbf{E}[X_1]$. If*

$$\min\{j \in \mathbb{N} \colon \mathbb{D}_j^\mu \cap A \neq \emptyset\} = j^* \tag{1.2.10}$$

and $d_{J_1}(A, \mathbb{D}_{<j^*}^\mu) > 0$, then there exists a measure C on $\mathbb{D}$ such that $(\mathbb{D}_{j^*}^\mu)^c$ is a C-null set and

$$C(A^\circ) \leq \varliminf_{n\to\infty} \frac{\mathbf{P}(\hat{S}_n \in A)}{(n\mathbf{P}(X_1 \geq n))^{j^*}} \leq \varlimsup_{n\to\infty} \frac{\mathbf{P}(\hat{S}_n \in A)}{(n\mathbf{P}(X_1 \geq n))^{j^*}} \leq C(A^-).$$

For the exact definition of C we refer to Chapter 2 below. Note that, in case the solution to the optimization problem in (1.2.10) equals 1, Theorem 1.2.2 gives us the principle of a single big jump. Moreover, analogs of Theorem 1.2.2

exist for heavy-tailed Lévy processes, see results in Section 3 of [105]. Next we revisit Example 1.2.1 under the heavy-tailed setup.

Example 1.2.1 (continued)**.** Recall we are interested in $\mathbf{P}(\bar{S}_n \in A)$ with

$$A = \{\xi \in \mathbb{D}\colon \sup_{t\in[0,1]} \xi(t) \geq a\}.$$

Let $X_1 \in \mathcal{R}_{-\alpha}$ for some $\alpha > 1$. W.l.o.g. we assume $a > \mathbf{E}[X_1] > 0$. Using the fact that any $\xi = \mu t + x_1 \mathbb{1}_{[u_1,1]}$ with $x_1 \geq a - \mathbf{E}[X_1]$ is contained in A, we conclude that j^* in (1.2.10) equals 1. Moreover, A is a C-continuous set, and hence

$$\lim_{n\to\infty} n^{\alpha-1}\mathbf{P}(\hat{S}_n \in A) = C(A),$$

where it has been shown, for example, in [53] that $C(A) = (a - \mathbf{E}X_1)^{-\alpha}$. In fact (see Corollary 4.1), given the rare event takes place, the scaled random walk behaves like a piecewise linear function with drift μ and one jump of size larger than $a - \mathbf{E}[X_1]$ almost surely since it leads to the rare event with a minimum number of jumps.

In Chapter 4, we go beyond the setting of [105] by considering stochastic processes whose increments are driven by some general Markov chain. To relate our problem with the existing theory of sample-path large deviations for stochastic processes, we first identify a sequence of regeneration times $\{r_n\}_{n\geq 1}$ (see [12]) and split the Markov chain into i.i.d. cycles. By aggregating the trajectory of $\bar{X}_n$ over regeneration cycles, we obtain a regenerative process with i.i.d. jump distributions and $\{r_n\}_{n\geq 1}$ as renewals. Under a set of classical assumptions originating from the works of [81] and [66], we adapt a large deviation change of measure argument and further establish a sample path large deviations result for the aggregated process. However, showing that the residual process is negligible is not straightforward, especially when the increments of $\bar{X}_n$ are in general dependent.

1.3 rare-event simulation

Not every mathematical model can be analyzed exactly. This is where numerical simulation is needed. Thanks to techniques such as parallel computing, GPU computing etc., simulation becomes an efficient way to obtain good approximations.

First let us introduce some basic concepts in rare-event simulation. Let A_n be a sequence of events on some probability space. In order to be *rare*, we should have $\lim_{n\to\infty} \mathbf{P}(A_n) = 0$ as $n \to \infty$. To obtain an unbiased estimator of $\mathbf{P}(A_n)$, one of the most intuitive methods is the so-called *crude Monte Carlo simulation.* That is, one samples N instances of the model. For each instance, one observes whether the event A_n occurs. Suppose we observe N_s occurrences, or, successes, then an unbiased estimator is given by N_s/N. Although the Monte Carlo method is very handy to execute, there is an issue with it in rare-event simulation. We consider the setup of [107]. For $\alpha \in (0,1)$, let z_α be such that $\mathbf{P}(-z_\alpha \leq \mathcal{N} \leq z_\alpha) = \alpha$, where $\mathcal{N}$ has a standard normal distribution. It can be shown that the sample variance is equal to $\mathbf{P}(A_n)(1-\mathbf{P}(A_n))/\sqrt{N}$, and hence, the radius of the asymptotic $\alpha\%$-confidence interval is given by

$$z_\alpha \sqrt{\frac{\mathbf{P}(A_n)(1-\mathbf{P}(A_n))}{N}},$$

by the central limit theorem. Moreover, the *relative error* of the estimator—which is defined as the ratio between the radius of the $\alpha\%$-confidence interval and the estimated probability—satisfies

$$z_\alpha \sqrt{\frac{1-\mathbf{P}(A_n)}{\mathbf{P}(A_n)N}} \sim \frac{z_\alpha}{\sqrt{\mathbf{P}(A_n)N}}, \quad \text{as } n \to \infty. \tag{1.3.1}$$

In practice, we often require that the relative error is smaller than some threshold, e.g., 5%. Thus, equation (1.3.1) implies that the number of simulation runs that are required to achieve a given accuracy is inversely proportional to $\mathbf{P}(A_n)$, and hence, increases drastically as the event of interest becomes rare.

One of many methods that try to address this issue is *importance sampling.* We explain this technique by generalizing the setting. Suppose that we are interested in computing $y = \mathbf{E}[\psi(Y)]$ for some $\mathbf{P}$-measurable function ψ. Importance sampling is basically generating Y from a modified probability measure $\tilde{\mathbf{P}}$ so that most of the samples are drawn from the part of the state space that contributes the most to y; this procedure is called *change of measure.* Especially, if we choose $\tilde{\mathbf{P}}$ such that $\mathbf{P}$ is absolutely continuous with respect to (w.r.t.) $\tilde{\mathbf{P}}$, then there exists a function L such that

$$y = \mathbf{E}[\psi(Y)] = \tilde{\mathbf{E}}[\psi(Y)L(Y)],$$

where $\tilde{\mathbf{E}}$ denotes the expectation operator w.r.t. $\tilde{\mathbf{P}}$. The output analysis is performed similarly to the crude Monte Carlo method, i.e., we generate N i.i.d. replicates of Y from $\tilde{\mathbf{P}}$ and we estimate y as the arithmetic mean of the replicates.

To describe the efficiency of a rare-event simulation algorithm, we adopt a widely applied criterion, which requires that the relative mean squared error of the associated estimator is appropriately controlled. To be more precise, suppose that we are interested in a sequence of rare events A_n, which becomes rare as n goes to infinity. For each $n \geq 1$, let L_n be an unbiased estimator of the rare-event probability $\mathbf{P}(A_n)$. L_n is said to be strongly efficient if

$$\lim_{n\to\infty} \mathbf{P}(A_n)^{-2}\mathbb{E}L_n^2 < \infty. \tag{1.3.2}$$

In particular, strong efficiency implies that the number of simulation runs required to estimate the target probability to a given relative accuracy is bounded w.r.t. n. For a more in-depth discussion on rare-event simulation, see [5].

In Chapter 2 we develop an efficient simulation algorithm for computing rare-event probabilities involving path functionals of heavy-tailed random walks and compound Poisson processes with regularly varying increments in a general large deviations regime. Our simulation estimator is based on an importance sampling strategy that hinges on the heavy-tailed sample path large deviations result in [105]. Next we provide a review of the theory and methods which are standard in rare-event simulation settings similar to those studied in this dissertation.

In the light-tailed context, large deviations theory can be used to design an importance sampling scheme. In fact, it is well known that an exponential change of measure—which is extracted from the proof of the asymptotic lower bound in large deviations analysis—can sometimes be efficient (for counterexamples see e.g. [63] and [64]). By connecting the design of efficient importance sampling estimators with a game theoretic formulation, [46], [47] and [48] provide the foundations for the use of large deviations theory in the construction and analysis of provably efficient rare-event simulation estimators. Moreover, a weakly efficient "universal" sampler has been proposed in [49] for a general class of hitting sets in arbitrary Jackson network topologies. Examples of additional recent papers are [25] and [114].

The setting of stochastic processes with heavy-tailed increments raises up additional challenges compared to its light-tailed counterpart discussed in the previous paragraph. These challenges were exposed in [9]. First of all, typically, the asymptotic conditional distribution of any particular increment given the rare event of interest converges to the underlying nominal distribution. Intuitively, if a rare event is caused by a large jump that may occur in a single "unlucky" increment out of many possible alternatives, then the chance that any specific increment is precisely the unlucky one is naturally small. So, any particular increment is likely to behave "normally". Therefore, in contrast to the light-tailed setting, there is no direct way in which one might attempt to bias a particular increment in order to stem the process towards the rare event of interest.

Moreover, as pointed out in [9], the asymptotic description of the most likely way in which a rare event may occur, for example due to the presence of a single big jump, does not lead to a valid change of measure for importance sampling because it is possible that several large jumps (or no large jump at all) might actually produce the event of interest under the nominal dynamics of the system. In other words, the natural biasing mechanism induced by directly approximating the zero-variance importance sampling distribution in the heavy-tailed setting assigns zero probability to events which are possible under the nominal dynamics leading to an ill-defined likelihood ratio.

The use of state-dependent importance sampling provides a way to deal with these difficulties. In [16], the authors explain how approximating Doob's h-transform can lead to a feasible change of measure which produces a strongly efficient importance sampling estimator in the setting of first passage time probabilities for one-dimensional random walks. A Lyapunov technique was introduced for the analysis of state-dependent importance sampling estimators. But the direct approximation of Doob's h-transform might be difficult to implement in higher dimensions both because of sampling implementation challenges and the evaluation of normalizing constants.

In the setting of one-dimensional compound sums of i.i.d. regularly varying random variables, [45] produced a state-dependent change of measure whose normalizing constant is straightforward to implement. Their idea can be described as follows: each increment is sampled by either the original measure or—with small probability, which is a design parameter—a different measure, which is essentially the original measure conditional on exhibiting a large jump. The advantage of the mixture samplers is that implementation challenges and the evaluation of normalizing constants can often be addressed by choosing a suitable set of parameters.

Under the setting where the time horizon is growing in large and moderate deviation schemes, Blanchet and Liu show in [21] how to use Lyapunov inequalities to address the parameter selection while enforcing a bounded relative error. A key step in the methodology is the construction of a suitable Lyapunov function (for an example of the technique in multidimensional settings, see [22]). Blanchet and Liu suggest using the type of fluid analysis which is prevalent in the large deviations literature of heavy-tailed stochastic processes (see e.g. [58] and [59]). However, the construction of the Lyapunov function and the verification of the Lyapunov inequality becomes highly non-trivial in settings involving multiple jumps and the presence of boundaries which are common in queueing systems, for an example of the types of complications which arise in queueing settings, see [17].

The idea of using mixtures suggested in [45] is also used in this thesis. For literature on simulation of heavy-tailed random walks from other perspectives, such as Markov chain Monte Carlo (MCMC) and cross-entropy, see e.g. [70], [76] and the references therein. While [45] treats a particular one-dimensional setting involving a rare event that is caused by a single big jump during a bounded time horizon, our setting is more general. Using the algorithm delevoped in Chapter 2, a wide range of rare events can be dealt with, which might be caused by multiple jumps during a growing time horizon in a large deviations scaling.

By considering the Markov chain defined by an i.i.d. sequence of iterated random functions, in Chapter 5 we develop an efficient simulation algorithm for the tail probability of the stationary distribution based on a state-dependent importance sampling scheme. While the results in Chapter 4 have regularly varying components, we investigate in Chapter 5 the case where the probability of interest is super heavy-tailed.

1.4 Applications

In this thesis we focus on rare events that appear in finance, actuarial science, and queueing systems. The first three applications will be considered in Chapters 2 and 3, and the algorithm developed in Chapter 2 serves as an efficient simulation method to quantify the probability of interest under the heavy-tailed setting. The last application, which can be found in Chapter 5, is to illustrate the case where there is a dependence structure in the increments of the underlying process.

1.4.1 Barrier option

In Chapter 2, we consider an application that arises in the context of financial mathematics; in particular we consider a down-in barrier option (see e.g. [111]). Simply put, the payoff of a barrier option depends on whether the price of the underlying asset crosses a predetermined *barrier* before maturity. The simplest barrier options are “knock-in” options, which are activated when the price of the underlying asset touches the barrier, and “knock-out” options, which are deactivated in that case. A down-in call has the same payoff as a European call if the price of the underlying asset remains above the barrier over the life of the option, but becomes worthless as soon as the price of the underlying asset falls below the barrier.

While the pricing of barrier options is analytically tractable under the classcial Black-Scholes model (see e.g. [108]), the calculation under a general Lévy market model becomes quite involved, for more details we refer to [97] and [38]. In this thesis, we focus on estimating the probability of exercising the barrier options under a large deviations regime, where the log-return of the underlying asset is regularly varying.

To be precise, let S_n, $n \geq 0$, denote the spot price of some underlying asset observed, say, at a daily basis. For $a > 0$ and $b > 0$, the probability of the event

$$A_n = \{S_n \geq bn, \min_{0\leq k\leq n} S_k \leq -an\},$$

can be interpreted as the chance of exercising a down-in barrier option. Intuitively, we need one upward jump and one downward jump in order to make the rare event happen. In Chapter 2, we derive the asymptotics of $\mathbf{P}(A_n)$ as $n \to \infty$.

1.4.2 Ruin probability under reinsurance policies

Another application that is used to test our algorithm is the finite-time ruin probability of an insurance company under three different types of reinsurance contracts. We first provide the probablistic framework. Let the *total claim amount process* of an insurance company be modelled by a compound Poisson process, denoted by $X(t)$, $t \in [0, \infty)$. Assume that $\mathbf{P}(X(1) > x)$ is regularly varying. Note that the tail asymptotics for the infinite-time ruin probability, which is given by

$$\mathbf{P}(u + ct - X(t) < 0 \quad \text{for some } t \in [0, \infty)), \quad \text{as } u \to \infty,$$

can be found in [55]. In this dissertation, we focus on the case of a finite horizon. To be precise, we want to estimate

$$\mathbf{P}(\sup_{t\in[0,1]} X(nt) - cnt - R(nt) > an),$$

for large values of n, where $R(t)$ denotes the *reinsured amount* at time t. We consider the following reinsurance forms. For a detailed description of each reinsurance form, we refer to [1].

Let Y_n, $n \geq 1$, be the claim size. We have

$$X(t) = \sum_{i=1}^{N(t)} Y_i.$$

The first example is the so-called (per risk) *excess of loss reinsurance*, which belongs to the category of non-proportional reinsurance forms. In this case, $R(t)$ is defined by

$$R(t) = \sum_{i=1}^{N(t)} (Y_i - b)_+,$$

for some pre-defined retention b. Basically, the reinsurer agrees to pay for each claim the excess over the retention. This reinsurance form is very popular in casualty and fire insurance, as it reduces the exposure of the ceding company in an effective way and has an intuitive and simple form. It is obvious that, for $a > b$, a single big jump is not enough to cause the rare event, and hence, one should expect that the principle of multiple big jumps takes place. Estimates for the infinite horizon case can be found in [8].

Another example is the *large claim reinsurance*. Consider the ordering of the claims $\{Y_i\}_{1\le i\le N(t)}$

$$Y_{1,N(t)} \ge Y_{2,N(t)} \ge \cdots \ge Y_{N(t),N(t)}.$$

In a large claim reinsurance contract the reinsurer agrees to cover the largest r claims, where $r \ge 1$ is a fixed number. To be precise,

$$R(t) = \sum_{i=1}^{r} Y_{i,N(t)},$$

where we make the convention that $Y_{i,N(t)} = 0$ for $i = N(t) + 1, \ldots, r$, in case $N(t) < r$.

Finally, we investigate a further variant of large claim reinsurance, which is called ECOMOR (Excédent du Coût Moyen Relatif). In an ECOMOR contract,

$$R(t) = \sum_{i=1}^{r} (Y_{i,N(t)} - Y_{r+1,N(t)})_+.$$

In particular, an ECOMOR contract constitutes an excess-of-loss treaty with a random retention, and the latter equals the $(r + 1)$st-largest claim. One should realize it can happen that the reinsured amount decreases, although the overall burden for the insurer has increased by the arrival of this additional claim. Such a feature can be undesirable sometimes in practice.

In Chapter 3, we will see that the last two reinsurance treaties fall under the framework of the principle of multiple big jumps. In fact, $r + 1$ big jumps are "required" to cause the ruin.

1.4.3 Stochastic fluid network

An application that arises in queueing theory is the single-class open stochastic fluid network with d nodes, each with a buffer of infinite capacity. Roughly speaking, the model assumes an exogenous fluid input arriving at each of the d nodes. At each node, the fluid is processed at a deterministic rate. After that, a proportion of the processed fluid is routed from each node to either one of the other nodes or out of the network. The object of interest is the d-dimensional buffer-content process. Now, let us give the mathematical formulation of the model. For a detailed introduction on the model, we refer to Section 14 of [117] and the references therein.

Let $J = (J^{(1)}, \ldots, J^{(d)})$ be the vector of exogenous input stochastic processes at the d nodes such that, in each coordinate, $J^{(i)}$ is a real-valued, non-decreasing and non-negative càdlàg function on $[0, \infty)$. Let $r = (r^{(1)}, \ldots, r^{(d)})$ be the vector of deterministic output rates at the d nodes, and let Q be the $d \times d$ routing matrix, that is, a proportion $Q_{ij} \geq 0$ of the fluid processed by the i-th node is immediately routed to the j-th node, while a proportion $1 - \sum_{j=1}^{d} Q_{ij}$ is routed out of the network. We assume $\lim_{n\to\infty} Q^n = 0$, so that all input eventually leaves the network. The dynamics of the model are expressed formally by the so-called Skorokhod map (for details see e.g. [109], [110], [71]), that is defined in terms of a pair of processes (Z, Y) satisfying a stochastic differential equation that we shall describe now. Let $X(t) = J(t) - (I - Q)^T rt$, and let $Z^{(i)}(t)$ denote the workload of the i-th station at time t. For given $Z^{(i)}(0)$, we have that

$$dZ(t) = dX(t) + (I - Q)^T dY(t),$$

where $Y(\cdot)$ is the *potential buffer-content process* that encodes the minimum amount of pushing required to keep $Z(\cdot)$ non-negative. In Chapter 2, we assume that J is a superposition of d independent compound Poisson processes with non-negative regularly varying jump size distributions. Moreover, the probability of the amount of fluids in a subset of the system crossing a high level is estimated. Although some particular cases exist that allow for an explicit analysis (see e.g. Section 13 in [41]), it is hard to come up with exact results for the distribution of the workload process in general. Hence, implementing our algorithm in such a context is particularly interesting.

1.4.4 Stochastic perpetuities

Consider the $\mathbb{R}$-valued Markov chain defined by the recursion

$$Z_{n+1} = R_{n+1} Z_n + M_{n+1}, \tag{1.4.1}$$

where (R_n, M_n), $n \geq 1$, is sequence of i.i.d. $\mathbb{R}^2$-valued random vectors, independent of the initial random variable Z_0. It is well known (see e.g. Chapter 2 of [26]) that, under some regularity condition, the Markov chain given by (1.4.1) has a unique stationary distribution, which is the same distribution of

$$Z = \sum_{j=0}^{\infty} M_{j+1} \prod_{k=1}^{j} R_k.$$

If one interprets $-\log R_n$ as the interest rate at time n, then Z is the present value of a bond that generates M_n unit of money at each time point n. Thus, Z is called stochastic perpetuity. In Chapter 5, we estimate the tail probability of Z under a super heavy-tailed setting.

1.5 Contributions

This section highlights the contributions that are made in this thesis.

In Chapter 2, we propose a class of strongly efficient rare-event simulation estimators for random walks and compound Poisson processes with regularly varying increments in a general large deviations regime. The proposed estimators are straightforward to implement and can be used to systematically evaluate the probability of a wide range of rare events with bounded relative error. They are "universal" in the sense that a single importance sampling scheme applies to a very general class of rare events that arise in heavy-tailed systems. In particular, our estimators can deal with rare events that are caused by multiple big jumps (therefore, beyond the usual principle of a single big jump) as well as multidimensional processes such as the buffer content process of a queueing network. We illustrate the versatility of our approach with several applications that arise in the context of mathematical finance, actuarial science, and queueing theory.

In Chapter 3, we investigate the probability that an insurance portfolio gets ruined within a finite time period under the assumption that the r largest claims are (partly) reinsured. We show that for regularly varying claim sizes the probability of ruin after reinsurance is also regularly varying in terms of the initial capital, and derive an explicit asymptotic expression for the latter. We establish this result by leveraging the findings in [105]. Our results allow, on the asymptotic level, for an explicit comparison between two well-known large-claim reinsurance contracts, namely LCR and ECOMOR. We finally assess the accuracy of the resulting approximations using state-of-the-art rare-event simulation techniques.

In Chapter 4, we extend Theorem 1.2.2 from random walks to a class of Markov additive processes. We present sample path large deviations results for $\bar{X}_n = \{1/n \sum_{i=0}^{\lfloor nt \rfloor - 1} X_n, t \in [0,1]\}$. We show that under a set of classical assumptions as in [81] and [66], $\mathbf{P}(\bar{X}_n \in A)$ is regularly varying as $n \to \infty$, when A satisfies certain topological properties that can be verified easily. We illustrate the usefulness of our results in an application in barrier option pricing.

In Chapter 5, we consider the stationary solution Z of the Markov chain that is defined by an i.i.d. sequence of iterated random functions f_n. We are interested in estimating the probability of the event $\{Z > x\}$ when x is large, and develop a state-dependent importance sampling estimator under a set of assumptions on f_n that, for large x, make the event $\{Z > x\}$ take place by a single big jump. Under natural conditions, we show that our estimator is strongly efficient. Special attention will be given to a class of perpetuities with heavy tails.

Chapter 2

Rare-Event Simulation for Multiple Jump Events

2.1 Introduction

In this chapter, we develop a strongly efficient importance sampling scheme for computing rare-event probabilities involving path functionals of heavy-tailed random walks and compound Poisson processes in a general large deviations regime.

We focus on stochastic processes with regularly varying increments. Our simulation algorithm is straightforward to implement and can be used to estimate the likelihood of a wide range of rare events with bounded relative error. In particular, such a single sampling scheme applies to a very general class of rare events whose occurrence is caused by one or several components in the system which exhibit extreme behavior, while the rest of the system is operating in "normal" circumstances (therefore, beyond the usual principle of a single big jump). In particular, our results apply to a large class of continuous functionals of multiple random walks and compound Poisson processes.

Our estimators are based on importance sampling, a Monte Carlo technique which consists in biasing the nominal distribution of the underlying process in order to induce the rare event of interest. Our goal is to find biasing techniques leading to estimators which have bounded coefficient of variation (see equation (1.3.2) above) uniformly as the probability of the event of interest tends to zero in a suitable large deviations regime.

The construction of our sampling scheme is driven by recently developed large deviations results in [105] for regularly varying Lévy processes. Specifically, let $X(t)$, $t \geq 0$ be a one-dimensional compensated compound Poisson process with unit arrival rate and a positive jump W that is regularly varying at infinity (see Definition 2.2.1 below). Define $\bar{X}_n = \{\bar{X}_n(t)\}_{t\in[0,1]}$, with $\bar{X}_n(t) = X(nt)/n$. For a measurable set $A \subseteq \mathbb{D}$ satisfying a specific topological property, the large deviations results derived in [105] establish that $\mathbf{P}(\bar{X}_n \in A) = \Theta\big(\big(n\mathbf{P}(W \geq n)\big)^{l^*}\big)$, where precise details can be found in Section 2.2 below. In practice, exact estimates are often demanded. Hence, we design a sampling scheme for rare events that take the form $\mathbf{P}(\bar{X}_n \in A)$. We illustrate our approach with several applications that arise in mathematical finance, actuarial science, and queueing theory.

Recall that we are interested in estimating $\mathbf{P}(\bar{X}_n \in A)$. The concept behind our sampling scheme can be described as follows. Based on the large deviations results derived in [105], we construct first an auxiliary set B^γ that is closely related to the optimization problem given by (2.2.2) below. Then, given a fixed mixture probability parameter $w \in (0, 1)$, we generate the sample path of $\bar{X}_n$ under the nominal measure. And, with probability $1 - w$ we generate the sample path of $\bar{X}_n$ under the measure $\mathbf{Q}^\gamma(\,\cdot\,) \triangleq \mathbf{P}(\,\cdot\,|\bar{X}_n \in B^\gamma)$. Finally, as a consequence of applying the importance sampling technique, we scale our samplers with a suitable likelihood ratio given as in (2.3.3) below. It should be noted that the set A can be as general as in the setting of [105]. Therefore, our methodological contribution in this chapter addresses precisely those types of challenges mentioned in Section 1.3, such as multiple jumps, time scales of order $\mathcal{O}(n)$, avoiding the evaluation of normalizing constants, and by-passing the verification of Lyapunov inequalities. The advantages of our sampling scheme are that the new estimators are strongly efficient and straightforward to implement. Moreover, they are "universal" in the sense that a single importance sampling scheme applies to a very general class of rare events involving multiple jumps that arise in heavy-tailed systems. As a final remark, it should be mentioned that constructing the auxiliary set B^γ requires choosing a set of suitable parameters γ whose existence is guaranteed by the topological property we impose on A. Hence, one of the main challenges is to select the set of parameters specifically for each application.

Our mathematical contributions in this chapter can be summarized as follows.

- We propose a simulation algorithm for estimating the rare-event probability of $\bar{X}_n \in A$, together with a sampling scheme for $\bar{X}_n \in \cdot$ given $\bar{X}_n \in B^\gamma$, which is based on a rejection sampling with an unconditional acceptance probability bounded away from zero as $n \to \infty$.

- By showing the existence of the parameter γ, we prove the strong efficiency of our sampling scheme under a very general setting (see Assumption 2.3.2 below).
- We showcase the versatility of the algorithm by illustrating the implementation of the proposed sampling scheme to the rare-events that arise in finance, actuarial science, and queueing theory.
- Especially in the application to queueing networks (see Section 2.6 below), we show that the tail index of the rare-event probability—which usually exhibits a complex boundary behavior due to the nonlinear nature of the associated Skorokhod mapping—can be determined by solving s knapsack problem with nonlinear constraints.

The rest of the chapter is organized as follows. Section 2.2 deals with basic background and notations required to state our contributions. Section 2.3 introduces our estimators and describes the main result. Applications and numerical implementations are discussed in Sections 2.4–2.6. All the proofs of results presented in this chapter are given in Section 2.7.

2.2 Notations and preliminaries

This section is split into two parts. The first discusses general notions that will be employed in this chapter. The second reviews recently developed results involving large deviations for regularly varying Lévy processes and random walks.

2.2.1 Notations

We start with a summary of notations that will be employed in this chapter. Let $\mathbb{Z}_+$ denote the set of non-negative integers, and let $\mathbb{R}_+$ denote the set of non-negative real numbers. Let A° and A^- denote the interior and the closure of A, respectively. Let $(\mathbb{D}_{[0,1],\mathbb{R}}, d)$ be the metric space of real-valued càdlàg functions on $[0,1]$, denoted by $\mathbb{D} = \mathbb{D}_{[0,1],\mathbb{R}}$, equipped with the Skorokhod J_1 metric on $\mathbb{D}$ that is defined by $d(x,y) = \inf_{\lambda\in\Lambda} ||\lambda - \mathrm{id}||_\infty \vee ||x \circ \lambda - y||_\infty, \quad x, y \in \mathbb{D}$, where id denotes the identity mapping, $||\cdot||_\infty$ denotes the uniform metric, i.e., $\|x\|_\infty \triangleq \sup_{t\in[0,1]} |x(t)|$, and Λ denotes the set of all strictly increasing, continuous bijections from $[0,1]$ to itself. Let $\mathbb{D}^k$ denote the k-fold product space of $\mathbb{D}$. Let $\mathbb{D}^k_\uparrow$ denote the subset of functions in $\mathbb{D}^k$ that are non-negative and nondecreasing in each coordinate. When it comes to the tail indices of a regularly varying distribution, we use β (or β_i in the multidimensional case) for the right tail and α for the left tail. Let $\mathbb{D}_l$ denote the subspace of $\mathbb{D}$ consisting of non-decreasing step functions vanishing at time zero with l jumps, and let $\mathbb{D}_{<l^*}$ denote the subspace of $\mathbb{D}$ consisting of non-decreasing step functions vanishing at 0 with at most $l-1$ jumps, i.e. $\mathbb{D}_{<l^*} = \bigcup_{l\leq l^*-1} \mathbb{D}_l$. Define

$$\mathbb{D}_{<(l_1^*,\ldots,l_d^*)} \triangleq \bigcup_{(l_1,\ldots,l_d)\in I_{<(l_1^*,\ldots,l_d^*)}} \prod_{i=1}^d \mathbb{D}_{l_i},$$

where

$$I_{<(l_1^*,\ldots,l_d^*)} \triangleq \left\{(l_1,\ldots,l_d)\in\mathbb{Z}_+^d \setminus \{(l_1^*,\ldots,l_d^*)\} \colon \mathcal{I}(l_1,\ldots,l_d) \leq \mathcal{I}(l_1^*,\ldots,l_d^*)\right\},$$

and $\mathcal{I}(l_1,\ldots,l_d) \triangleq (\beta_1-1)l_1+\ldots+(\beta_d-1)l_d$. Define a partial order $\prec$ on $\mathbb{Z}_+^d$ such that

$$(l_1,\ldots,l_d) \prec (m_1,\ldots,m_d) \text{ iff } \mathbb{C}_{(l_1,\ldots,l_d)} \subsetneq \mathbb{C}_{(m_1,\ldots,m_d)},$$

where $\mathbb{C}_{(l_1,\ldots,l_d)} \triangleq \bigcup_{i=1}^d \mathbb{D}^{i-1}\times\mathbb{D}_{<l_i}\times\mathbb{D}^{d-i}$. Define

$$J_{(j_1,\ldots,j_d)} \triangleq \{(l_1,\ldots,l_d)\in\mathbb{Z}_+^d\setminus I_{<(j_1,\ldots,j_d)} \colon (m_1,\ldots,m_d)\prec(l_1,\ldots,l_d) \text{ implies } (m_1,\ldots,m_d)\in I_{<(j_1,\ldots,j_d)}\}.$$

To get familiar with the notation, an illustration of $I_{<(l_1*,\ldots,l_d*)}$, $J_{(l_1*,\ldots,l_d*)}$, and the partial order $\prec$ is given in Figure 2.1. Let $\mathbb{D}_{l_-;l_+}$ denote the subspace of the Skorokhod space consisting of step functions vanishing at the origin with exactly l_- downward jumps and l_+ upward jumps, and define

$$\mathbb{D}_{<l_-^*;l_+^*} \triangleq \bigcup_{(l_-,l_+)\in I_{<l_-^*;l_+^*}} \mathbb{D}_{l_-;l_+},$$

where

$$I_{<l_-^*;l_+^*} \triangleq \{(l_-,l_+)\in\mathbb{Z}_+^2\setminus\{(l_-^*,l_+^*)\}\colon (\alpha-1)l_-+(\beta-1)l_+ \leq (\alpha-1)l_-^*+(\beta-1)l_+^*\}. \tag{2.2.1}$$

Given non-negative sequences of real numbers x_n and y_n, we write $x_n = \mathcal{O}(y_n)$, $x_n = o(y_n)$ and $x_n = \Theta(y_n)$, if $\limsup_{n\to\infty} x_n/y_n < \infty$, $\lim_{n\to\infty} x_n/y_n =$

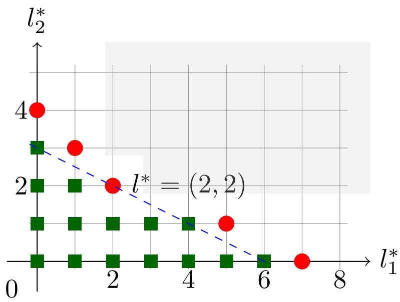

Figure 2.1: An example of important notations introduced in Section 2.2.1. For $(\beta_1 - 1)/(\beta_2 - 1) = 2$ and $(l_1^*, l_2^*) = (2, 2)$, we mark the elements in $I_{<(l_1*, l_2*)}$ and $J_{(l_1^*, l_2^*)}$ with squares and circles, respectively. Moreover, the shaded area contains all those points (l_1, l_2) such that $(l_1^*, l_2^*) \prec (l_1, l_2)$.

0 and $0 < \liminf_{n\to\infty} x_n/y_n \le \limsup_{n\to\infty} x_n/y_n < \infty$, respectively. Given two $\mathbb{R}$-valued functions f and g, we write $f \propto g$, if there exists $c \in \mathbb{R}$ such that $f = cg$. For $x = (x_1, \ldots, x_k)$, $y = (y_1, \ldots, y_k) \in \mathbb{R}^k$, we write $x \le y$, if $x_i \le y_i$, for all $i \in \{1, \ldots, k\}$. Let the cardinality of S be denoted by $|S|$ or $\#S$. Finally, let $\mathcal{C}(S, k)$ and $\mathcal{P}(S, k)$ denote the set of all k-combinations and k-permutations of a set S, respectively. Note that $|C(S, k)| = \binom{|S|}{k}$ and $|\mathcal{P}(S, k)| = k!|C(S, k)|$.

2.2.2 Preliminaries

As we will see, the simulation algorithm that we propose in this chapter is constructed based on the asymptotic behavior of rare-event probabilities, therefore we review some recently developed large deviations results for scaled Lévy processes with heavy-tailed Lévy measures, introduced in [105]. To begin with, we give the definition of regular variation for general random variables.

Definition 2.2.1. A random variable X is said to be regularly varying at infinity and minus infinity with index β, if $\mathbf{P}(X \ge x)$ and $\mathbf{P}(X \le -x)$ are regularly varying with index β, respectively.

Now, let X be a Lévy process with Lévy measure ν, where ν is spectrally positive and regularly varying (at infinity) with index $-\beta < -1$. Let $\bar{X}_n \triangleq \{X(nt)/n\}_{t\in[0,1]}$ denote the associated scaled process. Let ν_β^l denote the restriction of the l-fold product measures of ν_β to $\{x \in \mathbb{R}_+^l : x_1 \ge x_2 \ge \ldots \ge x_l\}$, where $\nu_\beta(x, \infty) \triangleq x^{-\beta}$. For $l \ge 1$, define a (Borel) measure $C_l(\,\cdot\,) \triangleq \mathbf{E}[\nu_\beta^l\{y \in$

$(0,\infty)^l\colon \sum_{i=1}^l y_i \mathbb{1}_{[U_i,1]} \in \cdot\,\}]$, where U_i, $i \geq 1$ are i.i.d. uniformly distributed on $[0,1]$. Note that C_l is concentrated on $\mathbb{D}_l$, i.e., $C_l(\mathbb{D}_l) = 1$. Moreover, we make the convention that C_0 is the Dirac measure concentrated on the zero function. The following result is useful in designing an efficient algorithm for rare events involving one-dimensional scaled processes. Throughout the rest of this chapter, all measurable sets are understood to be Borel measurable.

Result 2.2.1 (Theorem 3.1 of [105])**.** *Suppose that A is a measurable set. If A is bounded away from $\mathbb{D}_{<l^*}$, i.e., $d(A, \mathbb{D}_{<l^*}) > 0$, where*

$$l^* \triangleq \min\{l \in \mathbb{Z}_+ \colon \mathbb{D}_l \cap A \neq \emptyset\} < \infty,$$

then we have that

$$C_{l^*}(A^\circ) \leq \liminf_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in A)}{(n\nu[n,\infty))^{l^*}} \leq \limsup_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in A)}{(n\nu[n,\infty))^{l^*}} \leq C_{l^*}(A^-).$$

As one can see in Section 2.6 below, some applications can be interpreted as sample-path rare events in a multidimensional setting. Therefore, it is particularly interesting to consider large deviations results for multidimensional processes. Let $X^{(1)}, \dots, X^{(d)}$ be independent centered one-dimensional Lévy processes with spectrally positive Lévy measures $\nu_1(\cdot), \dots, \nu_d(\cdot)$, respectively, where each ν_i is regularly varying with index $-\beta_i < -1$ at infinity. Moreover, for the finite product of metric spaces, we use the maximum metric; i.e., we use $d_{\mathbb{S}_1\times\cdots\times\mathbb{S}_d}((x_1,\dots,x_d),(y_1,\dots,y_d)) \triangleq \max_{i=1,\dots,d} d_{\mathbb{S}_i}(x_i,y_i)$ for the product $\mathbb{S}_1\times\cdots\times\mathbb{S}_d$ of metric spaces $(\mathbb{S}_i, d_{\mathbb{S}_i})$. Finally, for $(l_1,\dots,l_d) \in \mathbb{Z}_+^d$, we define $C_{l_1}\times\cdots\times C_{l_d}(\cdot)$ (which is concentrated on $\prod_{i=1}^d \mathbb{D}_{l_i}$) as the product measure of

$$C_{l_i}(\,\cdot\,) \triangleq \mathbf{E}\left[\nu_{\beta_i}^{l_i}\{y \in (0,\infty)^{l_i} \colon \sum_{j=1}^{l_i} y_j \mathbb{1}_{[U_j,1]} \in \cdot\,\}\right].$$

Result 2.2.2 states a large deviations result for d-dimensional process $\bar{X}_n(t) \triangleq (X^{(1)}(nt)/n, \dots, X^{(d)}(nt)/n)$ for $t \in [0,1]$.

Result 2.2.2 (Theorem 3.6 of [105])**.** *Suppose that A is measurable. If A is bounded away from $\mathbb{D}_{<(l_1^*,\dots,l_d^*)}$, where*

$$(l_1^*,\dots,l_d^*) = \operatorname*{arg\,min}_{(l_1,\dots,l_d)\in\mathbb{Z}_+^d,\ \prod_{i=1}^d \mathbb{D}_{l_i}\cap A\neq\emptyset} \mathcal{I}(l_1,\dots,l_d), \tag{2.2.2}$$

and $\mathcal{I}(l_1,\ldots,l_d) = (\beta_1-1)l_1 + \ldots + (\beta_d-1)l_d$*, then we have that*

$$\begin{aligned} C_{l_1^*}\times\cdots\times C_{l_d^*}(A^\circ) &\le \liminf_{n\to\infty}\frac{\mathbf{P}(\bar{X}_n\in A)}{\prod_{i=1}^d\big(n\nu_i[n,\infty)\big)^{l_i^*}} \\ &\le \limsup_{n\to\infty}\frac{\mathbf{P}(\bar{X}_n\in A)}{\prod_{i=1}^d\big(n\nu_i[n,\infty)\big)^{l_i^*}} \le C_{l_1^*}\times\cdots\times C_{l_d^*}(A^-). \end{aligned}$$

Note that the assumption that A is bounded away from $\mathbb{D}_{<(l_1^*,\ldots,l_d^*)}$ guarantees the uniqueness of $(l_1^*,\ldots,l_d^*)$. Finally, we conclude this section with an extension of Result 2.2.2, which will be useful in constructing an efficient simulation algorithm for heavy-tailed random walks. Let S_k, $k\ge 0$, be a random walk, set $\bar{S}_n(t) = S_{\lfloor nt\rfloor}/n$, $t\ge 0$, and define $\bar{S}_n = \{\bar{S}_n(t),\ t\in[0,1]\}$. Let ν_β^l be as defined above. Similarly, let ν_α^m denote the restriction of m-fold product measures of ν_α to $\{x\in\mathbb{R}_+^m : x_1\ge x_2\ge\ldots\ge x_m\}$, where $\nu_\alpha(x,\infty)\triangleq x^{-\alpha}$. Let $C_{0,0}(\cdot)\triangleq\delta_{\mathbf{0}}(\cdot)$ be the Dirac measure concentrated on the zero function. For each $(l_-,l_+)\in\mathbb{Z}_+^2\setminus\{(0,0)\}$, define a measure (which is concentrated on $\mathbb{D}_{l_-;l_+}$) $C_{l_-;l_+}(\cdot)\triangleq\mathbf{E}[\nu_\alpha^{l_-}\times\nu_\beta^{l_+}\{(x,y)\in(0,\infty)^{l_-}\times(0,\infty)^{l_+} : \sum_{i=1}^{l_+}y_i1_{[V_i,1]} - \sum_{i=1}^{l_-}x_i1_{[U_i,1]}\in\cdot\}]$, where U_i's and V_i's are i.i.d. uniform on $[0,1]$.

Result 2.2.3. *Suppose that* $\mathbf{P}(S_1\le -x)$ *is regularly varying with index* $-\alpha$ *and* $\mathbf{P}(S_1\ge x)$ *is regularly varying with index* $-\beta$*. Let* A *be a measurable set bounded away from* $\mathbb{D}_{<l_-^*;l_+^*}$*, where*

$$(l_-^*,l_+^*) = \operatorname*{arg\,min}_{(l_-,l_+)\in\mathbb{Z}_+^2,\,\mathbb{D}_{l_-;l_+}\cap A\neq\emptyset}(\alpha-1)l_- + (\beta-1)l_+. \tag{2.2.3}$$

Then

$$\begin{aligned} C_{l_-^*;l_+^*}(A^\circ) &\le \liminf_{n\to\infty}\frac{\mathbf{P}(\bar{S}_n\in A)}{(n\mathbf{P}(S_1\le -n))^{l_-^*}(n\mathbf{P}(S_1\ge n)))^{l_+^*}} \\ &\le \limsup_{n\to\infty}\frac{\mathbf{P}(\bar{S}_n\in A)}{(n\mathbf{P}(S_1\le -n))^{l_-^*}(n\mathbf{P}(S_1\ge n)))^{l_+^*}} \le C_{l_-^*;l_+^*}(A^-). \end{aligned}$$

2.3 Main results

In this section we present our main results. Although the large deviations results reviewed in Section 2.2 are stated for Lévy processes, we focus on compensated compound Poisson processes for simulation purposes. Let X denote a d-dimensional compensated compound Poisson process, and recall that $\bar{X}_n$ is the scaled process with $\bar{X}_n(t) = X(nt)/n$, $t \in [0,1]$. For a measurable set $A \in \mathbb{D}^d$, we are interested in estimating the probability of the event $A_n \triangleq \{\bar{X}_n \in A\}$, when n is large. Note that, in view of the law of large numbers, one can expect that $\mathbf{P}(\bar{X}_n \in A) \to 0$ for A's that are bounded away from the zero function, and hence, A_n's are rare events for large n's. In Section 2.3.1, we first illustrate the idea of our algorithm in the special case for $d = 1$, where the notations are simpler. In Section 2.3.2 we extend this result to general d.

2.3.1 The one-dimensional case

Let $\{X(t)\}_{t\geq 0}$ be a one-dimensional compensated compound Poisson process with i.i.d. jump sizes $\{W(k)\}_{k\geq 1}$. That is, $X(t) = \sum_{k=1}^{N(t)} W(k) - \lambda t \mathbf{E} W(1)$, where $\{N(t)\}_{t\geq 0}$ is a Poisson process with arrival rate λ, and let

$$\bar{X}_n \triangleq \{X(nt)/n,\, t \in [0,1]\}$$

denote the associated scaled process. Moreover, let $\mathbf{P}(W(1) > x)$ be regularly varying of index $-\beta < -1$. The following assumption is essential for analyzing the asymptotic behavior of the rare-event probability, and hence, deriving the strong efficiency of our estimator.

Assumption 2.3.1. Let A be a measurable set in $\mathbb{D}$. We assume that A is bounded away from $\mathbb{D}_{<l^*}$, where $l^* = \min\{l \in \mathbb{Z}_+ : \mathbb{D}_l \cap A \neq \emptyset\}$ denotes the minimal number of upward jumps of a step function in A. Moreover, assume that $C_{l^*}(A^\circ) > 0$.

Remark 2.1. As one can see in Sections 2.4, 2.5, and 2.6, one of the typical settings that arises in applications is that the set A can be written as a finite combination of unions and intersections of $F_1^{-1}(A_1), \dots, F_m^{-1}(A_m)$, where each $F_i : \mathbb{D} \to \mathbb{S}_i$ is a continuous function, and all sets A_i are subsets of a general topological space $\mathbb{S}_i$. If we denote this operation of taking unions and intersections by Ψ (i.e., $A = \Psi(F_1^{-1}(A_1), \dots, F_m^{-1}(A_m))$), then it holds that $\Psi\left(F_1^{-1}(A_1^\circ), \dots, F_m^{-1}(A_m^\circ)\right) \subseteq A^\circ \subseteq A \subseteq A^- \subseteq \Psi\left(F_1^{-1}(A_1^-), \dots, F_m^{-1}(A_m^-)\right)$. Hence, $C_{l^*}(A^\circ) > 0$ holds if $\hat{T}_{l^*}^{-1}\left(\Psi\left(F_1^{-1}(A_1^\circ), \dots, F_m^{-1}(A_m^\circ)\right)\right)$ has positive Lebesgue measure, where $\hat{T}_j : \hat{S}_j \to \mathbb{D}_j$ is defined by $\hat{T}_j(x,u) \triangleq \sum_{i=1}^{j} x_i \mathbb{1}_{[u_i,1]}$ for $j \in \mathbb{Z}_+$, and

$$\hat{S}_j \triangleq \left\{(x,u) \in \mathbb{R}_+^j \times [0,1]^j : x_1 \geq \cdots \geq x_j, 0, 1, u_1, \dots, u_j \text{ are distinct}\right\}.$$

Analogously, one can derive a sufficient condition for $C_{l_1^*} \times \cdots \times C_{l_d^*}(A^\circ) > 0$ (see Assumption 2.3.2 below).

Remark 2.2. There are several examples that satisfy Assumption 2.3.1. For instance, considering $A = \{\xi \in \mathbb{D}_{[0,1]} : \xi(1) \geq a\}$ corresponds to estimating the rare-event probability $\mathbf{P}(X(n) \geq an)$. Another application that fits into this framework is the ruin probability of an insurance company, where the reinsurance policy is taken into account. For details of this application we refer to Section 2.4. Finally, for one of many examples of A in the multidimensional setting we refer to Section 2.6, where the workload in a queueing network is considered.

We design a simulation algorithm that estimates the probability of $A_n \triangleq \{\bar{X}_n \in A\}$ efficiently, based on an importance sampling strategy. To construct an importance distribution, we introduce a constant $\gamma > 0$ and define $B_n^\gamma \triangleq \{\bar{X}_n \in B^\gamma\}$, where $B^\gamma \triangleq \{\xi \colon \#\{t \,|\, \xi(t) - \xi(t^-) > \gamma\} \geq l^*\}$. In the construction of our rare-event simulation algorithm, we will take advantage of the fact that one can always choose γ so that B_n^γ is sufficiently "close" to A_n. The specific choice of γ will be further discussed later in Section 2.4, 2.5, and 2.6 for concrete examples. Let $\mathbf{Q}_\gamma(\,\cdot\,) \triangleq \mathbf{P}(\,\cdot\,|B_n^\gamma)$ denote the conditional distribution given $\bar{X}_n \in B^\gamma$. One should notice that $d\mathbf{Q}_\gamma/d\mathbf{P} = \mathbf{P}(B_n^\gamma)^{-1}\mathbb{1}_{B_n^\gamma}$. Moreover, by the Fubini-Tonelli theorem, a closed-form expression for $\mathbf{P}(B_n^\gamma)$ is given by

$$\mathbf{P}\left(B_n^\gamma\right) = 1 - \exp\left\{ -\lambda n \mathbf{P}(W(1) > n\gamma)\right\} \sum_{j=0}^{l^*-1} \frac{(\lambda n)^j}{j!} \mathbf{P}(W(1) > n\gamma)^j. \quad (2.3.1)$$

From (2.3.1) one should recognize that B_n^γ can be interpreted as the event of a Poisson distributed random variable with rate $\lambda n \mathbf{P}(W(1) > \gamma n)$ crossing the level l^*. Now, let $w \in (0,1)$ be arbitrary but fixed. We propose an importance distribution $\mathbf{Q}_{\gamma,w}$ that is absolutely continuous w.r.t. $\mathbf{P}$ and is given by

$$\mathbf{Q}_{\gamma,w}(\,\cdot\,) \triangleq w\mathbf{P}(\,\cdot\,) + (1-w)\mathbf{Q}_\gamma(\,\cdot\,). \quad (2.3.2)$$

We give here an algorithm for generating the sample path of $\bar{X}_n$ under the probability measure $\mathbf{Q}_\gamma(\,\cdot\,)$. Since $\{\bar{X}_n \in B^\gamma\} \subseteq \{N(n) \geq l^*\}$, we observe that $\mathbf{Q}_\gamma(\bar{X}_n \in \,\cdot\,) = \mathbf{P}(B_n^\gamma)^{-1}\mathbf{P}(\bar{X}_n \in \,\cdot\,, B_n^\gamma) = \sum_{m=l^*}^\infty h_m \, \mathbf{P}(\bar{X}_n \in \,\cdot \,|\, B_n^\gamma, N(n) = m)$, where $h_m = h_m(n) \triangleq \mathbf{P}(B_n^\gamma, N(n) = m)/\mathbf{P}(B_n^\gamma)$ satisfies $\sum_{m \geq l^*} h_m = 1$. Hence, it remains to discuss sampling from $\mathbf{P}(\bar{X}_n \in \,\cdot \,|\, B_n^\gamma, N(n) = m)$. It turns out that we can use the acceptance-rejection method, where drawing from the proposal distribution can be achieved as follows: first sample $\{b_k\}_{k \leq l^*}$ uniformly from $\mathcal{C}\left(\{1,\ldots,m\}, l^*\right)$; then sample each $W(b_k)$, $k \leq l^*$, conditional on $W(b_k) > n\gamma$; finally sample $W(m')$, $m' \leq m$, $m' \notin \{b_k\}_{k \leq l^*}$, under the nominal measure. Note that the target density $f_{\text{target};m}$, defined by

$$
\begin{aligned}
& f_{\text{target};m}(w_1, \ldots, w_m)\, dw_1 \cdots dw_m \\
& \triangleq \mathbf{P}\left(W(1) \in w_1 + dw_1, \ldots, W(m) \in w_m + dw_m \mid B_n^\gamma, N(n) = m\right),
\end{aligned}
$$

can be bounded by $M_m f_{\text{proposal};m}(w_1, \ldots, w_m)$, where

$$
\begin{aligned}
& f_{\text{target};m}(w_1, \ldots, w_m) \\
& \propto \frac{1}{\mathbf{P}(B_n^\gamma | N(n) = m)} \prod_{j=1}^{m} \frac{d}{dw_j} \mathbf{P}\left(W(j) \leq w_j\right) \mathbb{1}_{B_n^\gamma}(w_1, \ldots, w_m),
\end{aligned}
$$

$$
\begin{aligned}
& f_{\text{proposal};m}(w_1, \ldots, w_m) \\
& = \frac{1}{\binom{m}{l^*} \mathbf{P}(W(1) > n\gamma)^{l^*}} \prod_{j=1}^{m} \frac{d}{dw_j} \mathbf{P}(W(j) \leq w_j) \sum_{\substack{(b_1, \ldots, b_{l^*}) \in \\ \mathcal{C}(\{1, \ldots, m\}, l^*)}} \mathbb{1}_{\{W(b_k) > n\gamma, \forall k \leq l^*\}},
\end{aligned}
$$

and hence, $M_m = M_m(n) \triangleq \binom{m}{l^*} \mathbf{P}(W(1) > n\gamma)^{l^*} \mathbf{P}(B_n^\gamma | N(n) = m)^{-1}$. Now, it is natural to accept $(W(1), \ldots, W(m))$ with probability

$$
a(W(1), \ldots, W(m)) = \binom{\#\{i \in \{1, \ldots, m\} \colon W(i) > n\gamma\}}{l^*}^{-1}.
$$

Finally, we are able to formulate the pseudocode for generating $\bar{X}_n$ under $\mathbf{Q}_\gamma$ in Algorithm 1. Moreover, we show in Proposition 2.3.1 that the expected running time of Algorithm 1 is uniformly bounded from above w.r.t. n.

Proposition 2.3.1. *Let $T_{alg1}(n)$ denote the expected running time of Algorithm 1. Under the condition that $W(1)$ is regularly varying of index $-\beta < -1$, we have that $T_{alg1}(n) = \sum_{m \geq l^*} h_m(n) M_m(n)$ is uniformly bounded from above w.r.t. n, i.e. $\max_{n \geq 0} T_{alg1}(n) < \infty$.*

In view of the observations we made so far, we propose an estimator Z_n for $\mathbf{P}(A_n)$ that is given by

$$
Z_n = \mathbb{1}_{A_n} \frac{d\mathbf{P}}{d\mathbf{Q}_{\gamma,w}} = \frac{\mathbb{1}_{A_n}}{w + (1-w)(\mathbf{P}(B_n^\gamma))^{-1} \mathbb{1}_{B_n^\gamma}}. \tag{2.3.3}
$$

Intuitively, an importance sampling technique is used to get more samples from the interesting region, by sampling from a distribution that overweighs the

Algorithm 1 Generating the sample path of $\bar{X}_n$ under $\mathbf{Q}_\gamma$

sample $m \sim h_m$ $\triangleright$ $m = m'$ with probability $h_{m'} = \mathbf{P}(N(n) = m' \,|\, B_n^\gamma)$
$R \leftarrow$ **true**
while $R =$ **true** **do**
  sample $\{b_k\}_{k \le l^*} \sim \mathrm{unif}\left(\mathcal{C}\left(\{1, \ldots, m\}, k\right)\right)$ $\triangleright$ uniform distribution on $\mathcal{C}\left(\{1, \ldots, m\}, k\right)$
  for $i \in \{b_k\}_{k \le l^*}$ **do**
    sample $W(i) \sim W(1) \,\big|\, W(1) > n\gamma$
  for $i \notin \{b_k\}_{k \le l^*}$ **do**
    sample $W(i) \sim W(1)$
  $c \leftarrow \#\{j \in \{1, \ldots, m\} \colon W(j) > n\gamma\}$; $a \leftarrow \binom{c}{l^*}^{-1}$; sample $u \sim \mathrm{uniform}[0, 1]$; $R \leftarrow$ **true**
  if $u < a$ **then**
    $R \leftarrow$ **false**
  return $\bar{X}_n$

important region. Based on this, the choice of B_n^γ can be "justified", since B_n^γ is mimicking the asymptotic behavior of the probability of interest. However, as one can see in the proof of strong efficiency (see Theorem 2.3.2 below), we should analyze the second moment of our estimator to avoid "backfire", yielding an estimator with larger or even infinite variance. It turns out that this intuition can be made rigorous by applying Result 2.2.1. We end this section with a theorem regarding the strong efficiency of our estimator.

Theorem 2.3.1. *Under Assumption 2.3.1, there exists a* $\gamma > 0$ *such that the estimator constructed in* (2.3.3) *is strongly efficient for estimating* $\mathbf{P}(A_n)$.

Remark 2.3. Note that, under Assumption 1, there exists $r > 0$ such that $d(A, \mathbb{D}_{<l^*}) > r$. One sufficient way to make Z_n in (2.3.3) strongly efficient is to choose γ such that $\mathbb{Z} \not\ni \lceil r/\gamma \rceil \ge l^* + 1$. Sometimes, finding r can be application specific, though generally r is the smallest size a big jump needs to take to make the rare event happen, and physical intuition—which can be obtained from solving the large-deviations problem—is helpful in making an educated guess on r. For more details about finding r as well as choosing γ we refer to Sections 2.4–2.6.

2.3.2 Extension to general d

In this section we extend the results in Section 2.3.1 to the case with general d. To be precise, let $X \triangleq \left(X^{(1)}, \ldots, X^{(d)}\right)$ be a superposition of d independent compensated compound Poisson processes with upward jumps, where $\{N^{(i)}(t)\}$ is a Poisson process with arrival rate λ_i, and $X^{(i)}(t) = \sum_{k=1}^{N^{(i)}(t)} W^{(i)}(k) - \lambda_i t\mathbf{E}W^{(i)}(1)$. Moreover, let $\mathbf{P}(X^{(i)}(1) > x)$ be regularly varying of index $-\beta_i < -1$ at infinity. Finally, let $\bar{X}_n$ denote the corresponding scaled process. As we can see in Result 2.2.2, the large deviations results for $o\mathbf{P}(\bar{X}_n \in A)$ depend heavily on the value of $\mathcal{I}(l_1^*, \ldots, l_d^*)$, where $(l_1^*, \ldots, l_d^*)$ is as defined in (2.2.2). However, for $c \in \mathbb{R}$, the grid $(l_1, \ldots, l_d) \in \mathbb{Z}_+^d$ satisfying $\mathcal{I}(l_1, \ldots, l_d) = c$ is not unique in general. Therefore assuming A being bounded away from $\prod_{i=1}^d \mathbb{D}_{<l_i}$ is not sufficient for our purposes. The following assumption, which is slightly different from Assumption 2.3.1 corresponds to the extension of Result 2.2.1 to Result 2.2.2.

Assumption 2.3.2. Let A be a measurable set. Assume that A is bounded away from $\mathbb{D}_{<(l_1^*,\ldots,l_d^*)}$, where $(l_1^*, \ldots, l_d^*)$ is the unique solution of the minimization problem given by (2.2.2). Moreover, assume that $C_{l_1^*} \times \cdots \times C_{l_d^*}(A^\circ) > 0$.

If the solution to (2.2.2) is not unique, we may partition A. As in Section 2.3.1, we focus now on constructing the auxiliary set B^γ for the importance distribution. Define $A_n \triangleq \left\{\bar{X}_n \in A\right\}$ and $B_n^\gamma \triangleq \{\bar{X}_n \in B^\gamma\}$. As one can see in the proof of Theorem 2.3.2, controlling the probability of $A_n \cap (B_n^\gamma)^c$ should be taken into account in choosing the auxiliary set B^γ. In the one-dimensional case, letting B^γ mimic the optimal path leading to the rare event makes us capable of controlling the relative error of our estimator. By "mimicking the optimal path" we mean that the minimal number of jumps l^* that are needed for $\mathbb{D}_{l^*} \cap A \neq \emptyset$ is used as parameter in the construction of B^γ. However, the same strategy would fail in the multidimensional case, since the rare event can be reached through other feasible but not necessarily optimal paths. Thus, we require a more complicated construction of B^γ.

Definition 2.3.1. Let A be a measurable set in $\mathbb{D}^d$, and let $(l_1^*, \ldots, l_d^*)$ denote the unique solution to (2.2.2). Let $\gamma \in \mathbb{R}^d$ with $\gamma_i > 0$ for all $i \in \{1, \ldots, d\}$, and define

$$B^\gamma \triangleq \bigcup_{(l_1,\ldots,l_d)\in J_{(l_1^*,\ldots,l_d^*)}} B^{\gamma;l}, \tag{2.3.4}$$

where $B^{\gamma;l}$ is the set of càdlàg functions on $\mathbb{R}^d$ that have at least l_i jumps with size larger than γ_i in its i-th coordinate, i.e., $B^{\gamma;l} \triangleq \{(\xi^{(1)}, \ldots, \xi^{(d)}) \in \mathbb{D}^d \colon \#\{t \colon \xi^{(i)}(t) - \xi^{(i)}(t^-) > \gamma_i\} \geq l_i,\ \forall i \in \{1, \ldots, d\}\}$.

Remark 2.4. Note that the cardinality of $J_{(l_1^*,\ldots,l_d^*)}$ is finite. To design a strongly efficient simulation algorithm for estimating $\mathbf{P}(A_n)$, we will take advantage of an important property of $J_{(l_1^*,\ldots,l_d^*)}$. That is, for all $\xi \in A$ with A being bounded away from $\mathbb{D}_{<(l_1^*,\ldots,l_d^*)}$, there exists $(l_1,\ldots,l_d) \in J_{(l_1^*,\ldots,l_d^*)}$, such that the path of ξ in its i-th coordinate is bounded away from $\mathbb{D}_{<l_i^*}$, for every $i \in \{1,\ldots,d\}$.

Let $\mathbf{Q}_\gamma(\,\cdot\,) \triangleq \mathbf{P}(\,\cdot\,|B_n^\gamma)$ and let $\mathbf{Q}_{\gamma,w}$ be as defined in (2.3.2), following the same strategy as in Section 2.3.1 we propose an estimator that takes the same form as in (2.3.3). Before turning to the efficiency analysis of our estimator, we summarize the findings above in Algorithm 2.

Algorithm 2 Efficient sampling of $\mathbf{P}\left(\bar{X}_n \in A\right)$

```
1: sample u ~ uniform[0,1]                          ▷ uniform distribution on [0,1]
2: sample X̄_n ~ P(X̄_n ∈ · | X̄_n ∈ B^γ)
3: if u < w then
4:     sample X̄_n ~ P(X̄_n ∈ ·)
5: if X̄_n ∈ A then
6:     L ← [w + (1 − w)𝟙_{B_n^γ} / P(B_n^γ)]^{−1}
7: else
8:     L ← 0
       return L
```

In order to complete our algorithm, we need to discuss the computation of $\mathbf{P}(B_n^\gamma)$, as well as the strategy of sampling from the conditional distribution $\mathbf{Q}_\gamma(\,\cdot\,)$. Since B^γ constructed in Definition 2.3.1 is the union of $B^{\gamma;l}$ with $l = (l_1,\ldots,l_d) \in J_{(l_1^*,\ldots,l_d^*)}$, by the inclusion-exclusion principle, it is sufficient to discuss computing the probability of sets of the form $\bigcap_{(l_1,\ldots,l_d)\in I} B^{\gamma;l}$, where I is a finite collection of elements in $\mathbb{Z}_+^d$. It turns out that the probability of such a set can be computed similarly as in Section 2.3.1. Based on this observation, we give the following proposition.

Proposition 2.3.2. *The probability of B_n^γ is equal to $\sum_{k=1}^{|J_{(l_1^*,\ldots,l_d^*)}|}(-1)^{k-1}c_k$, where*

$$c_k = \sum_{\substack{|I|=k \\ I\subseteq J_{(l_1^*,\ldots,l_d^*)}}} \prod_{i=1}^{d} \left(1 - \exp\{-\lambda_i n \mathbf{P}(W^{(i)}(1) > n\gamma_i)\} \sum_{j=0}^{\hat{l}_{i;I}-1} \frac{(\lambda_i n)^j}{j!} \mathbf{P}(W^{(i)}(1) > n\gamma_i)^j\right),$$

and $\hat{l}_{i;I} \triangleq \max_{(l_1,\ldots,l_d)\in I} l_i$.

Remark 2.5. It should be mentioned that the complexity of computing $\mathbf{P}(B_n^\gamma)$ can be reduced significantly in the case, where, for example, one is able to take a smaller (in the sense of cardinality) set than $J_{(l_1^*,\ldots,l_d^*)}$ (see e.g. Corollary 2.3.1, Sections 2.5 and 2.6 below).

As in Section 2.3.1, we now discuss generating the sample path of $\bar{X}_n$ under $\mathbf{Q}_\gamma$ in the next step. To begin with, we need the following lemma, which shows that B^γ can be decomposed into finitely many disjoint sets.

Lemma 2.3.1. *Let*

$$B^{\gamma;l}(i,j) \triangleq \left\{\xi \in \mathbb{D}^d \colon \#\{t \,|\, \xi^{(i)}(t) - \xi^{(i)}(t^-) > \gamma_i\} \geq (l(j))_i\right\}.$$

Let the elements in $J_{(l_1^*,\ldots,l_d^*)}$, denoted by $l(1),\ldots,l(|J_{(l_1^*,\ldots,l_d^*)}|)$, be ordered such that $(l(1))_d \leq \cdots \leq (l(|J_{(l_1^*,\ldots,l_d^*)}|))_d$. Define

$$\Delta B^{\gamma;l}(i,j) \triangleq B^{\gamma;l}(i,j) \setminus \left(\bigcup_{m=1}^{j-1} B^{\gamma;l}(i,m)\right), \quad i \in \{1,\ldots,d-1\}. \tag{2.3.5}$$

Then, we have that

$$B^\gamma = \bigcup_{m_1=1}^{|J_{(l_1^*,\ldots,l_d^*)}|} \bigcup_{m_2=1}^{m_1} \cdots \bigcup_{m_{d-1}=1}^{m_{d-2}} \left(\left(\bigcap_{i=1}^{d-1} \Delta B^{\gamma;l}(i,m_i)\right) \cap B^{\gamma;l}(d,1)\right).$$

Lemma 2.3.1 shows that B^γ can be decomposed into finitely many disjoint sets. This implies that

$$\mathbf{Q}_\gamma(\bar{X}_n \in \cdot\,) = \sum_{m_1=1}^{|J_{(l_1^*,\ldots,l_d^*)}|} \sum_{m_2=1}^{m_1} \cdots \sum_{m_{d-1}=1}^{m_{d-2}} h_{1;m_1,\ldots,m_{d-1}} \mathbf{P}(\bar{X}_n \in \cdot \,|\, \bar{X}_n \in B^\gamma(m_1,\ldots,m_{d-1})),$$

where

$$B^\gamma(m_1,\ldots,m_{d-1}) \triangleq \left(\bigcap_{i=1}^{d-1} \Delta B^{\gamma;l}(i,m_i)\right) \cap B^{\gamma;l}(d,1),$$

and $h_{1;m_1,\ldots,m_{d-1}} \triangleq \mathbf{P}(\bar{X}_n \in B^\gamma(m_1,\ldots,m_{d-1}))\mathbf{P}(\bar{X}_n \in B^\gamma)^{-1}$ satisfying

$$\sum_{m_1=1}^{|J_{(l_1^*,\ldots,l_d^*)}|} \sum_{m_2=1}^{m_1=1} \cdots \sum_{m_{d-1}}^{m_{d-2}} h_{1;m_1,\ldots,m_{d-1}} = 1.$$

Hence, it remains to design a sampling scheme for generating the sample path of $\bar{X}_n$ under $\mathbf{P}(\,\cdot\,|\,\bar{X}_n \in B^\gamma(m_1,\ldots,m_{d-1}))$ (for details about generating multi-dimensional discrete random numbers, see e.g. [73]). Due to the special structure of $B^\gamma(m_1,\ldots,m_{d-1})$, we are able to generate $\bar{X}_n^{(1)},\ldots,\bar{X}_n^{(d)}$ independently under $\mathbf{P}(\,\cdot\,|\,\bar{X}_n \in B^\gamma(m_1,\ldots,m_{d-1}))$. To see this, first note that sampling $\bar{X}_n^{(d)}$ is trivial due to the discussion in Section 2.3.1. Define $\check{l}(m_i;i) \triangleq \min_{\xi\in\Delta B^{\gamma;l}(i,m_i)} \#\{t\colon \xi(t)-\xi(t^-) > \gamma_i\}$, and

$$\hat{l}(m_i;i) \triangleq \max_{\xi\in\Delta B^{\gamma;l}(i,m_i)} \#\{t\colon \xi(t)-\xi(t^-) > \gamma_i\},$$

for $i \in \{1,\ldots,d-1\}$. By (2.3.5), we have that

$$\begin{aligned}
&\mathbf{P}(\bar{X}_n^{(i)} \in \cdot\,|\,\bar{X}_n \in B^\gamma(m_1,\ldots,m_{d-1}))\\
&= \sum_{q_i=\check{l}(m_i;i)}^{\infty} h_{2;q_i}\, \mathbf{P}(\bar{X}_n^{(i)} \in \,\cdot\,|\,\Delta B^{\gamma;l}(i,m_i), N^{(i)}(n)=q_i),
\end{aligned}$$

where

$$h_{2;q_i} \triangleq \mathbf{P}(\Delta B^{\gamma;l}(i,m_i), N^{(i)}(n)=q_i)/\mathbf{P}(\Delta B^{\gamma;l}(i,m_i))$$

satisfies $\sum_{q_i\geq \check{l}(m_i;i)} h_{2;q_i} = 1$. Note that

$$\begin{aligned}
&\mathbf{P}(\Delta B^{\gamma;l}(i,m_i), N^{(i)}(n)=q_i)\\
&= \frac{(\lambda n)^{q_i}}{e^{\lambda_i n} q_i!}\left(\sum_{i=\check{l}(m_i;i)}^{\hat{l}(m_i;i)\wedge q_i} \binom{q_i}{i} \mathbf{P}(W^{(i)}(1) > n\gamma)^i \mathbf{P}(W^{(i)}(1) \leq n\gamma)^{q_i-i}\right).
\end{aligned}$$

Therefore, it suffices to consider sampling $\bar{X}_n^{(i)}$ under

$$\mathbf{P}(\,\cdot\,|\,\Delta B^{\gamma;l}(i,m_i), N^{(i)}(n)=q_i).$$

Again, we can proceed using a similar approach as in Section 2.3.1: sample $\{b_k\}_{k\leq l}$ uniformly from $\mathcal{C}\left(\{1,\ldots,q_i\}, \check{l}(m_i;i)\right)$; sample each $W^{(i)}(b_k)$, $k \leq q_i$, conditional on $W^{(i)}(b_k) > n\gamma_i$; sample $W^{(i)}(q_i')$, $q_i' \leq q_i$, $q_i' \notin \{b_k\}_{k\leq l^*}$, under the nominal measure; accept $(W^{(i)}(1),\ldots,W^{(i)}(q_i))$ with probability

$$\begin{aligned}
&a(W^{(i)}(1),\ldots,W^{(i)}(q_i))\\
&= \binom{\#\left\{j\in\{1,\ldots,q_i\}\colon W^{(i)}(j) > n\gamma_i\right\}}{\check{l}(m_i;i)}^{-1} \mathbb{1}_{\{\#\{j\in\{1,\ldots,q_i\}\colon W^{(i)}(j)>n\gamma_i\}\leq \hat{l}(m_i;i)\}}.
\end{aligned}$$

Finally, we are able to give the pseudocode of this sampling scheme in Algorithm 3 below. For its expected running time, an analogous result to Proposition 2.3.1 is formulated in Proposition 2.3.3.

Algorithm 3 Generating the sample path of $\bar{X}_n^{(1)}, \dots, \bar{X}_n^{(d)}$ under $\mathbf{Q}_\gamma$

```
sample (m_1, ..., m_{d-1}) ~ h_{1;m_1,...,m_{d-1}}
for i = 1 to d do
    sample q_i ~ h_{2;q_i}; R ← true
    while R = true do
        sample {b_k}_{k ≤ ľ(m_i;i)} ~ unif( C( {1, ..., q_i}, ľ(m_i;i) ) )
        for j ∈ {b_k}_{k ≤ ľ(m_i;i)} do
            sample W^(i)(j) ~ W^(i)(1) | W^(i)(1) > nγ_i
        for j ∉ {b_k}_{k ≤ ľ(m_i;i)} do
            sample W^(i)(j) ~ W^(i)(1)
        c ← #{ j ∈ {1, ..., q_i} : W^(i)(j) > nγ_i }; a ← 0
        if c < l̂(m_i;i) then
            a ← binom(c, ľ(m_i;i))^{-1}
        sample u ~ uniform[0,1]; R ← true
        if u < a then
            R ← false
  return X̄_n^(1), ..., X̄_n^(d)
```

Proposition 2.3.3. *Let $T_{alg3}(n)$ denote the expected running time of Algorithm 3. Under the assumption that $W^{(i)}(1)$ is regularly varying of index $-\beta_i < -1$, for all $i \in \{1, \dots, d\}$, we have that $T_{alg3}(n)$ is uniformly bounded from above w.r.t. n, i.e. $\max_{n \geq 0} T_{alg3}(n) < \infty$.*

The discussion above shows that sampling from the conditional distribution $\mathbf{Q}_\gamma(\,\cdot\,)$ is tractable. As we mentioned in the introduction, our estimator is straightforward to implement. Moreover, its strong efficiency, which is formulated in Theorem 2.3.2, can be proved based on Lemma 2.7.1. Moreover, we state in Theorem 2.3.2 that our estimator is strongly efficient. Without introducing any new notations, we formulate a corollary to address a special case, where it is sufficient to consider a smaller (in the sense of cardinality) set than $J_{(l_1^*, \dots, l_d^*)}$ as in Definition 2.3.1. Note that Corollary 2.3.1 can be shown by following similar arguments as in the proofs of Lemma 2.7.1 and Theorem 2.3.2, thus the proof is omitted.

Theorem 2.3.2. *Let $B_n^\gamma \triangleq \{\bar{X}_n \in B^\gamma\}$, where B^γ is as defined in* (2.3.4)*. Under Assumption 2.3.2, there exists γ such that the estimator given by* (2.3.3) *is strongly efficient for estimating $\mathbf{P}(A_n)$.*

Corollary 2.3.1. *Along with Assumption 2.3.2, we assume additionally that there exists an index set $I \subseteq J_{(l_1^*,\ldots,l_d^*)}$ and $r > 0$ such that, for every $\xi \in A$, there exists $(l_1, \ldots, l_d) \in I$ satisfying $d\left(\xi, \mathbb{C}_{(l_1,\ldots,l_d)}\right) \geq r$. Set $\tilde{J}_{(l_1^*,\ldots,l_d^*)} = I \setminus \Delta I$, where $(l_1, \ldots, l_d) \in \Delta I$ if and only if*

- *$(l_1, \ldots, l_d) \in I$ satisfies that $\mathcal{I}(l_1, \ldots, l_d) > 2\mathcal{I}(l_1^*, \ldots, l_d^*)$; and*
- *for every $(l_1', \ldots, l_d') \in I \setminus \{(l_1, \ldots, l_d)\}$, we have that $\mathcal{I}(l_1, \ldots, l_d) \neq \mathcal{I}(l_1', \ldots, l_d')$.*

Setting $B_n^\gamma = \{\bar{X}_n \in B^\gamma\}$ with $B^\gamma \triangleq \bigcup_{(l_1,\ldots,l_d) \in \tilde{J}_{(l_1^*,\ldots,l_d^*)}} B^{\gamma;l}$, there exists γ such that the estimator given by (2.3.3) *is strongly efficient for estimating $\mathbf{P}(A_n)$.*

Remark 2.6. Even though our simulation algorithm is constructed in the context of Poisson processes with positive jump distributions, it can be easily generalized to the case, where the jump distributions are regularly varying at both $-\infty$ and ∞ (for details see the proof of Theorem 3.5 in [105] and the references therein).

Remark 2.7. We end this section with a final remark to point out the connection between the one-dimensional case and the multidimensional case. That is, if we set $d = 1$ then Assumption 2.3.2 coincides with Assumption 2.3.1 and no additional conditions are imposed on the set A. Moreover, the auxiliary sets B^γ in both cases are essentially the same. Thus, Theorem 2.3.2 can be considered as a special case of Theorem 2.3.1.

2.3.3 Extension to random walks

Let S_k, $k \geq 0$, be a centered random walk with increments $\{Y_k\}_{k\geq 1}$. Let $\mathbf{P}(Y_1 \leq -x)$ be regularly varying with index $-\alpha$ and let $\mathbf{P}(Y_1 \geq x)$ be regularly varying with index $-\beta$. Define $\bar{S}_n(t) \triangleq S_{\lfloor nt \rfloor}/n$, $t \geq 0$. In this subsection, we want to design an efficient simulation algorithm for estimating the probability of $\bar{S}_n \in A$. As in Sections 2.3.1 and 2.3.2, we make the following assumption for the set A.

Assumption 2.3.3. Assume that A is a measurable set bounded away from $\mathbb{D}_{<l_-^*;l_+^*}$, where (l_-^*, l_+^*) is the unique solution of (2.2.3). Moreover, assume that $C_{l_-^*,l_+^*}(A^\circ) > 0$.

Then, we construct the auxiliary set B^γ as follows.

Definition 2.3.2. Let (l_-^*, l_+^*) denote the unique solution to (2.2.3), and let

$$\begin{aligned} J_{l_-^*;l_+^*} \triangleq \{&(l_-, l_+) \in \mathbb{Z}_+^2 \setminus I_{<l_-^*;l_+^*} : \\ &(m_-, m_+) \prec (l_-, l_+) \text{ implies } (m_-, m_+) \in I_{<l_-^*;l_+^*}\}, \end{aligned}$$

where $I_{<l_-^*;l_+^*}$ is as in (2.2.1). For $\gamma_- > 0$ and $\gamma_+ > 0$, define

$$B^\gamma \triangleq \bigcup_{(l_-,l_+)\in J_{l_-^*;l_+^*}} B^{\gamma;l_-^*;l_+^*}, \tag{2.3.6}$$

where $B^{\gamma;l_-^*;l_+^*} \triangleq \{\xi \in \mathbb{D}\colon \#\{t \,|\, \xi(t^-) - \xi(t) > \gamma_-\} \geq l_-^*, \#\{t \,\big|\, \xi(t) - \xi(t^-) > \gamma_+\} \geq l_+^*\}$.

Defining $A_n \triangleq \{\bar{S}_n \in A\}$ and $B_n^\gamma \triangleq \{\bar{S}_n \in B^\gamma\}$, we propose an estimator for $\mathbf{P}(\bar{S}_n \in A)$ that is given by (2.3.3). Note that, computing $\mathbf{P}(\bar{S}_n \in B^\gamma)$, as well as generating the sample path $\bar{S}_n$ under $\mathbf{Q}^\gamma$ can be achieved by following similar strategies as in Sections 2.3.1 and 2.3.2. Hence, the details are omitted (for examples, see Sections 2.4 and 2.5 below). We state the strong efficiency of our estimator in the following theorem without giving the proof, since it is analogous to the proof of Theorem 2.3.2.

Theorem 2.3.3. *Let $B_n^\gamma \triangleq \{\bar{X}_n \in B^\gamma\}$, where B^γ is as defined in* (2.3.6). *Under Assumption 2.3.3, there exist γ_- and γ_+ such that the estimator given by* (2.3.3) *is strongly efficient for estimating $\mathbf{P}(A_n)$.*

With the results presented in this section at hand, we are able to apply our general simulation algorithm to three examples in the next sections. These examples can be found in the applications of mathematical finance, actuarial science and queueing networks.

2.4 An application to finite-time ruin probabilities

2.4.1 Problem setting

Let S_k, $k \geq 0$, be a centered random walk with increments $\{Y_k\}_{k\geq 1}$. Moreover, let $\mathbf{P}(Y_1 > x)$ be regularly varying at infinity with index $-\beta$. For $c \geq 0$, define

$A_n \triangleq \{\max_{0\leq k\leq n} Y_k \leq nb, \max_{0\leq k\leq n} S_k - ck \geq na\}$. Additionally, we make a technical assumption that $a/b \notin \mathbb{Z}$. We are interested in computing $\mathbf{P}(A_n)$. This probability is particularly interesting, since it is related to, for example, insurance, where huge claims may be reinsured and therefore are irrelevant in the sense of estimating the finite-time ruin probability of an insurance company.

2.4.2 Large deviations results

The rare-event probability can be estimated efficiently using the technique introduced in Section 2.3. To see this, define $A \triangleq \{\xi \in \mathbb{D}\colon \sup_{t\in[0,1]}[\xi(t) - ct] \geq a; \sup_{t\in[0,1]}[\xi(t) - \xi(t^-)] \leq b\}$ and $\bar{S}_n \triangleq \{\bar{S}_n(t)\}_{t\in[0,1]}$, where $\bar{S}_n(t) = S_{\lfloor nt\rfloor}/n$ for $t \geq 0$. Note that $\mathbf{P}(A_n) = \mathbf{P}\left(\bar{S}_n \in A\right)$. Set $l^* = \lceil a/b\rceil$. Intuitively, l^* should be the key parameter, as it takes at least l^* jumps of size b to cross level a. This intuition has been made rigorous by Rhee et al. in [105, Section 5.1], where the authors show that A is bounded away from $\mathbb{D}_{<l^*}$, and hence, $\mathbf{P}(A_n) = \Theta\left(n^{l^*}\mathbf{P}(S_1 \geq n)^{l^*}\right)$.

2.4.3 Construction of B^γ

We set $B^\gamma = \{\xi \in \mathbb{D}\colon \#\{t \,|\, \xi(t) - \xi(t^-) > \gamma\} \geq l^*\}$ and $B_n^\gamma = \left\{\bar{S}_n \in B^\gamma\right\} = \{\#\{k \in \{1, \ldots, n\}\colon Y_k > n\gamma\} \geq l^*\}$, where γ is the parameter that needs to be tuned. For the completeness of our algorithm, we give a closed-form expression for $\mathbf{P}(B_n^\gamma)$. Let p denote the probability of $\mathbf{P}(Y_1 > \gamma n)$, then we have that

$$\mathbf{P}(B_n^\gamma) = \sum_{i=l^*}^{n} \binom{n}{i} p^i (1-p)^{n-i} = 1 - \sum_{i=0}^{l^*-1} \binom{n}{i} p^i (1-p)^{n-i}, \qquad (2.4.1)$$

where the latter representation in (2.4.1) is for numerical purposes.

2.4.4 Choice of γ

As we mentioned in Remark 2.3, a strategy of choosing the parameters γ needs to be discussed in the next step. From the proof of Theorem 2.3.2, it is sufficient to select γ such that $\mathbf{P}\left(A_n \cap (B_n^\gamma)^c\right) = o\left(\mathbf{P}(A_n)^2\right)$. We propose to select γ such that $\left(a - (l^*-1)b\right)/\gamma \notin \mathbb{Z}_+$, and that

$$\left\lceil \frac{a - (l^*-1)b}{\gamma} \right\rceil > l^* + 1. \qquad (2.4.2)$$

In view of Theorem 2.3.3, it is sufficient to show that $A \cap (B^\gamma)^c$ is bounded away from $\mathbb{D}_{<2l^*+1}$ with γ satisfying (2.4.2). To see this, choose θ with $d(\theta, \mathbb{D}_{<2l^*+1}) < r$. This implies that there exists $\xi \in \mathbb{D}_{<2l^*+1}$ satisfying $d(\theta, \xi) < r$ and $\xi(t) = \sum_{j=1}^{2l^*} x_j \mathbb{1}_{[u_j,1]}(t)$. In particular, there exists a homeomorphism $\lambda\colon [0,1] \to [0,1]$ satisfying $||\, \lambda - \mathrm{id}\, ||_\infty \vee ||\, \xi \circ \lambda - \theta\, ||_\infty < r$. Hence, for $\theta \in A$, using the identity $\xi \circ \lambda = \theta + (\xi \circ \lambda - \theta)$, we conclude that the following holds:

1. $x_j < b + 2r$, for every $j \in \{1, \dots, 2l^*\}$; and
2. there exists t' such that $\sum_{u_j \le 1} x_j \ge \sum_{u_j \le \lambda(t')} x_j > a - 2r$.

To see this, note that $\xi \circ \lambda(t) = \sum_{j=1}^{2l^*} x_j \mathbb{1}_{[u_j,1]}(\lambda(t)) = \theta + (\xi \circ \lambda - \theta)$ and $\|\xi \circ \lambda - \theta\|_\infty < r$. By the fact that $\sup_{t\in[0,1]}[\theta(t) - \theta(t^-)] \le b$, conclusion (1) follows. Moreover, by the fact that $\sup_{t\in[0,1]}[\theta(t) - ct] \ge a$, there exists t' such that $\xi \circ \lambda(t') = \sum_{u_j \le \lambda(t')} x_j > a - 2r$, and hence, conclusion (2) is obtained. This implies that $\sum_{j \ge l^*} x_j > a - 2r - (l^* - 1)(b + 2r)$. Moreover, for $\theta \in (B^\gamma)^c$, every jump of ξ should be bounded by $\gamma + 2r$ after having $l^* - 1$ jumps with size bigger than b. Due to the fact that γ satisfies (2.4.2) and a is not a multiple of b, we obtain the result by choosing r small.

2.4.5 Sampling from $\mathbf{Q}_\gamma$

Summarizing the discussion in the previous paragraphs, we are able to propose a strongly efficient estimator for $\mathbf{P}(A_n)$ that is given by (2.3.3). As the last ingredient of our simulation algorithm, a strategy of sampling from $\mathbf{Q}_\gamma(\,\cdot\,)$ $(= \mathbf{P}(\,\cdot\,|B_n^\gamma))$ needs to be discussed. We use a similar strategy as in Algorithm 3 and formulate the pseudocode in Algorithm 4.

2.4.6 Numerical results

Finally, we investigate our algorithm numerically based on a concrete example. Let $Y_1 = Y_1' - \mathbf{E}Y_1'$, where $\mathbf{P}(Y_1' > t) = (1/t)^\beta$. In Table 2.1 we select $c = 0.05$, $w = 0.05$ (for a heuristic of the choice of w and its impact on the empirical performance see Section 2.5 below) and summarize the estimated probability and the level of precision (ratio between the radius of the 95% confidence interval and the estimated value) for different combinations of n, β, a, b, and c (based on 10^6 samples). We observe that, for different values of β, a and b, the precision stays roughly constant as n grows. This confirms our theoretical results.

Algorithm 4

$R \leftarrow$ **true**
while $R =$ **true** **do**
  sample $(i_1, \ldots, i_{l^*})$ uniformly from $\mathcal{C}(\{1, \ldots, n\}, l^*)$
  for $j \in \{i_1, \ldots, i_{l^*}\}$ **do**
    sample $Y_j \sim Y_1 : \gamma n < Y_1 \leq bn$
  for $j \notin \{i_1, \ldots, i_{l^*}\}$ **do**
    sample $Y_j \sim Y_1$
  sample $u \sim$ uniform$[0, 1]$; $c \leftarrow \#\{m \in \{1, \ldots, n\} : \gamma n < Y_1 \leq bn\}$; $a \leftarrow \binom{c}{l^*}$; $R \leftarrow$ **true**
  if $u < a^{-1}$ **then**
    $R \leftarrow$ **false**
  return $(Y_1, \ldots, Y_n)$

Est	$n = 80$		$n = 100$		$n = 200$	
Pr	$\beta = 1.5$	$\beta = 2.0$	$\beta = 1.5$	$\beta = 2.0$	$\beta = 1.5$	$\beta = 2.0$
$a = 2, b = 1.2$	1.171×10^{-3}	3.904×10^{-5}	1.043×10^{-3}	2.361×10^{-5}	6.316×10^{-4}	5.167×10^{-6}
$(l^* = 2)$	2.053×10^{-2}	3.133×10^{-2}	2.057×10^{-2}	3.376×10^{-2}	2.130×10^{-2}	3.975×10^{-2}
$a = 4, b = 1.2$	5.099×10^{-7}	3.778×10^{-10}	3.860×10^{-7}	1.592×10^{-10}	1.326×10^{-7}	8.911×10^{-12}
$(l^* = 4)$	1.799×10^{-2}	2.278×10^{-2}	1.761×10^{-2}	2.366×10^{-2}	1.717×10^{-2}	2.780×10^{-2}
$a = 2, b = 0.3$	1.635×10^{-10}	1.147×10^{-12}	1.795×10^{-10}	3.983×10^{-13}	1.202×10^{-10}	6.775×10^{-15}
$(l^* = 7)$	6.441×10^{-2}	1.662×10^{-2}	5.456×10^{-2}	1.635×10^{-2}	3.535×10^{-2}	1.826×10^{-2}

Table 2.1: Estimated rare-event probability and level of precision for the application as described in Section 2.4.

2.5 An application in barrier option pricing

In this section we consider an application that arises in the context of financial mathematics; in particular we consider a down-in barrier option (see Section 11.3 in [111]).

2.5.1 Problem setting

Let S_k, $k \geq 0$, be a centered random walk with increments $\{Y_k\}_{k \geq 1}$. Let $\mathbf{P}(Y_1 \leq -x)$ be regularly varying with index $-\alpha$ and let $\mathbf{P}(Y_1 \geq x)$ be regularly varying with index $-\beta$. Let a, b and c be positive real numbers. We provide a strongly efficient estimator for the probability of

$$A_n \triangleq \left\{ S_n \geq bn, \min_{0 \leq k \leq n} S_k + ck \leq -an \right\},$$

which can be interpreted as the chance of exercising a down-in barrier option. This application is interesting, since, as we will see, the large deviations behavior of $\mathbf{P}(A_n)$ is caused by two large jumps.

2.5.2 Large deviations results

Define $A \triangleq \{\xi \in \mathbb{D} : \xi(1) \geq b,\ \inf_{0\leq t\leq 1} \xi(t) + ct \leq -a\}$. Obviously, we have that $(l_-^*, l_+^*) = (1,1)$, where (l_-^*, l_+^*) denotes the solution to (2.2.3). To verify the topological property of A, we define $m, \pi_1 : \mathbb{D} \to \mathbb{R}$ by $m(\xi) = \inf_{0\leq t\leq 1}\{\xi(t)+ct\}$, and $\pi_1(\xi) = \xi(1)$. Note that F, π_1 and m are continuous, therefore $F^{-1}(A) = m^{-1}(-\infty, -a] \cap \pi_1^{-1}[b, \infty)$ is a closed set. By adapting the results in [105, Section 5.2], it can be shown that, for any arbitrary $i \geq 0$, $\mathbb{D}_{i;0}$ and $\mathbb{D}_{0;i}$ are bounded away from $m^{-1}(-\infty, -a]$ and $\pi_1^{-1}[b, \infty)$, respectively. Hence, A is bounded away from $\mathbb{D}_{<1;1}$. Applying Result 2.2.3, we obtain that $\mathbf{P}(\bar{X}_n \in A) = \Theta\left(n^2 \mathbf{P}(S_1 \geq n)\mathbf{P}(S_1 \leq -n)\right)$.

2.5.3 Construction of B^γ

Now we are in the framework of Theorem 2.3.3. Note that by Definition 2.3.2 we have $J_{1;1} = \{(1,1), (l,0), (0,m)\}$, where $l = \min\{l' \in \mathbb{Z}_+ : (l'-1)(\beta-1) > (\alpha-1)\}$ and $m = \min\{m' \in \mathbb{Z}_+ : (m'-1)(\alpha-1) > (\beta-1)\}$. However, adapting the idea behind Corollary 2.3.1 together with the fact that A is bounded away from both $\mathbb{D}_{i;0}$ and $\mathbb{D}_{0;i}$, it is sufficient to consider $\tilde{J}_{1;1} = \{(1,1)\}$. Hence, we can set $B^\gamma = \big\{\xi \in \mathbb{D} \colon \#\big\{t \,\big|\, \xi(t^-) - \xi(t) > \gamma_-\big\} \geq 1,\ \#\big\{t \,\big|\, \xi(t) - \xi(t^-) > \gamma_+\big\} \geq 1\big\}$. As we mentioned in the introduction, it is possible that estimators may be crafted specifically for the events of interest, in order to obtain (up to constant factors) better performance. Due to the fact that at least one downward jump should happen *before* upward jumps, without introducing new notations, we can modify B^γ such that $B^\gamma = \big\{\xi \in \mathbb{D} \colon \exists\, t_1 < t_2 : \xi(t_1^-) - \xi(t_1) > \gamma_-,\ \xi(t_2) - \xi(t_2^-) > \gamma_+\big\}$. This implies that $B_n^\gamma = \big\{\exists\, i < j : Y_i < -\gamma_- n,\ Y_j > \gamma_+ n\big\}$. By a straightforward computation, we obtain that $\mathbf{P}\left(B_n^\gamma\right) = 1 - \frac{p_2}{p_2 - p_1}(1-p_1)^n + \frac{p_1}{p_2-p_1}(1-p_2)^n$, where $p_1 \triangleq \mathbf{P}(Y_1 > \gamma_+ n)$ and $p_2 \triangleq \mathbf{P}(Y_1 < -\gamma_- n)$.

2.5.4 Choice of γ_- and γ_+

We discuss here the strategy of choosing the parameters γ_- and γ_+. From the proof of Theorem 2.3.2, it is sufficient to select γ_-, γ_+ such that $\mathbf{P}\left(A_n \cap (B_n^\gamma)^c\right) =$

$o\left(\mathbf{P}(A_n)^2\right)$. Hence, we propose to choose γ_- and γ_+ such that

$$((a+b)/\gamma_+, a/\gamma_-) \notin \mathbb{Z}_+^2,$$

and that

$$\min\left\{(\alpha-1)+\left\lceil\frac{a+b}{\gamma_+}\right\rceil(\beta-1), \left\lceil\frac{a}{\gamma_-}\right\rceil(\alpha-1)+(\beta-1)\right\} > 2(\alpha+\beta-2). \tag{2.5.1}$$

W.l.o.g. we assume that $\lceil a/\gamma_-\rceil(\alpha-1)+(\beta-1)$ is the unique minimum of (2.5.1). It suffices to prove that $A\cap(B^\gamma)^c$ is bounded away from $\mathbb{D}_{<\lceil a/\gamma_2\rceil;1}$. To show that $\bigcup_{(l_-,l_+)}\mathbb{D}_{<l_-;l_+}$ with $l_- \leq \lceil a/\gamma_2\rceil - 1$ is bounded away from $A\cap(B^\gamma)^c$, choose θ with $d(\theta, \bigcup_{(l_-,l_+)}\mathbb{D}_{<l_-;l_+}) < r$. This implies that there exists $\xi \in \bigcup_{(l_-,l_+)}\mathbb{D}_{<l_-;l_+}$ satisfying $d(\theta,\xi) < r$, where $\xi = \sum_{k=1}^{l_+} x_k \mathbb{1}_{[u_k,1]}(t) - \sum_{k=1}^{l_-} y_k \mathbb{1}_{[v_k,1]}(t)$. In particular, there exists a homeomorphism $\lambda\colon [0,1]\to[0,1]$ satisfying

$$||\,\lambda - \mathrm{id}\,||_\infty \vee ||\,\xi\circ\lambda - \theta\,||_\infty < r. \tag{2.5.2}$$

Using (2.5.2) and the identity $\xi\circ\lambda = \theta + (\xi\circ\lambda - \theta)$, we conclude that, for $\theta \in (B^\gamma)^c$ and $t \in [0,1]$, at least one of the following holds:

- $x_k \leq \gamma_+ + 2r$, for every $u_k \geq t$; or
- $y_k \leq \gamma_- + 2r$, for every $v_k < t$.

For $\theta \in m^{-1}(-\infty,-a]$, by (2.5.2) there exists t' such that

$$\sum_{u_j\leq\lambda(t')} x_j - \sum_{v_j\leq\lambda(t')} y_j < -a+3r. \tag{2.5.3}$$

Moreover, we can assume that $y_j \leq \gamma_- + 2r$ for j satisfying $v_j \leq \lambda(t')$. Otherwise x_j is bounded by $\gamma_+ + 2r$ for j satisfying $v_j > \lambda(t')$. By choosing r sufficiently small, it contradicts the fact that $\theta \in \pi_1^{-1}[b,\infty)$ and $\lceil a/\gamma_-\rceil(\alpha-1)+(\beta-1)$ is the minimum of (2.5.1). Hence, (2.5.3) implies that $(\lceil a/\gamma_-\rceil - 1)(\gamma_- + 2r) > a - 3r$. Since $(\lceil a/\gamma_-\rceil - 1)\gamma_- < a$, choosing r sufficiently small we obtain the result. Similarly, it can be shown that $A\cap(B^\gamma)^c$ is bounded away from $\bigcup_{(l_-,l_+)}\mathbb{D}_{<l_-;l_+}$ for $l_+ \leq \lceil(a+b)/\gamma_+\rceil - 1$.

2.5.5 Sampling from $\mathbf{Q}_\gamma$

As in Section 2.4, a strategy of sampling from $\mathbf{Q}_\gamma(\,\cdot\,)$ needs to be discussed. Even though B^γ is modified to obtain smaller relative error, a similar strategy as in Algorithm 3 can be used here. Hence we omit the details.

2.5.6 Numerical results

We end this section with some numerical investigations. First let $Y_1 = Y_1' - \mathbf{E}Y_1'$, where Y_1' is a random variable with density function f_Y that is given by $f_Y(y) = \frac{1}{3}\left(\frac{\beta-1}{y}\right)^{\beta} \mathbb{1}_{(1,\infty)}(y) + \frac{1}{3}\left(-\frac{\alpha-1}{y}\right)^{\alpha} \mathbb{1}_{(-\infty,-1)}(y) + \frac{1}{6}\mathbb{1}_{[-1,1]}(y)$. We apply our algorithm to estimate $\mathbf{P}(S_n \geq bn, \min_{0\leq k\leq n} S_k \leq -an)$ with $a = 2$ and $b = 1.5$. In Figure 2.2 we plot the precision of the estimated probability against the parameter w for different values of n. We observe that the estimated probabilities become more precise as w decreases. This heuristic suggests the upper bound we derive in (2.7.7), where the latter term in (2.7.7) is of the order $o(\mathbf{P}(A_n)^2)$, as long as w is strictly positive. Based on this observation, we choose $w = 0.05$ for all numerical investigations presented in this chapter. In Table 2.2 we compare the estimated rare-event probability and precision w.r.t. different values of n, α and β. We observe that the precision stays roughly constant as n increases for different combinations of α and β, which suggests the strong efficiency of our estimator.

Next, we make a comparison between the algorithms developed in this chapter and in [70], where a simulation algorithm is designed for estimating $\mathbf{P}(S_n \geq bn)$ from an MCMC perspective. First note that, instead of unbiased estimators, MCMC algorithms give us only consistent estimators. Furthermore, note that the event $\{S_n \geq bn\}$ is a special case of the event studied in this section with $a = 0$. Here we consider Y_1 with density function $f_Y(y) = 2(y+1)^{-3}$ for $y \geq 0$. In Table 2.3 we present the estimated rare-event probability, the level of precision, the computational time (in seconds), and the normalized workload—i.e. the (estimated) standard deviation multiplied by the computational time divided by the sample mean—produced by the two algorithms, based on 10^6 samples. Note that our algorithm typically outperforms the MCMC algorithm in terms of *computational time*—especially as n increases—while producing a slightly larger *coefficient of variation* compared to the MCMC algorithm. Overall, our algorithm seems to be more efficient for larger values of n, while MCMC seems to be more efficient for small values of n in terms of *normalized workload*. This can be explained by the fact that our estimator is state-independent, i.e., the increments of the random walk S_n can be updated simultaneously. On the other hand, in the MCMC case, the algorithm needs a burn-in period to converge, and the increments have to be simulated following a specific order. Moreover, updating the value of, say, Y_k relies on the values of $Y_1, \ldots, Y_{k-1}, Y_{k+1}, \ldots, Y_n$ to run the MCMC algorithm. It may be noted that the range of probabilities examined in Table 2.3 is smaller than the typical range of practical interest in applications.

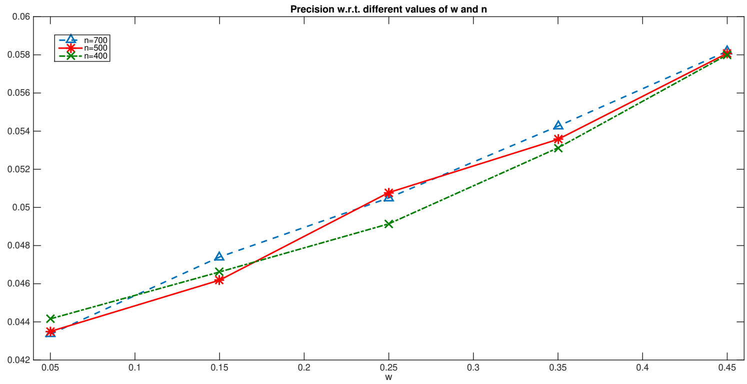

Figure 2.2: A plot of the precision of the estimators discussed in Section 2.5 w.r.t. w for different values of n.

For example, insurance companies (according to Solvency II) are suggested to have a capital reserve corresponding to bankruptcy events in the order of a 0.5% likelihood (on an annual basis). Nevertheless, when model uncertainty is taken into account, considering a bankruptcy probability of the putative (assumed) parametric model in the range of likelihood that is considerably smaller than nominal values suggested by regulation may be necessary. For example, calibrating any distribution of claims with a degree of precision corresponding to a bankruptcy probability of 0.5% in a non-parametric way is practically impossible since one would need (due to the central limit theorem) millions of observations on an annual basis. In such a case, an additional safety margin should be added to the capital requirement to account for model error as discussed in [19] (see Section 9.2.1). That is, one may have to consider the range of likelihood in the order of 10^{-4} to 10^{-5} for the putative parametric model to ensure the 0.5% likelihood for the true claims. Note that our importance sampling algorithm seems to be comparable or preferable to the MCMC algorithm in this range.

2.6 An application to queueing networks

In this section, an application to queueing networks is considered. More specifically, the probability of the number of customers in a subset of the system crossing a high level is estimated. Although some particular cases exist that

Est Pr	$n = 250$	$n = 500$	$n = 750$	$n = 1000$	$n = 1250$	$n = 1500$
$\alpha = 2, \quad \beta = 1.5$	3.913×10^{-7} 0.043	1.370×10^{-7} 0.043	6.992×10^{-8} 0.044	4.539×10^{-8} 0.044	3.305×10^{-8} 0.044	2.471×10^{-8} 0.044
$\alpha = 1.8, \beta = 1.7$	3.322×10^{-7} 0.037	1.154×10^{-7} 0.037	6.040×10^{-8} 0.038	3.840×10^{-8} 0.038	2.870×10^{-8} 0.038	2.225×10^{-8} 0.037
$\alpha = 2.3, \beta = 2$	1.923×10^{-9} 0.053	4.004×10^{-10} 0.053	1.491×10^{-10} 0.054	7.601×10^{-11} 0.054	4.632×10^{-11} 0.054	3.072×10^{-11} 0.054
$\alpha = 2.7, \beta = 1.8$	6.838×10^{-10} 0.068	1.121×10^{-10} 0.070	4.092×10^{-11} 0.070	2.079×10^{-11} 0.069	1.105×10^{-11} 0.071	6.896×10^{-12} 0.071

Table 2.2: Estimated rare-event probability and level of precision for the application as described in Section 2.5.

Est Pr Time (s) NW	$n = 5$		$n = 20$		$n = 200$		$n = 1000$	
	MCMC	IS	MCMC	IS	MCMC	IS	MCMC	IS
$b = 20$	5.340×10^{-4} 0.587×10^{-3} 25.5 7.633	5.286×10^{-4} 1.020×10^{-3} 19.6 10.196	1.375×10^{-4} 0.636×10^{-3} 67.9 22.046	1.369×10^{-4} 1.060×10^{-3} 21.5 11.624	1.384×10^{-5} 0.645×10^{-3} 561.0 184.488	1.384×10^{-5} 1.071×10^{-3} 44.4 24.267	2.770×10^{-6} 0.644×10^{-3} 3686.8 1211.854	2.769×10^{-6} 1.073×10^{-3} 145.0 79.370
$b = 150$	8.962×10^{-6} 2.052×10^{-4} 24.8 2.594	8.958×10^{-6} 8.287×10^{-4} 17.1 7.227	2.250×10^{-6} 2.239×10^{-4} 63.1 7.204	2.250×10^{-6} 8.331×10^{-4} 19.9 8.468	2.252×10^{-7} 2.289×10^{-4} 545.7 63.712	2.252×10^{-7} 8.353×10^{-4} 43.7 18.633	4.505×10^{-8} 2.289×10^{-4} 3687.0 430.399	4.503×10^{-8} 8.370×10^{-4} 146.4 62.543
$b = 1000$	2.002×10^{-7} 0.796×10^{-5} 27.8 1.130	2.001×10^{-7} 8.051×10^{-4} 17.5 7.206	5.008×10^{-8} 0.848×10^{-5} 65.9 2.854	5.005×10^{-8} 8.068×10^{-4} 20.4 8.384	5.009×10^{-9} 0.851×10^{-5} 577.4 25.070	5.011×10^{-9} 8.042×10^{-4} 45.5 18.684	1.002×10^{-9} 0.881×10^{-5} 4011.9 180.317	1.002×10^{-9} 8.052×10^{-4} 149.7 61.482

Table 2.3: Estimated rare-event probability, level of precision, and computational time for estimating $\mathbf{P}(S_n \geq bn)$ using the algorithms introduced in this chapter and in [70].

allow for an explicit analysis (see e.g. Section 13 in [41]), it is hard to come up with exact results for the distribution of the workload process in general. Hence, implementing our algorithm in such a context is particularly interesting.

2.6.1 Model description and preliminaries

To be specific, we consider a d-dimensional stochastic fluid model. Suppose that jobs arrive to the i-th station in the network according to a Poisson process with unit rate, which is denoted by $\{N^{(i)}(t)\}_{t\geq 0}$ and independent of $\{N^{(j)}(t)\}_{t\geq 0}$ for $j \neq i$. Moreover, the k-th arrival of the i-th station brings a job of size $W^{(i)}(k)$. We are assuming that $\{W(k) \triangleq (W^{(1)}(k), \ldots, W^{(d)}(k))^T\}_{k\geq 1}$ is a sequence of i.i.d. positive random vectors and that $\{W(k)\}_{k\geq 1}$ is independent of $\{N(t)\}_{t\geq 0}$. Therefore, the total amount of external work that arrives to the i-th station up to time t is given by $J^{(i)}(t) = \sum_{k=1}^{N^{(i)}(t)} W^{(i)}(k)$. Now, assume that the workload at the i-th station is processed as a fluid by the server at a rate r_i and that a proportion $Q_{ij} \geq 0$ of the fluid processed by the i-th station is routed to the j-th server. Moreover, we assume that Q is a substochastic matrix with $Q_{ii} = 0$ and that $Q^n \to 0$ as $n \to \infty$. The dynamics of the model are expressed formally by the so-called Skorokhod map (for details see e.g. [109], [110], [71]), that is defined in terms of a pair of processes (Z, Y) satisfying a stochastic differential equation that we shall describe now. Let $R = (I - Q)^T$, $r = (r_1, \ldots, r_d)^T$, $X(t) \triangleq J(t) - Rrt$ and $Z^{(i)}(t)$ denote the workload of the i-th station at time t. Then, for given $Z^{(i)}(0)$, we have that

$$dZ(t) = dX(t) + RdY(t), \tag{2.6.1}$$

where $Y(\cdot)$ encodes the minimal amount of pushing required to keep $Z(\cdot)$ non-negative. In order to describe how to characterize the solution (Z, Y) to (2.6.1), we need to introduce some notations. Let $\psi : \mathbb{D}^d \to \mathbb{D}^d_{\uparrow}$ be such that

$$\psi(x) \triangleq \inf\left\{w \in \mathbb{D}^d_{\uparrow}\colon\ x + Rw \geq 0\right\},$$

i.e., $\psi^{(i)}(x)(t) \triangleq \inf\{w^{(i)}(t) \in \mathbb{R}\colon w \in \mathbb{D}^d_{\uparrow},\ x + Rw \geq 0\}$, for all i and t, and $\phi : \mathbb{D}^d \to \mathbb{D}^d$ with $\phi(x) \triangleq x + R\psi(x)$. The following results summarize useful properties and characterizations of the Skorokhod mappings ψ, ϕ, as well as the workload process $Z(t)$:

Result 2.6.1 (Theorem 14.2.1, 14.2.5, and 14.2.7 of [117]). *The mappings ψ and ϕ are well-defined for all $x \in \mathbb{D}^d$. Moreover, ψ and ϕ are Lipschitz continuous* *w.r.t. both the uniform metric and the Skorokhod J_1 metric. If $Y(t) \triangleq \psi(X)(t)$ and $Z(t) \triangleq \phi(X)(t)$, then $(Y(t), Z(t))$ solve the Skorokhod problem given by (2.6.1).*

Result 2.6.2 (Lemma 14.3.3, Corollary 14.3.4 and Corollary 14.3.5 of [117])**.** *Let $x \in \mathbb{D}^d$. For the discontinuity points of $\psi(x)$ (denoted by Disc$(\psi(x))$) and $\phi(x)$, we have that Disc$(\psi(x)) \cup$ Disc$(\phi(x)) =$ Disc(x). Moreover, if x has only positive jumps, then $\psi(x)$ is continuous and $\phi(x)(t) - \phi(x)(t^-) = x(t) - x(t^-)$.*

Result 2.6.3 (Theorem 14.2.2 of [117])**.** *The regulator map $y = \psi(x)$ can be characterized as the unique fixed point of the map $\pi_{x,Q} : \mathbb{D}^d_\uparrow \to \mathbb{D}^d_\uparrow$, where $\pi_{x,Q}(w)(t) \triangleq \max\{0, \sup_{s\in[0,t]}(Q^T w(s) - x(s))\}$.*

Result 2.6.4 (Consequence of Theorem 4.1 of [101])**.** *Let $\Delta \in \mathbb{D}^d$ be a non-decreasing function such that $\Delta(0) \geq 0$. Then, for $x \in \mathbb{D}^d$, we have that $\psi(x) \geq \psi(x+\Delta)$, $\phi(x) \leq \phi(x+\Delta)$, and $\phi(x)(t_2)-\phi(x)(t_1) \leq \phi(x+\Delta)(t_1)-\phi(x+\Delta)(t_2)$, for any $0 \leq t_1 \leq t_2 \leq 1$.*

Finally, we assume that the right tail of $W^{(i)}(1)$ is regularly varying with index $-\beta_i$ and that the stability condition holds, i.e. $R^{-1}\rho < r$, where $\rho \triangleq \mathbf{E}J(1)$. Let $\bar{Z}_n(t) \triangleq Z(nt)/n$ and $\bar{X}_n(t) \triangleq X(nt)/n$. Let $c \in \{0,1\}^d$ be a binary vector, and let $\mathcal{J}_c$ denote the index set encoded by c, i.e., $j \in \mathcal{J}_c$ if $c_j = 1$. Set $\bar{Z}_n(t) \triangleq Z(nt)/n$ and $\bar{X}_n(t) \triangleq X(nt)/n$. Define $l_c : \mathbb{R}^d \to \mathbb{R}$ by $l_c(x) = c^T x$ and $\pi_1 : \mathbb{D}^d \to \mathbb{R}^d$ by $\pi_1(\xi) = \xi(1)$. We are interested in estimating the probability of the subset, encoded by c, of the scaled workload $\bar{Z}_n$ exceeds level a at time 1, that is, $\mathbf{P}\left(c^T \bar{Z}_n(1) \geq a\right)$. By Theorem 14.2.6 (iii) of [117], we have that $\bar{Z}_n = \phi(\bar{X}_n)$, and hence it holds that, for $a > 0$,

$$\mathbf{P}\left(c^T \bar{Z}_n(1) \geq a\right) = \mathbf{P}\left(F(\bar{X}_n) \geq a\right) = \mathbf{P}(\bar{X}_n \in A), \tag{2.6.2}$$

where $F \triangleq l_c \circ \pi_1 \circ \phi$ and $A \triangleq \{\xi \in \mathbb{D} : F(\xi) \geq a\}$.

2.6.2 Large deviations results

To obtain the large deviations asymptotics for the rare-event probability as in (2.6.2), we proceed in the manner, as follows.

- To determine the tail index of the rare-event probability, we first study the optimization problem given by (2.2.2) and transform it into a (nonstandard) knapsack problem with nonlinear constraints (see (2.6.8) and Proposition 2.6.1 below).

- Under a certain assumption (see Assumption 2.6.1 below), we show that A, as defined in (2.6.2), is bounded away from $\mathbb{D}_{<(l_1^*,\dots,l_d^*)}$, where $l_1^*,\dots,l_d^*$ is the optimal solution to the knapsack problem derived in the first step.
- Finally, we derive a large deviations result for $\mathbf{P}(\bar{X}_n \in A)$ by applying Result 2.2.2.

We start with the optimization problem given by (2.2.2). Due to the fact that $X(t)$ is in general not a compensated compound Poisson process but one with a certain drift, it is convenient to consider a slightly different problem, which is given by

$$\operatorname*{arg\,min}_{(l_1,\dots,l_d)\in\mathbb{Z}_+^d,\ \prod_{i=1}^d \mathbb{L}_{l_i}(\mu_i)\cap A\neq\emptyset} \mathcal{I}(l_1,\dots,l_d), \tag{2.6.3}$$

where $\mu \triangleq \mathbf{E}X(1) = \rho - Rr$, $r' = r - R^{-1}\rho > 0$ due to the stability condition, and $\mathbb{L}_{l_i}(\mu_i) \triangleq \{\xi\colon \exists \xi' \in \mathbb{D}_{l_i} : \xi(t) = \xi'(t) + \mu_i t = \xi'(t) - (Rr')_i t\}$. Define $E_0 \triangleq \big\{(l_1,\dots,l_d) \in \mathbb{Z}_+^d : l_i = 0,\ \forall i \in \mathcal{J}_c\big\}$ and $E_1 \triangleq \{e_i : i \in \mathcal{J}_c\}$, where e_i denotes the unit vector with entries 0 except for the i-th coordinate. By Result 2.6.2, instead of (2.6.3) we can solve two separate problems that are given by

$$\begin{aligned} &\operatorname*{arg\,min}_{(l_1,\dots,l_d)\in E_0,\ \prod_{i=1}^d \mathbb{L}_{l_i}(\mu_i)\cap A\neq\emptyset} \mathcal{I}(l_1,\dots,l_d),\\ \text{and}\quad &\operatorname*{arg\,min}_{(l_1,\dots,l_d)\in E_1,\ \prod_{i=1}^d \mathbb{L}_{l_i}(\mu_i)\cap A\neq\emptyset} \mathcal{I}(l_1,\dots,l_d). \end{aligned} \tag{2.6.4}$$

Note that the latter problem in (2.6.4) can be solved easily by considering $\min_{i\in\mathcal{J}_c} \beta_i - 1$, therefore we focus on the first problem in (2.6.4). Let $\mathcal{J}$ be a subset of $(\mathcal{J}_c)^c$. Moreover, let $\theta \in \mathbb{D}_1$ and let $\xi \in \mathbb{D}^d$ be such that

$$\xi^{(i)}(t) = \begin{cases} -(Rr')_i t,\ t\in[0,1], & \text{for } i \notin \mathcal{J},\\ \theta^{(i)} - (Rr')_i t,\ t \in [0,1], & \text{for } i \in \mathcal{J}. \end{cases} \tag{2.6.5}$$

A necessary and sufficient condition for the existence of $\xi \in A$ is given in the following proposition.

Proposition 2.6.1. *Let $\mathcal{J} \subseteq (\mathcal{J}_c)^c$. Moreover, let $\{r_i^*\}_{i\notin\mathcal{J}}$ be such that*

$$r_i^* = \max\left\{ r_i' - \sum_{j\neq i} Q_{ji} r_j' + \sum_{\substack{j\neq i\\ j\notin\mathcal{J}}} Q_{ji} r_j^*, 0 \right\}, \quad \textit{for } i \notin \mathcal{J}. \tag{2.6.6}$$

Define

$$\partial_z(\mathcal{J}) \triangleq \sum_{i\in\mathcal{J}_c}\left(r_i^* - r_i' + \sum_{j\neq i} Q_{ji}r_j' - \sum_{\substack{j\neq i\\ j\notin\mathcal{J}}} Q_{ji}r_j^*\right). \tag{2.6.7}$$

If $\partial_z(\mathcal{J}) \neq a$, then there exists ξ satisfying (2.6.5) *and $c^T\phi(\xi)(1) \geq a$, if and only if $\partial_z(\mathcal{J}) > a$. Additionally, if $\mathcal{J}_1 \subseteq \mathcal{J}_2 \subseteq (\mathcal{J}_c)^c$, then we have that $\partial_z(\mathcal{J}_1) \leq \partial_z(\mathcal{J}_2)$.*

We give a sketch of the proof and refer to Section 2.7 for details. Note that $\partial_z(\mathcal{J})$ given by (2.6.7) is the increasing rate of the subset $\mathcal{I}_c$ of the workload process, whose associated input process does not have any jumps but starts with a sufficiently large initial value. Based on this observation, a ξ can be constructed for the "if"-part of the first statement. For the "only if"-part, suppose that there is a ξ satisfying $c^T\phi(\xi)(1) \geq a$. By Result 2.6.4, enlarging the size of jumps in ξ will preserve the fact that $c^T\phi(\xi)(1) \geq a$. Hence, we can construct a new ξ, such that

- the associated workload process $\phi(\xi)$ is piecewise linear between two neighboring discontinuity points; and
- the increasing rate of $c^T\phi(\xi)$ is always smaller or equal than $\partial_z(\mathcal{J})$ given by (2.6.7).

Remark 2.8. Note that (2.6.6) can be written in a matrix notation that is given by $r^* = \max\{((I-Q^T)r')_{\notin\mathcal{J}} + (Q_{\notin\mathcal{J}})^T r^*, 0\} = \max\{(Rr-\rho)_{\notin\mathcal{J}} + (Q_{\notin\mathcal{J}})^T r^*, 0\}$, where $(Rr-\rho)_{\notin\mathcal{J}}$ and $Q_{\notin\mathcal{J}}$ denote the vector and matrix respectively with its i-th row and column being removed for all $i \in \mathcal{J}$. Using the Banach fixed-point theorem, we obtain that $r^* = \lim_{n\to\infty}\underline{\pi}^n(0)$, where $\underline{\pi}^n \triangleq \underline{\pi} \circ \underline{\pi}^{n-1}$ and $\underline{\pi}(x) \triangleq \max\{(Rr-\rho)_{\notin\mathcal{J}} + (Q_{\notin\mathcal{J}})^T x, 0\}$.

Define $E_{\mathcal{J}} \triangleq \{(l_1,\ldots,l_d) \in E_0' \colon \partial_z\left(\mathcal{J}_{(l_1,\ldots,l_d)}\right) > a\}$, where $E_0' \triangleq E_0 \cap \{(l_1,\ldots,l_d) \in \mathbb{Z}_+^d \colon l_i \in \{0,1\}, \forall i \notin \mathcal{J}_c\}$ and $\partial_z\left(\mathcal{J}_{(l_1,\ldots,l_d)}\right)$ is as defined in (2.6.7) with $\mathcal{J}_{(l_1,\ldots,l_d)}$ denoting the index set encoded by $(l_1,\ldots,l_d) \in E_0'$. By Proposition 2.6.1, we conclude that the first problem in (2.6.4) is equivalent to

$$\operatorname*{arg\,min}_{(l_1,\ldots,l_d)\in E_{\mathcal{J}}} \mathcal{I}(l_1,\ldots,l_d). \tag{2.6.8}$$

Thanks to the last statement of Proposition 2.6.1, it is unnecessary to check every $(l_1,\ldots,l_d) \in E_{\mathcal{J}}$ for solving (2.6.8). In the following example, we consider

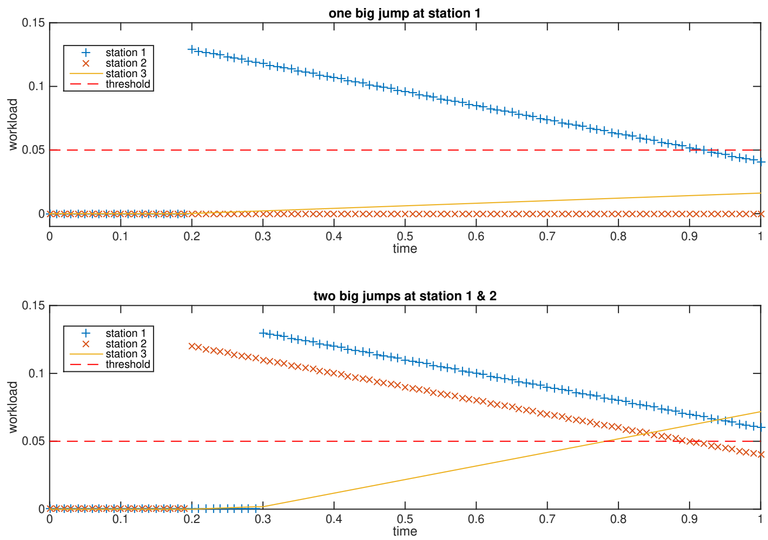

Figure 2.3: An illustration of two different sample paths of workload processes (in the setting of Example 2.6.1), whose associated input processes have the form as in (2.6.5).

a specific fluid network and illustrate how to solve (2.6.8) using Proposition 2.6.1.

Example 2.6.1. Consider the fluid network given by $\rho = (0.8\ 0.8\ 1)^T$, $r = (1\ 1\ 2.5)^T$ and

$$Q = \begin{bmatrix} 0 & 0.1 & 0.8 \\ 0.1 & 0 & 0.8 \\ 0 & 0 & 0 \end{bmatrix}.$$

We are interested in the probability of the rare event that the third station crosses the level na at time n for large n, i.e. $\mathcal{J}_c = \{3\}$. It is easy to check that the stability condition holds. By an easy computation, we obtain that $\partial_z(\{1, 2\}) = 0.1$ and $\partial_z(\{1\}) = \partial_z(\{2\}) = 0.02$. For $a = 0.05$, the optimal solution to (2.6.8) is given by $(1, 1, 0)$.

Suppose that we have solved (2.6.8). To obtain the large deviations results, the following technical assumption needs to be made.

Assumption 2.6.1. Assume that (2.6.8) satisfies the following conditions.

a) The optimization problem given by (2.6.8) has a unique solution.

b) For every $\mathcal{J} \subseteq (\mathcal{J}_c)^c$, it holds that $\partial_z(\mathcal{J}) \neq a$.

c) Let $(l_1^*, \ldots, l_d^*)$ denote the optimal solution to (2.6.8). We assume that $\mathcal{I}(l_1^*, \ldots, l_d^*) < \min_{i \in \mathcal{J}_c} \beta_i - 1$.

By Result 2.2.2, Assumption 2.6.1 c) implies that the objective value of the first problem in (2.6.4) is strictly less than the objective value of the latter one in (2.6.4), and hence the optimal solution $(l_1^*, \ldots, l_d^*)$ to (2.6.8) solves (2.6.3). In view of this observation, the rare event is caused by multiple large jumps. Throughout the rest of this section, we assume that Assumption 2.6.1 holds. We end this subsection with a large deviations result for $\mathbf{P}(\bar{X}_n \in A) = \mathbf{P}\left(c^T \bar{Z}_n(1) \geq a\right)$, which is formulated in the following proposition:

Proposition 2.6.2. *Suppose that Assumption 2.6.1 holds. Let F be as defined in* (2.6.2)*. Then $A = F^{-1}[a, \infty)$ is bounded away from*

$$\bigcup_{(l_1, \ldots, l_d) \in I_{<(l_1^*, \ldots, l_d^*)}} \prod_{i=1}^{d} \mathbb{L}_{l_i}(\mu_i),$$

where $(l_1^*, \ldots, l_d^*)$ denotes the unique optimal solution of (2.6.8)*. Moreover,*

$$\begin{aligned}
&C_{l_1^*} \times \cdots \times C_{l_d^*}\left((F^{-1}[a, \infty))^\circ\right) \\
&\leq \liminf_{n \to \infty} \frac{\mathbf{P}\left(\bar{X}_n \in A\right)}{\prod_{i=1}^{d} \left(n\nu_i[n, \infty)\right)^{l_i^*}} \leq \limsup_{n \to \infty} \frac{\mathbf{P}\left(\bar{X}_n \in A\right)}{\prod_{i=1}^{d} \left(n\nu_i[n, \infty)\right)^{l_i^*}} \\
&\leq C_{l_1^*} \times \cdots \times C_{l_d^*}\left(F^{-1}[a, \infty)\right).
\end{aligned}$$

2.6.3 Simulation

Again, we are in the setting of Theorem 2.3.2. To be able to discuss the choice of $J_{(l_1^*, \ldots, l_d^*)}$ and the parameter γ in a more precise context, let us consider the stochastic fluid network introduced in Example 2.6.1.

Example 2.6.1 (continued)**.** Recall that, for $a = 0.05$, the optimal solution of (2.6.3) is given by $\beta_1 + \beta_2 - 2$, if we assume that $\beta_1 + \beta_2 - 2 < \beta_3 - 1$. Moreover, it can be easily shown that A is bounded away from both $\mathbb{D}_{<i} \times \mathbb{D}_0 \times \mathbb{D}_0$ and

$\mathbb{D}_0 \times \mathbb{D}_{<j} \times \mathbb{D}_0$. Combining this with $\mathcal{I}(1,1,1) > 2\mathcal{I}(1,1,0)$, as well as Corollary 2.3.1, it is sufficient to take $\tilde{J}_{(l_1^*,\dots,l_d^*)} = \{(1,1,0),(0,0,1)\}$. This implies that

$$\begin{aligned} B_n^\gamma = &\Big\{\#\{k \,|\, W^{(i)}(k) > n\gamma_i, k \le N^{(i)}(n)\} \ge 1, i \in \{1,2\}\Big\} \\ &\cup \Big\{\#\{k \,|\, W^{(3)}(k) > n\gamma_3, k \le N^{(3)}(n)\} \ge 1\Big\}, \end{aligned}$$

and hence, $(B_n^\gamma)^c = \{\exists i \in \{1,2\} : W^{(i)}(k) \le n\gamma_i, \forall k \le N^{(i)}(n)\} \cap \{W^{(3)}(k) \le n\gamma_3, \forall k \le N^{(3)}(n)\}$.

We choose γ such that $\mathbf{P}(A_n \cap (B_n^\gamma)^c) = o(\mathbf{P}(A_n)^2)$. To begin with, we assume w.l.o.g. that $\beta_3 - 1 \le 2(\beta_1 + \beta_2 - 2)$, otherwise we can simply set $J_{(l_1^*,\dots,l_d^*)} = \{(1,1,0)\}$, since $\mathcal{I}(0,0,1) > 2\mathcal{I}(1,1,0)$. Now the parameter γ_3 can be chosen such that $\lceil \frac{1}{20}/\gamma_3 \rceil(\beta_3 - 1) > 2(\beta_1 + \beta_2 - 2)$. For the choice of γ_1, we observe that the job arriving at the second station can have an arbitrarily large size. Hence, it is sufficient to consider the inequality $\partial_z(\{1,2\})t' + \partial_z(\{2\})(1-t') > a$, where $\partial_z(\{1,2\}) = 0.1$ and $\partial_z(\{2\}) = 0.02$. Solving the inequality we obtain that $t' < 3/8$. This simply means that the workload of the third station cannot exceed the level a at time 1 if we keep both of the first and the second stations overloaded less than 3/8 of the time. Since the workload process of the first station decays at rate 1/10, one can choose γ_1 such that $\lceil \frac{3}{80}/\gamma_1 \rceil(\beta_1 - 1) + (\beta_2 - 1) > 2(\beta_1 + \beta_2 - 2)$. Analogously, it is sufficient to set γ_2 such that $(\beta_1 - 1) + \lceil \frac{3}{80}/\gamma_2 \rceil(\beta_2 - 1) > 2(\beta_1 + \beta_2 - 2)$.

We give a closed-form expression for $\mathbf{P}(B_n^\gamma)$. By assumption $\{W^{(i)}(k)\}_{1 \le i \le d}$ are mutually independent. Therefore

$$\begin{aligned} \mathbf{P}((B_n^\gamma)^c) &= \mathbf{P}\left(\exists i \in \{1,2\} : W^{(i)}(k) \le n\gamma_i, \forall k \le N^{(i)}(n)\right) \\ &\quad \times \mathbf{P}\left(W^{(3)}(k) \le n\gamma_3, \forall k \le N^{(3)}(n)\right) \\ &= \left[1 - \prod_{i=1}^{2}\left(1 - \mathbf{P}\left(W^{(i)}(k) \le n\gamma_i, \forall k \le N^{(i)}(n)\right)\right)\right] \\ &\quad \times \mathbf{P}\left(W^{(3)}(k) \le n\gamma_3, \forall k \le N^{(3)}(n)\right). \end{aligned}$$

Conditionally on $N^{(i)}(n)$, we obtain

$$\mathbf{P}(W^{(i)}(k) \le n\gamma_i, \forall k \le N^{(i)}(n)) = \exp\{-n(1 - \mathbf{P}(W^{(i)}(1) \le n\gamma_i))\},$$

since $N^{(i)}(n)$ is Poisson distributed. Summarizing the above findings, we are able to propose a strongly efficient estimator for $\mathbf{P}(A_n)$ that is given by

Est Pr	$n = 1200$	$n = 1600$	$n = 2000$	$n = 2400$
$\beta_1 = 1.5, \beta_2 = 1.5, \beta_3 = 2.2$	7.719×10^{-2} 0.045	6.228×10^{-2} 0.058	4.541×10^{-2} 0.057	3.973×10^{-2} 0.057
Est Pr	$n = 800$	$n = 1200$	$n = 1600$	$n = 2000$
$\beta_1 = 2.5, \beta_2 = 2.3, \beta_3 = 4$	2.894×10^{-2} 0.325	1.686×10^{-2} 0.404	6.153×10^{-3} 0.445	2.023×10^{-3} 0.448
Est Pr	$n = 600$	$n = 1000$	$n = 1400$	$n = 1800$
$\beta_1 = 2.2, \beta_2 = 2.9, \beta_3 = 4.5$	5.139×10^{-2} 0.249	1.858×10^{-2} 0.347	9.987×10^{-3} 0.351	1.028×10^{-3} 0.377

Table 2.4: Estimated rare-event probability and level of precision for the application as described in Section 2.6.3.

(2.3.3). Moreover, Algorithm 3 can be used to sample from $\mathbf{Q}^{\gamma}$. To see this, we decompose B_n^{γ} into two disjoint sets $B_n^{\gamma}(1)$ and $B_n^{\gamma}(2)$ that are given by $B_n^{\gamma}(1) \triangleq \{\#\{k \mid W^{(3)}(k) > n\gamma_3,\, k \le N^{(3)}(n)\} \ge 1\}$, and

$$B_n^{\gamma}(2) \triangleq \Big\{\#\{W^{(i)}(k) > n\gamma_i,\, k \le N^{(i)}(n)\} \ge 1, \forall i \in \{1,2\}\Big\} \\ \cap \{W^{(3)}(k) \le n\gamma_3,\, \forall 1 \le k \le N^{(3)}(n)\},$$

respectively. Using Algorithm 3, the sample paths of $\bar{X}_n^{(1)}, \bar{X}_n^{(2)}, \bar{X}_n^{(3)}$ can be simulated independently on both $B_n^{\gamma}(1)$ and $B_n^{\gamma}(2)$. We present the numerical results based on 20000 samples in Table 2.4. We choose $W^{(i)}(1)$ such that $\mathbf{P}(W^{(i)}(1) > t) = (t_{r,i}/t)^{\beta_i}$ and $t_{r,i} = \rho_i(\beta_i - 1)/\beta_i$, for $i \in \{1,2,3\}$. As one can see, the numerical results show that the level of precision stay stable as n increases, and hence, confirm again what our theory predicts.

2.7 Proofs

In this section we provide proofs of the results presented in this chapter.

Proof of Proposition 2.3.1. Recall that the expected running time of the rejection sampling (see Algorithm 1 above), which is used to generate the jumps of $\bar{X}_n$, is given by $M_l = \binom{l}{l^*}\mathbf{P}(W(1) > n\gamma)^{l^*}\mathbf{P}(B_n^{\gamma}|N(n) = l)^{-1}$. Hence, for the expected running time of Algorithm 1, denoted by $T_{\mathrm{alg1}}(n)$, we have that

$$\begin{aligned}
T_{\mathrm{alg1}}(n) &= \sum_{l\geq l^*} h_l M_l = \mathbf{P}(B_n^\gamma)^{-1}\sum_{l\geq l^*}\mathbf{P}(B_n^\gamma|N(n)=l)\mathbf{P}(N(n)=l)M_l \\
&= \mathbf{P}(B_n^\gamma)^{-1}\sum_{l\geq l^*}\mathbf{P}(N(n)=l)\binom{l}{l^*}\mathbf{P}(W(1)>n\gamma)^{l^*} \\
&= \frac{n^{l^*}(\lambda\mathbf{P}(W(1)>n\gamma))^{l^*}}{\mathbf{P}(B_n^\gamma)}e^{-\lambda n}\sum_{l\geq l^*}\frac{(\lambda n)^{l-l^*}}{(l-l^*)!} = \frac{n^{l^*}(\lambda\mathbf{P}(W(1)>n\gamma))^{l^*}}{\mathbf{P}(B_n^\gamma)}.
\end{aligned}$$

Recall that $B_n^\gamma \triangleq \{\bar{X}_n \in B^\gamma\}$, where $B^\gamma \triangleq \{\xi\colon \#\{t \,|\, \xi(t)-\xi(t^-)>\gamma\} \geq l^*\}$. Noting that B_n^γ is bounded away from $\mathbb{D}_{<l^*}$ and $l^* = \min\{l\in\mathbb{Z}_+ \colon \mathbb{D}_l \cap B_n^\gamma\}$, by Result 2.2.1 we obtain that

$$\limsup_{n\to\infty} T_{\mathrm{alg1}}(n) \leq n^{l^*}(\lambda\mathbf{P}(W(1)>n\gamma))^{l^*}\mathbf{P}(B_n^\gamma)^{-1} \leq C_{l^*}((B^\gamma)^\circ)^{-1} < \infty.$$

$\square$

Proof of Proposition 2.3.2. Recall

$$B^\gamma \triangleq \bigcup_{(l_1,\ldots,l_d)\in J_{(l_1^*,\ldots,l_d^*)}} B^{\gamma;l}$$

was defined in (2.3.4). Let $I \subseteq J_{(l_1^*,\ldots,l_d^*)}$. Define $B_I^{\gamma;l} \triangleq \bigcap_{(l_1,\ldots,l_d)\in I} B^{\gamma;l}$. By the inclusion-exclusion principle, we have that

$$\mathbf{P}(\bar{X}_n \in B^\gamma) = \sum_{k=1}^{|J_{(l_1^*,\ldots,l_d^*)}|}(-1)^{k-1}\sum_{|I|=k,\, I\subseteq J_{(l_1^*,\ldots,l_d^*)}}\mathbf{P}\left(\bar{X}_n \in B_I^{\gamma;l}\right). \tag{2.7.1}$$

Moreover, for any finite collection I of elements in $\mathbb{Z}_+^d$ with $I \subseteq J_{(l_1^*,\ldots,l_d^*)}$, we have that

$$\begin{aligned}
B_I^{\gamma;l} &= \bigcap_{i=1}^d \bigcap_{(l_1,\ldots,l_d)\in I}\left\{\left(\xi^{(1)},\ldots,\xi^{(d)}\right) : \#\{t \,|\, \xi^{(i)}(t)-\xi^{(i)}(t^-)>\gamma_i\} \geq l_i\right\} \\
&= \bigcap_{i=1}^d \left\{\left(\xi^{(1)},\ldots,\xi^{(d)}\right) : \#\{t \,|\, \xi^{(i)}(t)-\xi^{(i)}(t^-)>\gamma_i\} \geq \hat{l}_{i;I}\right\}, \qquad (2.7.2)
\end{aligned}$$

where $\hat{l}_{i;I} \triangleq \max_{(l_1,\ldots,l_d)\in I} l_i$. Since $\bar{X}_n^{(1)},\ldots,\bar{X}_n^{(d)}$ are independent processes, we obtain that

$$\begin{aligned}&\mathbf{P}(B_I^{\gamma;l})\\&=\prod_{i=1}^d\left(1-\exp\{-\lambda_i n\mathbf{P}(W^{(i)}(1)>n\gamma_i)\}\sum_{j=0}^{\hat{l}_{i;I}-1}\frac{(\lambda_i n)^j}{j!}\mathbf{P}(W^{(i)}(1)>n\gamma_i)^j\right).\end{aligned}$$

□

Proof of Lemma 2.3.1. Recall that $B^{\gamma;l}(i,j) \triangleq \{\xi\in\mathbb{D}^d\colon \#\{t \,|\, \xi^{(i)}(t)-\xi^{(i)}(t^-)>\gamma_i\}\geq (l(j))_i\}$. Hence, we have that

$$B^\gamma=\bigcup_{j=1}^{|J_{(l_1^*,\ldots,l_d^*)}|}\bigcap_{i=1}^d B^{\gamma;l}(i,j)=\bigcup_{j=1}^{|J_{(l_1^*,\ldots,l_d^*)}|}\left(B^{\gamma;l}(1,j)\cap\bigcap_{i=2}^d B^{\gamma;l}(i,j)\right). \tag{2.7.3}$$

By definition $\Delta B^{\gamma;l}(i,j)\triangleq B^{\gamma;l}(i,j)\setminus\left(\bigcup_{m=1}^{j-1}B^{\gamma;l}(i,m)\right)$. Therefore, we have that

$$B^{\gamma;l}(i,j)=\bigcup_{m_i=1}^{j}\Delta B^{\gamma;l}(i,j). \tag{2.7.4}$$

Plugging (2.7.4) into (2.7.3), we obtain that

$$\begin{aligned}B^\gamma&=\bigcup_{j=1}^{|J_{(l_1^*,\ldots,l_d^*)}|}\left(\left(\bigcup_{m_1=1}^{j}\Delta B^{\gamma;l}(1,m_1)\right)\cap\bigcap_{i=2}^d B^{\gamma;l}(i,j)\right)\\&=\bigcup_{m_1=1}^{|J_{(l_1^*,\ldots,l_d^*)}|}\left(\bigcup_{j=m_1}^{|J_{(l_1^*,\ldots,l_d^*)}|}\left(\Delta B^{\gamma;l}(1,m_1)\cap\bigcap_{i=2}^d B^{\gamma;l}(i,j)\right)\right)\\&=\bigcup_{m_1=1}^{|J_{(l_1^*,\ldots,l_d^*)}|}\left(\Delta B^{\gamma;l}(1,m_1)\cap\left(\bigcup_{j=1}^{m_1}\bigcap_{i=2}^d B^{\gamma;l}(i,j)\right)\right).\end{aligned}$$

Applying the same procedure to $\bigcup_{j=1}^{m_1}\bigcap_{i=2}^d B^{\gamma;l}(i,j)$, we obtain that

$$B^\gamma=\bigcup_{m_1=1}^{|J_{(l_1^*,\ldots,l_d^*)}|}\bigcup_{m_2=1}^{m_1}\left(\Delta B^{\gamma;l}(1,m_1)\cap\Delta B^{\gamma;l}(2,m_2)\cap\left(\bigcup_{j=1}^{m_2}\bigcap_{i=3}^d B^{\gamma;l}(i,j)\right)\right).$$

Iterating the same procedure $d-1$ times, we obtain that

$$B^\gamma = \bigcup_{m_1=1}^{|J_{(l_1^*,\ldots,l_d^*)}|} \bigcup_{m_2=1}^{m_1} \cdots \bigcup_{m_{d-1}=1}^{m_{d-2}} \left(\left(\bigcap_{i=1}^{d-1} \Delta B^{\gamma;l}(i, m_i) \right) \cap \left(\bigcup_{j=1}^{m_{d-1}} B^{\gamma;l}(d,j) \right) \right). \tag{2.7.5}$$

Since $l(1), \ldots, l(|J_{(l_1^*,\ldots,l_d^*)}|)$ are ordered such that $(l(1))_d \le \cdots \le (l(|J_{(l_1^*,\ldots,l_d^*)}|))_d$, we obtain that

$$\bigcup_{j=1}^{m_{d-1}} B^{\gamma;l}(d,j) = B^{\gamma;l}(d,1). \tag{2.7.6}$$

Plugging (2.7.6) into (2.7.5), we obtain that

$$B^\gamma = \bigcup_{m_1=1}^{|J_{(l_1^*,\ldots,l_d^*)}|} \bigcup_{m_2=1}^{m_1} \cdots \bigcup_{m_{d-1}=1}^{m_{d-2}} \left(\left(\bigcap_{i=1}^{d-1} \Delta B^{\gamma;l}(i, m_i) \right) \cap B^{\gamma;l}(d,1) \right).$$

□

Proof of Theorem 2.3.2. For the second moment of Z (under the change of measure) we have that

$$\begin{aligned} \mathbf{E}^{\mathbf{Q}_{\gamma,w}}[Z_n^2] &= \mathbf{E}[Z_n] = \mathbf{E}\left[Z_n \mathbb{1}_{B_n^\gamma}\right] + \mathbf{E}\left[Z_n \mathbb{1}_{(B_n^\gamma)^c}\right] \\ &\le \frac{1}{1-w}\mathbf{P}(A_n \cap B_n^\gamma)\mathbf{P}(B_n^\gamma) + \frac{1}{w}\mathbf{P}(A_n \cap (B_n^\gamma)^c) \\ &\le \frac{1}{1-w}\mathbf{P}(A_n)\mathbf{P}(B_n^\gamma) + \frac{1}{w}\mathbf{P}(A_n \cap (B_n^\gamma)^c). \end{aligned} \tag{2.7.7}$$

Combining this with Lemma 2.7.1 below we obtain the strong efficiency of our estimator. □

Lemma 2.7.1. *Let B^γ be as defined in* (2.3.4)*. Under Assumption 2.3.2, we have that $\mathbf{P}\big(\bar{X}_n \in B^\gamma\big) = \mathcal{O}\big(\mathbf{P}(\bar{X}_n \in A)\big)$. Moreover, there exists a γ, such that $\mathbf{P}\big(\bar{X}_n \in A \cap (B^\gamma)^c\big) = o\big(\mathbf{P}(\bar{X}_n \in A)^2\big)$.*

Proof of Lemma 2.7.1. First, note that $\mathbf{P}\big(\bar{X}_n \in B^\gamma\big) = \mathcal{O}\big(\mathbf{P}(\bar{X}_n \in A)\big)$ follows immediately from Result 2.2.2. It remains to show the existence of γ such that $\mathbf{P}\big(\bar{X}_n \in A \cap (B^\gamma)^c\big) = o(\mathbf{P}(\bar{X}_n \in A)^2)$. Since A is bounded away from $\mathbb{D}_{<(l_1^*,\ldots,l_d^*)}$

by assumption, there exists an r such that $d\left(A, \mathbb{D}_{<(l_1^*,\ldots,l_d^*)}\right) \geq r$. On the one hand, from [105] we have that

$$A \subseteq \Bigg\{ (\xi^{(1)}, \ldots, \xi^{(d)}) : \\ \exists (l_1, \ldots, l_d) \in J_{(l_1^*,\ldots,l_d^*)} : d\left(\xi^{(i)}, \mathbb{D}_{<l_i}\right) \geq r,\, \forall i \in \{1, \ldots, d\} \Bigg\}. \tag{2.7.8}$$

On the other hand, we have that

$$(B^\gamma)^c = \Bigg\{ (\xi^{(1)}, \ldots, \xi^{(d)}) : \\ \forall (l_1, \ldots, l_d) \in J_{(l_1^*,\ldots,l_d^*)} : \exists i : \#\big\{ t \,\big|\, \xi^{(i)}(t) - \xi^{(i)}(t^-) > \gamma_i \big\} \leq l_i - 1 \Bigg\}. \tag{2.7.9}$$

Let $\xi = \left(\xi^{(1)}, \ldots, \xi^{(d)}\right) \in A \cap (B^\gamma)^c$ be a step function in the set $\prod_{i=1}^d \mathbb{D}_{l_i'}$. By (2.7.8), there exists $(l_1, \ldots, l_d) \in J_{(l_1^*,\ldots,l_d^*)}$, such that $\xi^{(i)} = \sum_{j=1}^{l_i+m_i} c_j^{(i)} \mathbb{1}_{[t_j^{(i)},1]}$, $m_i \in \mathbb{Z}_+$ and $d\left(\xi^{(i)}, \mathbb{D}_{<l_i}\right) \geq r$ for all $i \in \{1, \ldots, d\}$ with $l_i \neq 0$. Combining $d\left(\xi^{(i)}, \mathbb{D}_{<l_i'}\right) \geq r$ with the fact that $\xi^{(i)} = \sum_{j=1}^{l_i-1} c_j^{(i)} \mathbb{1}_{[t_j^{(i)},1]} \in \mathbb{D}_{<l_i}$, we conclude that

$$\sum_{j=l_i}^{l_i+m_i} c_j^{(i)} \geq d\left(\sum_{j=1}^{l_i+m_i} c_j^{(k)} \mathbb{1}_{[t_j^{(i)},1]}, \sum_{j=1}^{l_i-1} c_j^{(i)} \mathbb{1}_{[t_j^{(i)},1]} \right) \geq r, \tag{2.7.10}$$

or in other words, the sum of the $m_i + 1$ smallest jumps is bounded from below by r for each $\xi^{(i)}$ of $\{\xi^{(i)}\}_{i \in \{1,\ldots,d\}}$ satisfying $l_i \neq 0$. Combining (2.7.9) with (2.7.10), as well as choosing γ_k sufficiently small, there exists at least one $k \in \{1, \ldots, d\}$ such that the smallest jump of $\xi^{(k)}$ is bounded from below by $r' > 0$ for arbitrary but fixed m_k. Repeating the same procedure as described above, for each $(l_1, \ldots, l_d) \in J_{(l_1^*,\ldots,l_d^*)}$, we can construct $(m_1, \ldots, m_d)$, such that the optimization problem given by

$$\operatorname*{arg\,min}_{(l_1,\ldots,l_d) \in \mathbb{Z}_+^d,\, \prod_{i=1}^d \mathbb{D}_{l_i} \cap A \cap (B^\gamma)^c \neq \emptyset} \mathcal{I}(l_1, \ldots, l_d), \tag{2.7.11}$$

has a unique solution $(l_1^{**}, \ldots, l_d^{**})$ satisfying $\mathcal{I}(l_1^{**}, \ldots, l_d^{**}) > 2\mathcal{I}(l_1^*, \ldots, l_d^*)$. We denote this specific choice of $(m_1, \ldots, m_d)$ for every $(l_1, \ldots, l_d) \in J_{(l_1^*,\ldots,l_d^*)}$ by

$\{m^{(l_1,\dots,l_d)}\}_{(l_1,\dots,l_d)\in J_{(l_1^*,\dots,l_d^*)}}$. It should be noted that the existence and the uniqueness of $(l_1^{**},\dots,l_d^{**})$ can be guaranteed by enlarging the set A (since we are looking for an upper bound for $\mathbf{P}(\bar{X}_n \in A\cap(B^\gamma)^c)$), together with choosing the corresponding γ_i sufficiently small. By Result 2.2.2, it remains to be shown that, under the chosen γ, the set $A\cap(B^\gamma)^c$ is bounded away from $\mathbb{D}_{<(l_1^{**},\dots,l_d^{**})}$. Select ξ satisfying $d\left(\xi,\mathbb{D}_{<(l_1^{**},\dots,l_d^{**})}\right)<\delta$, and hence, there exists $\theta\in\mathbb{D}_{<(l_1^{**},\dots,l_d^{**})}$ such that $d(\xi,\theta)<\delta$. On the one hand, combining $d(\xi,\theta)<\delta$ with (2.7.8), there exists $(l_1,\dots,l_d)\in J_{(l_1^*,\dots,l_d^*)}$ such that $d\left(\theta^{(i)},\mathbb{D}_{<l_i}\right)>r-\delta$, for all $i\in\{1,\dots,d\}$. Hence, we have that $\theta^{(i)}=\sum_{j=1}^{l_i+m_i}e_j^{(i)}\mathbb{1}_{[t_j^{(i)},1]}$, $m_i\in\mathbb{Z}_+$, satisfying

$$\sum_{j=l_i}^{l_i+m_i}e_j^{(i)}\ge d\left(\sum_{j=1}^{l_i+m_i}e_j^{(k)}\mathbb{1}_{[t_j^{(i)},1]},\sum_{j=1}^{l_i-1}e_j^{(i)}\mathbb{1}_{[t_j^{(i)},1]}\right)\ge r-\delta,\tag{2.7.12}$$

for all $i\in\{1,\dots,d\}$ with $l_i\neq 0$. On the other hand, there exist homeomorphisms $\{\lambda_i\}_{i\in\{1,\dots,d\}}$ such that

$$||\,\lambda_i-\mathrm{id}\,||_\infty\vee||\,\theta^{(i)}\circ\lambda_i-\xi^{(i)}\,||_\infty<\delta,\qquad\text{for all } i\in\{1,\dots,d\}.\tag{2.7.13}$$

Combining (2.7.13) with (2.7.9), we conclude the existence of at least one $i\in\{1,\dots,d\}$ such that

$$\#\left\{t\,\middle|\,\theta^{(i)}(t)-\theta^{(i)}(t^-)>\gamma_i-\delta\right\}\le l_i-1.\tag{2.7.14}$$

Since $\theta\in\mathbb{D}_{<(l_1^{**},\dots,l_d^{**})}$, we have that

$$m_i\le(m^{(l_1,\dots,l_d)})_i-1,\tag{2.7.15}$$

for some $i\in\{1,\dots,d\}$ with $l_i\neq 0$. On the one hand, (2.7.12) says that the sum of the m_i+1 smallest jumps is bounded from below by $r-\delta$ for all $i\in\{1,\dots,d\}$. On the other hand, (2.7.14) says that each one of the m_i+1 smallest jumps is bounded by $\gamma_i-\delta$ for all $i\in\{1,\dots,d\}$. Hence, by choosing δ sufficiently small, m_i+1 has to be so large that (2.7.15) gets violated. $\square$

Proof of Proposition 2.6.1. We derive a necessary and sufficient condition for $c^T\phi(\xi)(1)\ge a$ with ξ as in (2.6.5).

For the "only if"-part, suppose that $\partial_z(\mathcal{J}) > a$. Let $(v_1, \dots, v_d) \in \mathbb{R}^d_+$, $\delta \in (0,1)$ and ξ be such that

$$\xi^{(i)}(t) = \begin{cases} -(Rr')_i t,\ t \in [0,1], & \text{for } i \notin \mathcal{J}, \\ v_i \mathbb{1}_{[\delta,1]}(t) - (Rr')_i t,\ t \in [0,1], & \text{for } i \in \mathcal{J}. \end{cases}$$

Obviously ξ satisfies (2.6.5). For $t \in [0,\delta)$, by Result 2.6.3, the regulator process $y_\xi \triangleq \psi(\xi)$ should satisfy the fixed point equation that is given by $y^{(i)}_\xi(t) = \max\left\{0, \sup_{s\in[0,t]} \sum_{j\neq i} Q_{ji} y^{(j)}_\xi(s) + (Rr')_i s\right\}$, for all $i \in \{1, \dots, d\}$. Using the fact that $r' > 0$, we obtain that $y_\xi(t) = r't$, for $t \in [0,\delta)$. For $t \in [\delta, 1]$, again by Result 2.6.3, it holds that

$$y^{(i)}_\xi(t) = \max\left\{0, -v_i + \sup_{s\in[0,t]} r'_i s + \sum_{j\neq i} Q_{ji}(y^{(j)}_\xi(s) - r'_j s)\right\}, \quad \text{for all } i \in \mathcal{J}, \tag{2.7.16}$$

and

$$y^{(i)}_\xi(t) = \max\left\{0, \sup_{s\in[0,t]} r'_i s + \sum_{j\in\mathcal{J}} Q_{ji}(y^{(j)}_\xi(s) - r'_j s) + \sum_{\substack{j\neq i \\ j\notin\mathcal{J}}} Q_{ji}(y^{(j)}_\xi(s) - r'_j s)\right\}, \tag{2.7.17}$$

for all $i \notin \mathcal{J}$. Since $\{v_i\}_{i\in\mathcal{J}}$ are non-negative, by Result 2.6.2 we conclude that $y_\xi(s)$ and $r'_i s + \sum_{j\neq i} Q_{ji}(y^{(j)}_\xi(s) - r'_j s)$ are continuous in s on $[0,1]$. Using the Bolzano-Weierstrass theorem, there exists a set of sufficiently large $\{v_i\}_{i\in\mathcal{J}}$ (depending on y_ξ), such that $y^{(i)}_\xi(t) = y^{(i)}_\xi(\delta) = r'_i\delta$ for $i \in \mathcal{J}$. Plugging this into (2.7.17) along with setting $y^{(i)}_\xi(t) = r'_i\delta + r^*_i(t-\delta)$ for $i \notin \mathcal{J}$, $t \in [\delta, 1]$, we obtain that

$$\begin{aligned} & r'_i\delta + r^*_i(t-\delta) \\ &= \max\left\{0, \sup_{s\in[0,t]} r'_i s + \sum_{j\in\mathcal{J}} Q_{ji}(y^{(j)}_\xi(s) - r'_j s) + \sum_{\substack{j\neq i \\ j\notin\mathcal{J}}} Q_{ji}(y^{(j)}_\xi(s) - r'_j s)\right\} \end{aligned}$$

$$= \max\left\{ r_i'\delta,\, r_i'\delta + \max_{s\in[\delta,t]} r_i'(s-\delta) - \sum_{j\neq i} Q_{ji} r_j'(s-\delta) + \sum_{\substack{j\neq i\\ j\notin\mathcal{J}}} Q_{ji} r_j^*(s-\delta) \right\}, \tag{2.7.18}$$

for $i \in \mathcal{J}$. Note that (2.7.18) is solved by r_i^* satisfying (2.6.6). Moreover, by a straightforward computation, for the workload process $z_\xi \triangleq \phi(\xi)$, we obtain that $c^T z_\xi(1) = \partial_z(\mathcal{J})(1-\delta)$. Since by assumption $\partial_z(\mathcal{J}) > a$, we can choose δ such that $c^T z_\xi(1) \geq a$.

For the other direction of the proof, suppose that $c^T\phi(\xi)(1) \geq a$ for some ξ satisfying (2.6.5). Let the jump sizes and the associated jump times of ξ be denoted by $\{u_i\}_{i\in\mathcal{J}}$ and $\{t_i\}_{i\in\mathcal{J}}$, respectively. First we should mention that, by Result 2.6.4, enlarging $\{u_i\}_{i\in\mathcal{J}}$ will preserve the fact that $c^T\phi(\xi)(1) \geq a$. Moreover, let $d_1 < \cdots < d_m$ denote the discontinuity points of ξ with $m \leq |\mathcal{J}|$ and define $\mathcal{J}_i \triangleq \{k\colon t_k \leq d_i\}$, for every $i \in \{1,\ldots,m\}$. Now observe that $y_\xi(t) = r't$, $t\in[0,d_1)$. Hence, we have that $z_\xi'(t) = 0 \leq \partial_z(\mathcal{J})$, for $t \in [0,d_1)$. For $y_\xi(t)$, $t\in[d_1,d_2)$, we can easily check that

$$y_\xi^{(i)}(t) = \begin{cases} r_i' d_1, & \text{for all } i \in \mathcal{J}_1, \\ r_i' d_1 + r_i^{*,1}(t-d_1), & \text{for all } i \notin \mathcal{J}_1, \end{cases}$$

by taking sufficiently large $\{u_i\}_{i\in\mathcal{J}_1}$, where $r_i^{*,1} = \max\{r_i' - \sum_{j\neq i} Q_{ji} r_j' + \sum_{\substack{j\neq i\\ j\notin\mathcal{J}_1}} Q_{ji} r_j^{*,1}, 0\}$, for $i \notin \mathcal{J}_1$. Since $\mathcal{J}_1 \subseteq \mathcal{J}$, by Result 2.6.4 and (2.6.6), we conclude that $z_\xi'(t) \leq \partial_z(\mathcal{J})$, for $t \in [d_1,d_2)$. Defining $\mathcal{J}_1' \triangleq \mathcal{J}_1 \cup \{k\colon r_k^{*,1} = 0\}$, we consider $y_\xi(t)$ for $t \in [d_2,d_3)$. Following a similar argument as above, we claim that

$$y_\xi^{(i)}(t) = \begin{cases} r_i' d_1, & \text{for all } i \in \mathcal{J}_1', \\ r_i' d_1 + r_i^{*,1}(d_2 - d_1), & \text{for all } i \in \mathcal{J}_2 \setminus \mathcal{J}_1', \\ r_i' d_1 + r_i^{*,1}(d_2-d_1) + r_i^{*,2}(t-d_2), & \text{for all } i \notin \mathcal{J}_1' \cup \mathcal{J}_2, \end{cases}$$

for sufficiently large $\{u_i\}_{i\in\mathcal{J}_1\cup\mathcal{J}_2}$, where

$$r_i^{*,2} = \max\left\{ r_i' - \sum_{j\neq i} Q_{ji} r_j' + \sum_{j\neq i,\, j\notin\mathcal{J}_1'\cup\mathcal{J}_2} Q_{ji} r_j^{*,2}, 0 \right\},$$

for $i \notin \mathcal{J}_1' \cup \mathcal{J}_2$. Consider the fixed point equation that is given by

$$\tilde{r}_i^{*,2} = \max\left\{ r_i' - \sum_{j\neq i} Q_{ji} r_j' + \sum_{\substack{j\neq i \\ j\notin \mathcal{J}_1\cup\mathcal{J}_2}} Q_{ji}\tilde{r}_j^{*,2}, 0 \right\}, \quad \text{for } i \notin \mathcal{J}_1 \cup \mathcal{J}_2. \quad (2.7.19)$$

Since $\mathcal{J}_1 \subseteq \mathcal{J}_1 \cup \mathcal{J}_2$, by Result 2.6.4, we obtain that $\tilde{r}_k^{*,2} = 0$, for every $k \in \mathcal{J}_1' \setminus \mathcal{J}_1$. By making the convention that $r_k^{*,2} = 0$ for $k \in \mathcal{J}_1' \setminus \mathcal{J}_1$, we claim that $r_i^{*,2} = \tilde{r}_i^{*,2}$, for $i \notin \mathcal{J}_1 \cup \mathcal{J}_2$. Since $\mathcal{J}_1 \cup \mathcal{J}_2 \subseteq \mathcal{J}$, by Result 2.6.4, (2.7.19) and (2.6.6), we conclude that $z_\xi'(t) \leq \partial_z(\mathcal{J})$, for $t \in [d_2, d_3)$. Iterating the same procedure m more times, we can construct a ξ (by taking $\{u_i\}_{i\in\mathcal{J}}$ sufficiently large) such that z_ξ is piecewise linear between neighboring discontinuity points. Moreover, the increasing rate of z_ξ is less than $\partial_z(\mathcal{J})$, i.e. $z_\xi'(t) \leq \partial_z(\mathcal{J})$, for $t \in [0, 1]$. Therefore, we obtain that $\partial_z(\mathcal{J}) > a$.

The last statement of Proposition 2.6.1 is a consequence of Result 2.6.4. □

Proof of Proposition 2.6.2. Let the unique optimal solution of (2.6.8) be denoted by $(l_1^*, \dots, l_d^*)$. To prove that A is bounded away from

$$\bigcup_{(l_1,\dots,l_d)\in\mathcal{I}_{<l_1^*,\dots,l_d^*}} \prod_{i=1}^d \mathbb{L}_{l_i}(\mu_i),$$

it is sufficient to show that $A = F^{-1}[a, \infty)$ is bounded away from $\prod_{i=1}^d \mathbb{L}_{l_i}(\mu_i)$ for all $(l_1, \dots, l_d) \in \mathcal{I}_{<l_1^*,\dots,l_d^*}$. To begin with, let $(l_1, \dots, l_d) \in \mathcal{I}_{<l_1^*,\dots,l_d^*}$. Under Assumption 2.6.1 we have that $\partial_z(\mathcal{I}) < a$, where $j \in \mathcal{I}$ if and only if $l_j \neq 0$. Applying a similar approach as in the proof of Proposition 2.6.1, it can be shown that $F\left(\prod_{i=1}^d \mathbb{L}_{l_i}(\mu_i)\right) \subseteq (-\infty, \partial_z(\mathcal{I})]$. This implies that there exists $\delta > 0$ satisfying

$$d\left(F\left(\prod_{i=1}^d \mathbb{L}_{l_i}(\mu_i)\right), [a, \infty)\right) > \delta. \quad (2.7.20)$$

Moreover, by Result 2.6.1 we conclude that the mapping F as composition of Lipschitz continuous mappings (for continuity of π_1 see e.g. Theorem 12.5 in [14]) is again Lipschitz continuous. Let K_F denote the Lipschitz constant of F. Combining this with (2.7.20) we conclude that $d\left(\prod_{i=1}^d \mathbb{L}_{l_i}(\mu_i), F^{-1}([a, \infty))\right) > \delta/K_F$, hence the second statement is obtained by applying Result 2.2.2. □

Chapter 3

Ruin probabilities under reinsurance treaties

3.1 Introduction

We consider the following ruin problem of the classical Cramér-Lundberg model in risk theory; see e.g. [4]. Let $\{X_1, X_2, \dots\}$ be a sequence of i.i.d. positive random variables representing successive claim sizes that arrive according to a homogeneous Poisson process $N(t)$, $t \geq 0$, with rate λ. Premiums are received continuously at a constant rate $p > \lambda \mathbf{E}X$. We assume that there is also a reinsurance agreement in place, where $R(t)$ is the reinsured amount at time t. More precisely, if $S(t) = \sum_{i=1}^{N(t)} X_i$ is the aggregate claim amount at time t and p_D is the remaining premium for the insurer after reinsurance has been purchased, then the aggregate loss minus premiums at time t for the insurer is equal to $S(t) - p_D t - R(t)$. If $u \geq 0$ is the initial capital, then the probability of ruin before time T is defined as

$$\psi(u,T) = \mathbf{P}\left(\sup_{0 \leq t \leq T} \{S(t) - p_D t - R(t)\} > u \right). \tag{3.1.1}$$

We will restrict our attention to two forms of large claims reinsurance, namely LCR and ECOMOR. In an LCR (largest claim reinsurance) contract (see e.g. [3] for an early reference), the reinsurer agrees to cover the largest r claims, where $r \geq 1$ is a fixed number, while in an ECOMOR (excédent du coût moyen relatif) contract [113], the reinsurer covers the excess of the r largest claims over the $(r+1)$st largest claim; see [1] for more details on this type of reinsurance contracts.

We assume that the distribution of the claim sizes belongs to a class of distributions with a regularly varying tail, which is valid for many applications [54]. It is well known that the *principle of one big jump* holds in the heavy-tailed claim setting, i.e. ruin is typically caused by a single large claim. However, under the presence of large claim reinsurance contracts, ruin probabilities are typically harder to analyse because the largest claims are covered by the reinsurer and thus multiple claims may be responsible for the event of ruin.

Several papers have studied properties of large claim reinsurance contracts. For example, when claim sizes are light-tailed, the asymptotic tail behavior of the reinsured amounts is considered in [72, 78] and their joint tail behavior in [100]. For asymptotic properties of the reinsured amounts when the claim size distribution is heavy-tailed, see [2, 85]. For dependence between claim sizes and interarrival times in this context, see [86]. An interesting recent link between large claim treaties and risk measures is given in [28]. However, none of these contributions deal with the ruin probability, which is considered here.

In this chapter, we suggest to leverage recent new tools developed in the context of sample-path large deviations for heavy-tailed stochastic processes for the study of ruin problems under LCR and ECOMOR treaties. Concretely, for a centered Lévy process $Y(t)$, $t \geq 0$, with regularly varying Lévy measure ν, sample-path large deviations were developed in [105]. Consider the process $\bar{Y}_n = \{\bar{Y}_n(t), t \in [0,1]\}$, where $\bar{Y}_n(t) = Y(nt)/n$, $t \geq 0$. Then, asymptotic estimates of $\mathbf{P}(\bar{Y}_n \in A)$ for a large collection of sets A were derived. For Lévy processes with only positive jumps that are regularly varying with index $-\alpha$, $\alpha > 1$, these results take the form

$$\begin{aligned} C_{\mathcal{J}(A)}(A^\circ) &\leq \liminf_{n\to\infty} \frac{\mathbf{P}(\bar{Y}_n \in A)}{\big(n \cdot \nu[n,\infty)\big)^{\mathcal{J}(A)}} \\ &\leq \limsup_{n\to\infty} \frac{\mathbf{P}(\bar{Y}_n \in A)}{\big(n \cdot \nu[n,\infty)\big)^{\mathcal{J}(A)}} \leq C_{\mathcal{J}(A)}(A^-), \end{aligned} \tag{3.1.2}$$

where A° and A^- are the interior and closure of A, $\mathcal{J}(A)$ is interpreted as the minimum number of jumps in the Lévy process that are needed to cause the event A, and C_j is a measure. We will show how the reinsurance problem fits in the above framework. For this, we resolve several technical challenges such as showing how ruin probabilities in the reinsurance setting can be written as continuous maps of the input process in a suitable Skorokhod space.

Apart from the fact that reinsurance contracts are an interesting object of study in their own right, the present application seems to be the first example for which it is possible to compute the pre-factors in the asymptotics (3.1.2) explicitly. More precisely, we show for both the LCR and ECOMOR treaty that $C_{\mathcal{J}(A)}(A^{\circ}) = C_{\mathcal{J}(A)}(A^{-})$ and we provide an explicit expression for this value.

The rest of the chapter is organised as follows. In Section 3.2, we provide some preliminary results and introduce the necessary notation. Section 3.3 develops the main result, i.e. the tail asymptotics for finite-time ruin probabilities. For this, we are required to write (3.1.1) in terms of (3.1.2). This leads to the need to show continuity of certain mappings, as well as several additional technical requirements. In Section 3.4, we validate our asymptotic results with numerical experiments.

3.2 Model description and preliminaries

Following the notation and terminology used in Section 3.1, let F denote the distribution function of the claim sizes and $\mathbf{E}X$ be their expectation. We assume that F is regularly varying with index $-\alpha$, i.e. there exists a slowly varying function $L(x)$ such that $\bar{F}(x) := 1 - F(x) = L(x)x^{-\alpha}$, with $\alpha > 1$. Let further $X^{\star}_{1,N(t)} \geq X^{\star}_{2,N(t)} \geq \cdots \geq X^{\star}_{N(t),N(t)}$ denote the order statistics of $X_1, X_2, \ldots X_{N(t)}$.

In an LCR treaty, the reinsured amount $R(t)$ is equal to

$$L_r(t) := \sum_{i=1}^{r} X^{\star}_{i,N(t)}, \tag{3.2.1}$$

i.e. the r largest claims are paid by the reinsurer. On the other hand, the reinsured amount $R(t)$ in an ECOMOR treaty takes the form

$$E_r(t) := \sum_{i=1}^{r} X^{\star}_{i,N(t)} - rX^{\star}_{r+1,N(t)} = \sum_{i=1}^{N(t)} \left(X_i - X^{\star}_{r+1,N(t)}\right)_+. \tag{3.2.2}$$

That is, the ECOMOR constitutes an excess-of-loss treaty with a random retention, and the latter is the $(r+1)$st-largest claim. For more details and background on such reinsurance contracts, see [1]. In either treaty, the number of reinsured claims is equal to r.

Assumption 3.2.1. If $N(t) \leq r$, we set $X^{\star}_{i,N(t)} = 0$, for $i = N(t)+1, \ldots, r+1$. This means that in case there are less than $r+1$ claims, the reinsurer pays all the claims in the ECOMOR treaty.

Another modeling assumption is concerned with the way the reinsurance is affecting the capital position of the insurance company under consideration.

Assumption 3.2.2. We assume that at each time t, the currently applicable reinsured amount $R(t)$ is considered in the determination of the available surplus. In particular, this means that before the arrival of the $(r+1)$-st claim, the random retention in the ECOMOR treaty is considered to be zero. As a consequence in the ECOMOR treaty, the arrival of a new claim can lead to a modification of $R(t)$ of either sign, as the excess over the $(r+1)$-st claim may also decrease.

Note also that the setup we study here is that the duration of the reinsurance contract is T, and the implied premium for the reinsurance contract over the period $[0, T]$ is uniformly spread over this time interval. We will study the asymptotic behavior of the finite-time ruin probabilities (3.1.1) utilizing (3.1.2). Therefore, we formulate in the next section the large deviations problem that arises in our reinsurance context.

3.2.1 Large deviations in reinsurance

In [105], the large deviations results in (3.1.2) were derived in the Skorokhod J_1 topology. Correspondingly, we let $\mathbb{D} = \mathbb{D}([0,1], \mathbb{R})$ be a Skorokhod space, i.e. a space of real-valued càdlàg (right continuous with left limits) functions on $[0, 1]$, equipped with the J_1-metric defined by

$$d(\xi, \zeta) = \inf_{h \in \Lambda} \{\|h - id\| \vee \|\xi - \zeta \circ h\|\}, \qquad (\xi, \zeta) \in \mathbb{D}^2, \tag{3.2.3}$$

where Λ denotes the set of all strictly increasing continuous bijections from $[0, 1]$ to itself, id denotes the identity mapping, and $\|\cdot\|$ denotes the uniform (sup) norm on $[0, 1]$. Thus, A and C_j in (3.1.2) are a measurable set and a measure on $\mathbb{D}$, respectively. Furthermore, if $\phi : \mathbb{D} \to \mathbb{R}$ is a continuous functional on $\mathbb{D}$ and $B \in \mathcal{B}(\mathbb{R})$ is a Borel set such that $A = \phi^{-1}(B)$, where ϕ^{-1} stands for the inverse of ϕ, it holds that

$$\mathbf{P}\big(\phi(\bar{Y}_n) \in B\big) = \mathbf{P}\big(\bar{Y}_n \in \phi^{-1}(B)\big) = \mathbf{P}(\bar{Y}_n \in A). \tag{3.2.4}$$

The above relation portrays how it is possible to use the result (3.1.2) to study continuous functionals of $\bar{Y}_n$. To connect this to our ruin problem, we define $\bar{S}_n := \{\bar{S}_n(t), t \in [0, 1]\}$ as the centred and scaled process

$$\bar{S}_n(t) = \frac{1}{n} S(nt) - \lambda \mathbf{E} X t = \frac{1}{n} \sum_{i=1}^{N(nt)} X_i - \lambda \mathbf{E} X t, \quad t \geq 0. \tag{3.2.5}$$

Moreover, we assume that the capital u increases linearly in n, i.e. there exists an $a > 0$ such that $u = na$. We now formulate the large deviations problem to estimate the probabilities

$$\begin{aligned}
&\mathbf{P}\left(\sup_{t\in[0,1]} \{S(nt) - p_D nt - R(nt)\} \geq na\right) \\
&= \mathbf{P}\left(\sup_{t\in[0,1]} \{S(nt) - \lambda \mathbf{E} X nt - (p_D - \lambda \mathbf{E} X)nt - R(nt)\} \geq na\right) \\
&= \mathbf{P}\left(\sup_{t\in[0,1]} \{n\bar{S}_n(t) - cnt - R(nt)\} \geq na\right) \\
&= \mathbf{P}\left(\sup_{t\in[0,1]} \{\bar{S}_n(t) - ct - \frac{1}{n}R(nt)\} \geq a\right),
\end{aligned} \tag{3.2.6}$$

where $c = p_D - \lambda \mathbf{E} X$. As a next step, we must identify a continuous functional ϕ such that

$$\sup_{t\in[0,1]} \{\bar{S}_n(t) - ct - \frac{1}{n}R(nt)\} = \phi(\bar{S}_n), \tag{3.2.7}$$

so that we can write

$$\mathbf{P}\left(\sup_{t\in[0,1]} \{\bar{S}_n(t) - ct - \frac{1}{n}R(nt)\} \geq a\right) = \mathbf{P}\big(\phi(\bar{S}_n) \geq a\big) = \mathbf{P}\Big(\bar{S}_n \in \phi^{-1}\big([a,\infty)\big)\Big). \tag{3.2.8}$$

However, it is not immediately obvious from (3.2.7) what the functional ϕ looks like because $R(nt)$ is not expressed in terms of $\bar{S}_n$. We focus first on the LCR treaty and observe that

$$\begin{aligned}
\frac{1}{n}R(nt) = \frac{1}{n}L_r(nt) &= \frac{1}{n}\sum_{i=1}^{r} X^{\star}_{i,N(nt)} \\
&= \max_{\substack{(s_1,\dots,s_r)\in[0,t]^r \\ s_i \neq s_j, \forall i\neq j}} \sum_{i=1}^{r} \big(\bar{S}_n(s_i) - \bar{S}_n(s_i^-)\big), \quad t \in [0,1],
\end{aligned}$$

i.e. $L_r(nt)/n$ can be expressed as the sum of the r biggest jumps of the process $\bar{S}_n(t)$. For every $\xi \in \mathbb{D}$ and $m \in \mathbb{N}$, we define

$$\mathfrak{I}^m_\xi(t) = \sup_{\substack{(s_1,\dots,s_m)\in[0,t]^m \\ s_i \neq s_j, \forall i\neq j}} \sum_{i=1}^{m} \big(\xi(s_i) - \xi(s_i^-)\big)$$

$$= \max_{\substack{(s_1,\dots,s_m)\in[0,t]^m \\ s_i\neq s_j, \forall i\neq j}} \sum_{i=1}^{m} \big(\xi(s_i) - \xi(s_i^-)\big), \quad \text{for } t \in (0,1], \tag{3.2.9}$$

as the supremum of the sum of the m largest jumps of the function ξ. Naturally, $\mathfrak{J}_\xi^m(0) = 0$. Consequently, the functional ϕ we are looking for is a mapping $\phi_r : \mathbb{D} \to \mathbb{R}$ defined for every $\xi \in \mathbb{D}$ as

$$\phi_r(\xi) = \sup_{t\in[0,1]} \Big\{\xi(t) - ct - \mathfrak{J}_\xi^r(t)\Big\}. \tag{3.2.10}$$

Moreover, we denote the pre-image of $[a,\infty)$ under ϕ_r as $A_{c,a}^r = \phi_r^{-1}\big([a,\infty)\big)$ where

$$A_{c,a}^r = \left\{\xi \in \mathbb{D} : \sup_{t\in[0,1]} \Big\{\xi(t) - ct - \mathfrak{J}_\xi^r(t)\Big\} \geq a\right\}. \tag{3.2.11}$$

We next discuss the extension to ECOMOR treaties. By comparing (3.2.1) and (3.2.2), we observe that the relation between the reinsured amounts of the two treaties is

$$\begin{aligned} E_r(t) &= L_r(t) - rX_{r+1,N(t)}^\star = (r+1)L_r(t) - r\big(L_r(t) + X_{r+1,N(t)}^\star\big) \\ &= (r+1)L_r(t) - rL_{r+1}(t). \end{aligned}$$

Thus, in the ECOMOR treaty, the functional ϕ in (3.2.8) is the mapping $\varphi_r : \mathbb{D} \to \mathbb{R}$ defined for every $\xi \in \mathbb{D}$ as

$$\varphi_r(\xi) = \sup_{t\in[0,1]} \Big\{\xi(t) - ct - (r+1)\mathfrak{J}_\xi^r(t) + r\mathfrak{J}_\xi^{r+1}(t)\Big\}, \tag{3.2.12}$$

while the pre-image of $[a,\infty)$ under φ_r, i.e. $\mathcal{A}_{c,a}^r = \varphi_r^{-1}\big([a,\infty)\big)$, is defined as

$$\mathcal{A}_{c,a}^r = \left\{\xi \in \mathbb{D} : \sup_{t\in[0,1]} \Big\{\xi(t) - ct - (r+1)\mathfrak{J}_\xi^r(t) + r\mathfrak{J}_\xi^{r+1}(t)\Big\} \geq a\right\}. \tag{3.2.13}$$

3.2.2 Preliminaries on the Skorokhod topology and notation

Consider the complete metric space $\big(\mathbb{D}, d(,)\big)$. The functional $\mathfrak{J}_\xi^m(t)$ defined in (3.2.9) will play a significant role in the forthcoming analysis. Thus, it is important to confirm that it is well-defined. For this reason, let $\mathcal{D}(\xi)$ be the set of discontinuities of $\xi \in \mathbb{D}$, i.e.

$$\mathcal{D}(\xi) = \{t \in [0,1] : \xi(t^-) \neq \xi(t)\}, \tag{3.2.14}$$

and let $\mathcal{D}(\xi, \epsilon)$ be the set of discontinuities of magnitude at least ϵ, i.e.

$$\mathcal{D}(\xi, \epsilon) = \{t \in [0,1] : \left|\xi(t^-) - \xi(t)\right| \geq \epsilon\}. \tag{3.2.15}$$

The following result is standard.

Lemma 3.2.1 (Theorem 12.2.1 and Corollary 12.2.1 of [117])**.** *For any $\xi \in \mathbb{D}$ and $\epsilon > 0$, $\mathcal{D}(\xi, \epsilon)$ is a finite subset of $[0,1]$. In particular, $\mathcal{D}(\xi)$ is either finite or countably infinite.*

Consequently, the supremum in (3.2.9) is attained because only finitely many jumps can exceed a given positive number. As a result, $\mathfrak{J}_\xi^m(t)$ is well-defined.

Some important subspaces of $\mathbb{D}$ for our analysis are those restricted to step functions. We let $\mathbb{D}_{\mathcal{S}}^{\uparrow}$ be the set of all non-decreasing step functions vanishing at the origin. Furthermore, $\mathbb{D}_j$ is the subspace of $\mathbb{D}$ consisting of non-decreasing step functions, vanishing at the origin, with exactly j steps, and similarly, $\mathbb{D}_{\leqslant j} = \bigcup_{0 \leq i \leq j} \mathbb{D}_i$ consists of non-decreasing step functions, vanishing at the origin, with at most j steps. Finally, if $\mathcal{D}_+(\xi)$ denotes the number of discontinuities of $\xi \in \mathbb{D}$, we can then formally define the integer-valued set function $\mathcal{J}(A)$ appearing in (3.1.2) by

$$\mathcal{J}(A) = \inf_{\xi \in A \cap \mathbb{D}_{\mathcal{S}}^{\uparrow}} \mathcal{D}_+(\xi), \tag{3.2.16}$$

which we call the rate function. Observe that every $\xi \in \mathbb{D}_j$ is determined by the pair of jump sizes and jump times $(\boldsymbol{x}, \boldsymbol{u}) \in \mathbb{R}_+^j \times [0,1]^j$, i.e. $\xi(t) = \sum_{i=1}^j x_i \mathbb{1}_{\{u_i,1\}}(t)$, where $\mathbb{1}_B$ is the indicator function on the set B. For $\boldsymbol{x} = (x_1, \dots, x_j)$ and $\boldsymbol{u} = (u_1, \dots, u_j)$, we define the sets

$$\mathbb{R}_+^{j\downarrow} = \{\boldsymbol{x} \in \mathbb{R}_+^j : x_1 \geq x_2 \geq \cdots \geq x_j > 0\}, \tag{3.2.17}$$

and

$$S_j = \{(\boldsymbol{x}, \boldsymbol{u}) \in \mathbb{R}_+^{j\downarrow} \times (0,1)^j : u_1, \dots, u_j \text{ are all distinct}\}, \tag{3.2.18}$$

where the u_j's are not following the ordering of the x_j's, i.e. $x_k \geq x_l \not\Rightarrow u_k \geq u_l$. Thus, we can formally define the mapping $T_j : S_j \to \mathbb{D}_j$ by $T_j(\boldsymbol{x}, \boldsymbol{u}) = \sum_{i=1}^j x_i \mathbb{1}_{\{u_i,1\}}$.

Furthermore, let $\nu_\alpha(x,\infty) = x^{-\alpha}$ (i.e. the pure power decay part of the regularly varying claim sizes), and let ν_α^j denote the restriction to $\mathbb{R}_+^{j\downarrow}$ of the j-fold product measure of ν_α. We define for each $j \geq 1$ the measure C_j concentrated on $\mathbb{D}_j$ as

$$C_j(\bullet) = \mathbf{E}\big[\nu_\alpha^j\{\boldsymbol{y} \in \mathbb{R}_+^j : \sum_{i=1}^{j} y_i \mathbb{1}_{\{U_i,1\}} \in \bullet\}\big], \tag{3.2.19}$$

where the random variables U_i, $i = 1,\dots,j$, are i.i.d. uniform on $[0,1]$.

Finally, we say that a set $A \subseteq \mathbb{D}$ is bounded away from another set $B \subseteq \mathbb{D}$ if $\inf_{x\in A, y\in B} d(x,y) > 0$. Additionally, we let ${}_\delta A = \{\xi \in \mathbb{D} : d(\xi, A) \leq \delta\}$ for any $\delta > 0$.

3.3 Main result

Note that the parameter $c = p_D - \lambda \mathbf{E}X$ introduced in (3.2.1) can be either positive or negative. However, for $a \leq -c$, the rare event probability in (3.2.8) converges to one by the functional law of large numbers. For this reason, we focus only on the case $c + a > 0$. Letting

$${}_2F_1(b,e;d;z) = \sum_{k=0}^{+\infty} \frac{(b)_k (e)_k}{(d)_k} \frac{z^k}{k!}$$

be the hypergeometric function, with $(b)_k = b(b+1)\dots(b+k-1)$ denoting the Pochhammer symbol, we have the following theorem.

Theorem 3.3.1. *For $a > 0$, $c + a > 0$, and $r \in \mathbb{N}$, it holds that*

$$\psi(na, n) \sim \mathcal{C}_{r+1}\big(\lambda L(n)\big)^{r+1} n^{-(r+1)(\alpha-1)}, \qquad n \to \infty, \tag{3.3.1}$$

where

$$\mathcal{C}_{r+1} = \Bigg[a^{-(r+1)\alpha}\, {}_2F_1[r+1, (r+1)\alpha; r+2; -c/a] \cdot \mathbb{1}_{\{c>0\}} + (a+c)^{-(r+1)\alpha} \cdot \mathbb{1}_{\{c<0\}}\Bigg] \times \frac{1}{(r+1)!} \times \begin{cases} 1, & \textit{if } R(t) = L_r(t) \textit{ (LCR)}, \\ (r+1)^{(r+1)\alpha}, & \textit{if } R(t) = E_r(t) \textit{ (ECOMOR)}. \end{cases}$$

The proof of (3.3.1) is based on sample-path large-deviations results developed in [105]. Specifically, Theorems 3.1–3.2 in [105] provide the conditions under which the result (3.1.2) holds, and in addition the lim inf and lim sup are equal. Thus, to achieve our goal, we must verify that these conditions are satisfied for $\bar{Y}_n = \bar{S}_n$ and $A = A^r_{c,a}$ (LCR) or $A = \mathcal{A}^r_{c,a}$ (ECOMOR) defined in (3.2.11) and (3.2.13), respectively. However, their verification is rather involved. Hence, to make the proof of (3.3.1) more accessible, we split it in various steps after the aforementioned conditions and we provide additional explanations for each step.

Note that all of the forthcoming results are similar in the two treaties with possible deviations in small details. Therefore, we will first prove them for the LCR treaty and then show briefly how they can be extended to the ECOMOR treaty.

3.3.1 Proof of Theorem 3.3.1

The first step is to show that both mappings $\phi_r, \varphi_r : \mathbb{D} \to \mathbb{R}$ from Equations (3.2.10) and (3.2.12), respectively, are Lipschitz continuous. Due to their continuity, Equation (3.2.8) will hold and, consequently, we will be able to write $\mathbf{P}\big(\phi_r(\bar{S}_n) \geq a\big) = \mathbf{P}(\bar{S}_n \in A^r_{c,a})$ and $\mathbf{P}\big(\varphi_r(\bar{S}_n) \geq a\big) = \mathbf{P}(\bar{S}_n \in \mathcal{A}^r_{c,a})$. For this, we need the following intermediate result.

Lemma 3.3.1. *For every $(\xi, \zeta) \in \mathbb{D}^2$, $m \in \mathbb{N}$, and $h \in \Lambda$, it holds that*

$$\left|\mathfrak{I}^m_{\zeta\circ h}(t) - \mathfrak{I}^m_{\xi}(t)\right| \leq 2m\|\xi - \zeta \circ h\|, \qquad \forall t \in [0,1]. \tag{3.3.2}$$

Proof. By the definition of $\mathfrak{I}^m_{\zeta\circ h}(t)$, there exists $(\sigma_1, \ldots, \sigma_m) \in [0,t]^m$ with $\sigma_i \neq \sigma_j$ for all $i \neq j$, such that

$$\mathfrak{I}^m_{\zeta\circ h}(t) = \sum_{i=1}^{m} \big(\zeta \circ h(\sigma_i) - \zeta \circ h(\sigma_i^-)\big). \tag{3.3.3}$$

In addition, we have that

$$\mathfrak{I}^m_{\xi}(t) = \max_{\substack{(s_1,\ldots,s_m)\in[0,t]^m \\ s_i\neq s_j, \forall i\neq j}} \sum_{i=1}^{m} \big(\xi(s_i) - \xi(s_i^-)\big) \geq \sum_{i=1}^{m} \big(\xi(\sigma_i) - \xi(\sigma_i^-)\big). \tag{3.3.4}$$

Subtracting now Equations (3.3.3) and (3.3.4), we obtain

$$\mathfrak{I}^m_{\zeta\circ h}(t) - \mathfrak{I}^m_{\xi}(t) \leq \sum_{i=1}^{m} \big(\zeta \circ h(\sigma_i) - \zeta \circ h(\sigma_i^-) - \xi(\sigma_i) + \xi(\sigma_i^-)\big)$$

$$\leq \sum_{i=1}^{m} \left(|\zeta \circ h(\sigma_i) - \xi(\sigma_i)| + \left|\zeta \circ h(\sigma_i^-) - \xi(\sigma_i^-)\right| \right)$$
$$\leq 2m\|\xi - \zeta \circ h\|.$$

Following similar arguments, we can also show that $\mathfrak{J}_\xi^m(t) - \mathfrak{J}_{\zeta\circ h}^m(t) \leq 2m\|\xi - \zeta \circ h\|$, which completes the proof. □

We are now ready to establish the desired continuity.

Lemma 3.3.2 (Lipschitz continuity of the mapping)**.** *The mappings $\phi_r, \varphi_r : \mathbb{D} \to \mathbb{R}$ defined by* (3.2.10) *and* (3.2.12)*, respectively, are Lipschitz continuous w.r.t. J_1. More precisely, there exist $K \in [0, |c| + 2r + 1]$ and $L \in [0, |c| + 4r^2 + 4r + 1]$ such that $|\phi_r(\xi) - \phi_r(\zeta)| \leq Kd(\xi, \zeta)$ and $|\varphi_r(\xi) - \varphi_r(\zeta)| \leq Ld(\xi, \zeta)$, for all $(\xi, \zeta) \in \mathbb{D}^2$.*

Proof. W.l.o.g. we assume that $\phi_r(\xi) \geq \phi_r(\zeta)$, otherwise we switch the roles of ξ and ζ. For every $\epsilon > 0$, there exists $t_* \in [0, 1]$ such that

$$\xi(t_*) - ct_* - \mathfrak{J}_\xi^r(t_*) > \phi_r(\xi) - \epsilon. \tag{3.3.5}$$

On the other hand, by the definition of J_1, there exists $h = h(\xi, \zeta, \epsilon) \in \Lambda$ so that

$$d(\xi, \zeta) + \epsilon = \|h - id\| \vee \|\xi - \zeta \circ h\| \geq \big(h(t_*) - t_*\big) \vee \big(\xi(t_*) - \zeta \circ h(t_*)\big). \tag{3.3.6}$$

Furthermore, using the fact that h is a homeomorphism on $[0, 1]$, we obtain

$$\begin{aligned}
&\zeta \circ h(t_*) - ch(t_*) - \mathfrak{J}_{\zeta\circ h}^r(t_*) \\
&= \zeta \circ h(t_*) - ch(t_*) - \max_{\substack{(s_1,\ldots,s_r)\in[0,t_*]^r \\ s_i\neq s_j,\forall i\neq j}} \sum_{i=1}^{r} \left(\zeta \circ h(s_i) - \zeta \circ h(s_i^-)\right) \\
&= \zeta\big(h(t_*)\big) - ch(t_*) - \max_{\substack{(s_1,\ldots,s_r)\in[0,h(t_*)]^r \\ s_i\neq s_j,\forall i\neq j}} \sum_{i=1}^{r} \left(\zeta(s_i) - \zeta(s_i^-)\right) \\
&= \zeta\big(h(t_*)\big) - ch(t_*) - \mathfrak{J}_\zeta^r(h(t_*)) \leq \phi_r(\zeta).
\end{aligned} \tag{3.3.7}$$

Subtracting (3.3.7) from (3.3.5) yields

$$\begin{aligned}
\phi_r(\xi) - \phi_r(\zeta) &< \epsilon + \big(\xi(t_*) - \zeta \circ h(t_*)\big) + c\big(h(t_*) - t_*\big) + \big(\mathfrak{J}_{\zeta\circ h}^r(t_*) - \mathfrak{J}_\xi^r(t_*)\big) \\
&< \epsilon + \big(d(\xi, \zeta) + \epsilon\big) + |c| \,\big(d(\xi, \zeta) + \epsilon\big) + 2r\big(d(\xi, \zeta) + \epsilon\big) \\
&= (2 + |c| + 2r)\epsilon + (1 + |c| + 2r)d(\xi, \zeta),
\end{aligned}$$

where we have also used (3.3.6) and $\mathfrak{J}^r_{\zeta\circ h}(t_*) - \mathfrak{J}^r_\xi(t_*) \leq 2r\|\xi - \zeta \circ h\|$ by applying Lemma 3.3.1 with $t = t_*$ and $m = r$. Letting $\epsilon \to 0$, we conclude that $\phi_r(\xi) - \phi_r(\zeta) \leq (1 + |c| + 2r)d(\xi, \zeta)$, i.e. ϕ_r is Lipschitz continuous. The Lipschitz continuity for the φ_r mapping can be shown in an analogous manner. More precisely, for every $\epsilon > 0$, there exists $t_* \in [0, 1]$ such that

$$\xi(t_*) - ct_* - (r+1)\mathfrak{J}^r_\xi(t_*) + r\mathfrak{J}^{r+1}_\xi(t_*) > \varphi_r(\xi) - \epsilon. \tag{3.3.8}$$

For a homeomorphism h on $[0, 1]$ satisfying (3.3.6), we have

$$\begin{aligned}&\zeta \circ h(t_*) - ch(t_*) - (r+1)\mathfrak{J}^r_{\zeta\circ h}(t_*) + r\mathfrak{J}^{r+1}_{\zeta\circ h}(t_*)\\ &\quad = \zeta\big(h(t_*)\big) - ch(t_*) - (r+1)\mathfrak{J}^r_\zeta(h(t_*)) + r\mathfrak{J}^{r+1}_\zeta(h(t_*)) \leq \varphi_r(\zeta).\end{aligned} \tag{3.3.9}$$

We assume now w.l.o.g. that $\varphi_r(\xi) \geq \varphi_r(\zeta)$ and we subtract (3.3.9) from (3.3.8) to obtain

$$\begin{aligned}\varphi_r(\xi) - \varphi_r(\zeta) &< \epsilon + \big(\xi(t_*) - \zeta \circ h(t_*)\big) + c\big(h(t_*) - t_*\big)\\ &\qquad + (r+1)\big(\mathfrak{J}^r_{\zeta\circ h}(t_*) - \mathfrak{J}^r_\xi(t_*)\big) + r\big(\mathfrak{J}^{r+1}_\xi(t_*) - \mathfrak{J}^{r+1}_{\zeta\circ h}(t_*)\big)\\ &< \ \big(2 + |c| + 4r(r+1)\big)\epsilon + \big(1 + |c| + 4r(r+1)\big)d(\xi, \zeta),\end{aligned}$$

where we have also used (3.3.6) and Lemma 3.3.1 twice with $m = r, r+1$ and $t = t_*$. Letting $\epsilon \to 0$, the result is immediate. □

As a next step, we calculate the rate functions $\mathcal{J}(A^r_{c,a})$ and $\mathcal{J}(\mathcal{A}^r_{c,a})$ that appear in (3.1.2) and are formally defined in (3.2.16). We set for simplicity $c_+ = \max\{0, c\}$ and $c_- = \max\{0, -c\}$.

Lemma 3.3.3 (Evaluation of the rate function)**.** *The rate function defined by* (3.2.16) *is equal to* $r+1$ *in both treaties, i.e.*

$$\mathcal{J}(A^r_{c,a}) = \mathcal{J}(\mathcal{A}^r_{c,a}) = r + 1.$$

Proof. We need to show first that $\mathcal{J}(A^r_{c,a})$ cannot take any value smaller than or equal to r. Let us assume on the contrary that $\xi \in A^r_{c,a} \cap \mathbb{D}^\uparrow_{\mathcal{S}}$ such that $\mathcal{D}_+(\xi) = k \leq r$. This means that $\xi = \sum_{i\leq k} x_i \mathbb{1}_{\{u_i,1\}}$, with $x_1 \geq x_2 \geq \dots x_k > 0$ and $\{0, u_1, u_2, \dots, u_k, 1\}$ all distinct. By taking into account Assumptions 3.2.1, 3.2.2, we calculate

$$\phi_r(\xi) = \sup_{t\in[0,1]} \Big\{\xi(t) - ct - \mathfrak{J}^r_\xi(t)\Big\}$$

$$= \sup_{t\in[0,1]} \left\{ \sum_{i=1}^{k} x_i \mathbb{1}_{\{u_i,1\}}(t) - ct - \sum_{i=1}^{k} x_i \mathbb{1}_{\{u_i,1\}}(t) \right\} = c_-,$$

which states that $\xi \notin A^r_{c,a}$ because $\phi_r(\xi) = c_- \not\geq a$. As a result, $\mathcal{J}(A^r_{c,a}) \neq k$, for all $k \leq r$.

Let us assume now that $\xi \in A^r_{c,a} \cap \mathbb{D}^{\uparrow}_{\mathcal{S}}$ such that $\mathcal{D}_+(\xi) = r+1$, i.e. $\xi = \sum_{i=1}^{r+1} x_i \mathbb{1}_{\{u_i,1\}}$, with $x_1 \geq x_2 \geq \dots x_{r+1} > 0$ and $\{0, u_1, u_2, \dots, u_{r+1}, 1\}$ all distinct. To calculate $\phi_r(\xi)$, observe first that

$$\xi(t) - \mathfrak{I}^r_\xi(t) = \sum_{i=1}^{r+1} x_i \mathbb{1}_{\{u_i,1\}}(t) - \mathfrak{I}^r_\xi(t) = \begin{cases} 0, & t < \max\{u_1, \dots, u_{r+1}\} \\ x_{r+1}, & t \geq \max\{u_1, \dots, u_{r+1}\} \end{cases}, \tag{3.3.10}$$

because all the claims are "absorbed" according to Assumption 3.2.2 before the arrival of the $(r+1)$st claim, which happens at time $t^* = \max\{u_1, \dots, u_{r+1}\}$. Thus, we can write

$$\begin{aligned}\phi_r(\xi) &= \sup_{t\in[0,1]} \left\{ \xi(t) - ct - \mathfrak{I}^r_\xi(t) \right\} = \sup_{t\in[0,1]} \left\{ x_{r+1} \prod_{i=1}^{r+1} \mathbb{1}_{\{u_i,1\}}(t) - ct \right\} \\ &= x_{r+1} - c_+ \max\{u_1, \dots, u_{r+1}\} + c_-,\end{aligned}$$

since $x_{r+1} \prod_{i=1}^{r+1} \mathbb{1}_{\{u_i,1\}}(t)$ remains fixed at the value x_{r+1} from

$$t^* = \max\{u_1, \dots, u_{r+1}\}$$

onwards, while $-ct$ decreases or increases depending on the value of c. Due to the fact that $\xi \in A^r_{c,a}$, we get

$$\phi_r(\xi) \geq a \Rightarrow x_{r+1} \geq a + c_+ \max\{u_1, \dots, u_{r+1}\} - c_- \geq a - c_- > 0,$$

i.e. $A^r_{c,a} \cap \mathbb{D}^{\uparrow}_{\mathcal{S}} \neq \emptyset$ but contains all step functions with $r+1$ steps such that the $(r+1)$st largest step satisfies: $x_{r+1} \geq a + c_+ \max\{u_1, \dots, u_{r+1}\} - c_-$. Thus, $\mathcal{J}(A^r_{c,a}) = r+1$.

The proof for $\mathcal{J}(\mathcal{A}^r_{c,a}) = r+1$ in the ECOMOR treaty is similar. More precisely, it can easily be shown that there does not exist $\xi \in \mathcal{A}^r_{c,a} \cap \mathbb{D}^{\uparrow}_{\mathcal{S}}$ with $\mathcal{D}_+(\xi) = k \leq r$. Consequently, $\mathcal{J}(\mathcal{A}^r_{c,a}) \neq k$, $k \leq r$. Let us assume next that $\xi \in \mathcal{A}^r_{c,a} \cap \mathbb{D}^{\uparrow}_{\mathcal{S}}$ such that $\mathcal{D}_+(\xi) = r+1$, i.e. $\xi = \sum_{i=1}^{r+1} x_i \mathbb{1}_{\{u_i,1\}}$, with $x_1 \geq x_2 \geq \dots x_{r+1} > 0$ and $\{0, u_1, u_2, \dots, u_{r+1}, 1\}$ all distinct. It holds that

$$r\mathfrak{I}^{r+1}_\xi(t) - (r+1)\mathfrak{I}^r_\xi(t) = -\mathfrak{I}^r_\xi(t) + \begin{cases} 0, & t < \max\{u_1, \dots, u_{r+1}\} \\ rx_{r+1}, & t \geq \max\{u_1, \dots, u_{r+1}\} \end{cases}, \tag{3.3.11}$$

due to Assumption 3.2.1. By combining (3.3.10) and (3.3.11), we calculate

$$\begin{aligned}\varphi_r(\xi) &= \sup_{t\in[0,1]}\Big\{\xi(t) - ct - (r+1)\mathfrak{J}^r_\xi(t) + r\mathfrak{J}^{r+1}_\xi(t)\Big\}\\ &= \sup_{t\in[0,1]}\Big\{(r+1)x_{r+1}\prod_{i=1}^{r+1}\mathbb{1}_{\{u_i,1\}}(t) - ct\Big\}\\ &= (r+1)x_{r+1} - c_+\max\{u_1,\dots,u_{r+1}\} + c_-.\end{aligned}$$

Since $\xi \in \mathcal{A}^r_{c,a}$, we get $\varphi_r(\xi) \geq a \Rightarrow (r+1)x_{r+1} \geq a + c_+\max\{u_1,\dots,u_{r+1}\} - c_-$, i.e. $\mathcal{A}^r_{c,a} \cap \mathbb{D}^{\uparrow}_{\mathcal{S}} \neq \emptyset$ but contains all step functions with $r+1$ steps such that the $(r+1)$st largest step satisfies

$$x_{r+1} \geq \big(a + c_+\max\{u_1,\dots,u_{r+1}\} - c_-\big)/(r+1).$$

Thus, $\mathcal{J}(\mathcal{A}^r_{c,a}) = r+1$, and the proof is complete. □

Remark 3.1. The above lemma not only gives the value of the rate function, but it also provides the form of the minimal ξ that belongs to the sets $A^r_{c,a}$ and $\mathcal{A}^r_{c,a}$, i.e. all step functions with $r+1$ steps such that their $(r+1)$st greatest step is greater than or equal to the value $a + c_+\max\{u_1,\dots,u_{r+1}\} - c_-$ in the LCR treaty and the value $(a + c_+\max\{u_1,\dots,u_{r+1}\} - c_-)/(r+1)$ in the ECOMOR treaty.

An essential condition of Theorem 3.2 in [105] is that the sets ${}_{\delta}A^r_{c,a} \cap \mathbb{D}_{\leqslant \mathcal{J}(A^r_{c,a})}$ and ${}_{\delta}\mathcal{A}^r_{c,a} \cap \mathbb{D}_{\leqslant \mathcal{J}(\mathcal{A}^r_{c,a})}$ are bounded away from $\mathbb{D}_{\leqslant \mathcal{J}(A^r_{c,a})-1}$ and $\mathbb{D}_{\leqslant \mathcal{J}(\mathcal{A}^r_{c,a})-1}$, respectively. Verifying this condition allows us then to derive the result (3.1.2) for both treaties. We can directly use the value of the rate function in the following result due to Lemma 3.3.3.

Lemma 3.3.4 (Bounded away property)**.** *The sets* ${}_{\delta}A^r_{c,a} \cap \mathbb{D}_{\leqslant r+1}$ *and* ${}_{\delta}\mathcal{A}^r_{c,a} \cap \mathbb{D}_{\leqslant r+1}$ *are bounded away from* $\mathbb{D}_{\leqslant r}$ *for some* $\delta > 0$*.*

Proof. To simplify the notation in the proof, we write A instead of $A^r_{c,a}$ and $\mathcal{A}$ instead of $\mathcal{A}^r_{c,a}$, while the notation ${}_{\delta}A$, ${}_{\delta}\mathcal{A}$ follows naturally.

We start by showing that ${}_{\delta}A \cap \mathbb{D}_{\leqslant r+1}$ is bounded away from $\mathbb{D}_{\leqslant r}$ for some $\delta > 0$. Thanks to Lemma 3.3.1, we have that ${}_{\delta}A \subset A(\delta)$, where $A(\delta) = \phi_r^{-1}\big([a - (|c| + 2r + 1)\delta, \infty)\big)$. Hence, it suffices to show that $A(\delta) \cap \mathbb{D}_{\leqslant r+1}$ is bounded away from $\mathbb{D}_{\leqslant r}$. Let $\xi \in A(\delta) \cap \mathbb{D}_{\leqslant r+1}$. Since $\xi \in \mathbb{D}_{\leqslant r+1}$, we can write $\xi = \sum_{i=1}^{r+1} x_i \mathbb{1}_{\{u_i,1\}}$ with $x_1 \geq x_2 \geq \cdots \geq x_{r+1} \geq 0$, for which it holds that

$\phi_r(\xi) \le x_{r+1} - c_+ \max\{u_1,\dots,u_{r+1}\} + c_- \le x_{r+1} + c_-$ according to the proof of Lemma 3.3.3. Furthermore, $\xi \in A(\delta) \Leftrightarrow \phi_r(\xi) \ge a - (|c| + 2r + 1)\delta$. Combining the two inequalities, we obtain that $x_{r+1} \ge (a-c_-) - (|c|+2r+1)\delta \ge (a-c_-)/2$, for $\delta \le (a-c_-)/2(|c|+2r+1)$. In other words, for $\delta \le (a-c_-)/2(|c|+2r+1)$, $\xi \in A(\delta) \cap \mathbb{D}_{r+1} \subset A(\delta) \cap \mathbb{D}_{\leqslant r+1}$ with jump sizes bounded from below by $(a-c_-)/2$, which implies that $A(\delta) \cap \mathbb{D}_{\leqslant r+1}$ is bounded away from $\mathbb{D}_{\leqslant r}$.

In a similar manner, it suffices to show that $\mathcal{A}(\delta) \cap \mathbb{D}_{\leqslant r+1}$ is bounded away from $\mathbb{D}_{\leqslant r}$, where $\mathcal{A}(\delta) = \varphi_r^{-1}\big([a-(|c|+4r^2+4r+1)\delta, \infty)\big)$. Let $\xi \in \mathcal{A}(\delta) \cap \mathbb{D}_{\leqslant r+1}$. Since $\xi \in \mathbb{D}_{\leqslant r+1}$, we can write $\xi = \sum_{i=1}^{r+1} x_i \mathbb{1}_{\{u_i,1\}}$ with $x_1 \ge x_2 \ge \cdots \ge x_{r+1} \ge 0$, for which it holds that $\varphi_r(\xi) \le (r+1)x_{r+1} - c_+ \max\{u_1,\dots,u_{r+1}\} + c_- \le (r+1)x_{r+1} + c_-$. Furthermore, $\xi \in \mathcal{A}(\delta) \Leftrightarrow \varphi_r(\xi) \ge a - (|c| + 4r^2 + 4r + 1)\delta$. Combining the two inequalities, we obtain that $(r+1)x_{r+1} \ge (a-c_-) - (|c|+4r^2+4r+1)\delta \ge (a-c_-)/2$, for $\delta \le (a-c_-)/2(|c|+4r^2+4r+1)$. In other words, the jump sizes of ξ are bounded from below by $(a-c_-)/2(r+1)$, which implies that $\mathcal{A}(\delta) \cap \mathbb{D}_{\leqslant r+1}$ is bounded away from $\mathbb{D}_{\leqslant r}$ for $\delta \le (a-c_-)/2(|c|+4r^2+4r+1)$, and the proof is complete. $\square$

Let $\mathcal{C}^L_{r+1} := C_{r+1}(A^r_{c,a})$ and $\mathcal{C}^E_{r+1} := C_{r+1}(\mathcal{A}^r_{c,a})$. According to Section 3.1 in [105], the lim inf and lim sup in (3.1.2) yield the same result when

$$C_{\mathcal{J}(A)}(A^\circ) = C_{\mathcal{J}(A)}(A) = C_{\mathcal{J}(A)}(A^-).$$

However, the above equality holds when the set A is $C_{\mathcal{J}(A)}$-continuous, i.e. $C_{\mathcal{J}(A)}(\partial A) = 0$, where the boundary $\partial A = A^- \setminus A^\circ$ of a set A is the closure of A without its interior. We prove in the next lemma that the sets $A^r_{c,a}$ and $\mathcal{A}^r_{c,a}$ are both C_{r+1}-continuous.

Lemma 3.3.5 (Equality of the limits)**.** *The sets $A^r_{c,a}$ and $\mathcal{A}^r_{c,a}$ are C_{r+1}-continuous, i.e. $C_{r+1}(\partial A^r_{c,a}) = C_{r+1}(\partial \mathcal{A}^r_{c,a}) = 0$.*

Proof. To simplify the notation in the proof, we write again A instead of $A^r_{c,a}$ and $\mathcal{A}$ instead of $\mathcal{A}^r_{c,a}$, while the notation A°, $\mathcal{A}^\circ$, A^-, $\mathcal{A}^-$ follows naturally.

We start by showing the C_{r+1}-continuity of A. In line with the notation introduced in Section 3.2.2, we consider the function $T^{-1}_{r+1} : \mathbb{D}_{r+1} \to S_{r+1}$ such that

$$\begin{aligned}
T^{-1}_{r+1}\left(A^\circ\right) &= T^{-1}_{r+1}\left(\phi_r^{-1}\big((a,\infty)\big)\right) \\
&= \Big\{(\boldsymbol{x},\boldsymbol{u}) \in S_{r+1} : x_{r+1} > a + c_+ \max\{u_1,\dots,u_{r+1}\} - c_-\Big\}, \\
T^{-1}_{r+1}\left(A^-\right) &= T^{-1}_{r+1}\left(\phi_r^{-1}\big([a,\infty)\big)\right)
\end{aligned}$$

$$= \Big\{(\boldsymbol{x}, \boldsymbol{u}) \in S_{r+1} : x_{r+1} \geq a + c_+ \max\{u_1, \ldots, u_{r+1}\} - c_-\Big\}.$$

Obviously, the set $T_{r+1}^{-1}(A^-) \setminus T_{r+1}^{-1}(A^\circ)$ has zero Lebesgue measure. Combining this with $A^\circ \subseteq A \subseteq A^-$ and ϕ_r being a continuous function, we conclude that $C_{r+1}(\partial A) = 0$, i.e. A is C_{r+1}-continuous. To prove the C_{r+1}-continuity of $\mathcal{A}$, it suffices to observe that the set $T_{r+1}^{-1}(\mathcal{A}^-) \setminus T_{r+1}^{-1}(\mathcal{A}^\circ)$ has zero Lebesgue measure, where

$$\begin{aligned}
&T_{r+1}^{-1}(\mathcal{A}^\circ) \\
&= \Big\{(\boldsymbol{x}, \boldsymbol{u}) \in S_{r+1} : x_{r+1} > \big(a + c_+ \max\{u_1, \ldots, u_{r+1}\} - c_-\big)/(r+1)\Big\}, \\
&T_{r+1}^{-1}(\mathcal{A}^-) \\
&= \Big\{(\boldsymbol{x}, \boldsymbol{u}) \in S_{r+1} : x_{r+1} \geq \big(a + c_+ \max\{u_1, \ldots, u_{r+1}\} - c_-\big)/(r+1)\Big\},
\end{aligned}$$

which follows by the same reasoning. □

We calculate now the pre-constants $C_{\mathcal{J}(A^r_{c,a})}(A^r_{c,a})$ and $C_{\mathcal{J}(\mathcal{A}^r_{c,a})}(\mathcal{A}^r_{c,a})$.

Lemma 3.3.6 (Calculation of the pre-constant)**.** *The constants $\mathcal{C}^L_{r+1}$ and $\mathcal{C}^E_{r+1}$ are given by*

$$\mathcal{C}^L_{r+1} = \frac{1}{(r+1)!} \times \begin{cases} a^{-(r+1)\alpha} \cdot {}_2F_1[r+1, (r+1)\alpha; r+2; -c/a], & c > 0, \\ (a+c)^{-(r+1)\alpha}, & c < 0. \end{cases}$$

$$\mathcal{C}^E_{r+1} = \frac{(r+1)^{(r+1)\alpha}}{(r+1)!} \times \begin{cases} a^{-(r+1)\alpha} \cdot {}_2F_1[r+1, (r+1)\alpha; r+2; -c/a], & c > 0, \\ (a+c)^{-(r+1)\alpha}, & c < 0. \end{cases}$$

Proof. Recall that $\mathcal{C}^L_{r+1} := C_{r+1}(A^r_{c,a})$ and $\mathcal{C}^E_{r+1} := C_{r+1}(\mathcal{A}^r_{c,a})$. To calculate these constants, we use the definition of the measure $C_{r+1}(\bullet)$ in (3.2.19). We start with $\mathcal{C}^L_{r+1}$. It is known that for $U_1, \ldots, U_{r+1} \sim \mathcal{U}(0,1)$, the distribution of the r.v. $\max\{U_1, \ldots, U_{r+1}\}$ is given by the formula $\mathbf{P}(\max\{U_1, \ldots, U_{r+1} \leq t) = t^{r+1}$. Furthermore, by using that $\int_b^{+\infty} \alpha y^{-n\alpha-1} dy = b^{-n\alpha}/n$ with $b > 0$, we recursively calculate the following multiple integrals for $n \in \mathbb{N}$ and positive y_i's

$$\mathcal{I}_n = \int\limits_{y_1 \geq \cdots \geq y_{n+1}} \prod_{i=1}^{n} \alpha y_i^{-\alpha-1} dy_1 \ldots dy_n$$

$$= \int_{y_n=y_{n+1}}^{+\infty} \int_{y_{n-1}=y_n}^{+\infty} \cdots \int_{y_2=y_3}^{+\infty} \prod_{i=2}^{n} \alpha y_i^{-\alpha-1} \underbrace{\left(\int_{y_1=y_2}^{+\infty} \alpha y_1^{-\alpha-1} dy_1 \right)}_{=y_2^{-\alpha}} dy_2 \dots dy_n$$

$$= \int_{y_n=y_{n+1}}^{+\infty} \cdots \int_{y_3=y_4}^{+\infty} \prod_{i=3}^{n} \alpha y_i^{-\alpha-1} \underbrace{\left(\int_{y_2=y_3}^{+\infty} \alpha y_2^{-2\alpha-1} dy_2 \right)}_{=y_3^{-2\alpha}/2} dy_3 \dots dy_n$$

$$= \cdots = \frac{1}{n!} (y_{n+1})^{-n\alpha}.$$

Consequently, in case $c > 0$, we obtain by virtue of Remark 3.1

$$\begin{aligned}
\mathcal{C}_{r+1}^L &= \mathbf{E}\big[\nu_\alpha^{r+1}\{\boldsymbol{y} \in \mathbb{R}_+^{r+1} : \sum_{i=1}^{r+1} y_i \mathbb{1}_{\{U_i,1\}} \in A_{c,a}^r\}\big] \\
&= \mathbf{E}\left[\int_{y_1 \geq \cdots \geq y_{r+1} > 0} \prod_{i=1}^{r+1} \alpha y_i^{-\alpha-1} \mathbb{1}_{\{y_{r+1} \geq a + c \max\{U_1, \dots, U_{r+1}\}\}} dy_1 \dots dy_{r+1} \right] \\
&= \int_{t \in [0,1]} \int_{y_1 \geq \cdots \geq y_{r+1} > 0} \prod_{i=1}^{r+1} \alpha y_i^{-\alpha-1} \mathbb{1}_{\{y_{r+1} \geq a + ct\}} (r+1) t^r dy_1 \dots dy_{r+1} dt \\
&= \int_{t \in [0,1]} \int_{y_{r+1} > 0} \mathcal{I}_r \alpha y_{r+1}^{-\alpha-1} \mathbb{1}_{\{y_{r+1} \geq a + ct\}} (r+1) t^r dy_{r+1} dt \\
&= \int_{t \in [0,1]} \int_{y_{r+1} = a + ct}^{+\infty} \frac{1}{r!} (y_{r+1})^{-r\alpha} \alpha y_{r+1}^{-\alpha-1} (r+1) t^r dy_{r+1} dt \\
&= \frac{r+1}{r!} \int_0^1 t^r \left(\int_{a+ct}^{+\infty} \alpha (y_{r+1})^{-(r+1)\alpha - 1} dy_{r+1} \right) dt = \frac{1}{r!} \int_0^1 t^r (a + ct)^{-(r+1)\alpha} dt \\
&= \frac{a^{-(r+1)\alpha}}{(r+1)!} \cdot {}_2F_1[r+1, (r+1)\alpha; r+2; -c/a].
\end{aligned}$$

Analogously, we find

$$\begin{aligned}
\mathcal{C}^E_{r+1} &= \mathbf{E}\big[\nu_\alpha^{r+1}\{\boldsymbol{y} \in \mathbb{R}_+^{r+1} : \sum_{i=1}^{r+1} y_i \mathbb{1}_{\{U_i,1\}} \in \mathcal{A}^r_{c,a}\}\big] \\
&= \int\limits_{y_1 \geq \cdots \geq y_{r+1} > 0} \prod_{i=1}^{r+1} \alpha y_i^{-\alpha-1} \mathbb{1}_{\{y_{r+1} \geq (a + c\max\{U_1,\ldots,U_{r+1}\})/(r+1)\}} dy_1 \ldots dy_{r+1} \\
&= \int\limits_{t\in[0,1]} \int\limits_{y_{r+1}>0} \mathcal{I}_r \alpha y_{r+1}^{-\alpha-1} \mathbb{1}_{\{y_{r+1} \geq (a+ct)/(r+1)\}} (r+1) t^r dy_{r+1} dt \\
&= \frac{1}{r!} \int\limits_0^1 t^r \left(\frac{a+ct}{r+1}\right)^{-(r+1)\alpha} dt = (r+1)^{(r+1)\alpha} \frac{1}{r!} \int\limits_0^1 t^r (a+ct)^{-(r+1)\alpha} dt \\
&= (r+1)^{(r+1)\alpha} \frac{a^{-(r+1)\alpha}}{(r+1)!} \cdot {}_2F_1[r+1, (r+1)\alpha; r+2; -c/a].
\end{aligned}$$

In case $c < 0$, the coefficients simplify to

$$\begin{aligned}
\mathcal{C}^L_{r+1} &= \mathbf{E}\left[\int\limits_{y_1 \geq \cdots \geq y_{r+1} > 0} \prod_{i=1}^{r+1} \alpha y_i^{-\alpha-1} \mathbb{1}_{\{y_{r+1} \geq a+c\}} dy_1 \ldots dy_{r+1}\right] \\
&= \int\limits_{a+c}^{+\infty} \mathcal{I}_r \alpha y_{r+1}^{-\alpha-1} dy_{r+1} = \ldots = \frac{1}{(r+1)!}(a+c)^{-(r+1)\alpha}, \quad \text{and} \\
\mathcal{C}^E_{r+1} &= (r+1)^{(r+1)\alpha} \frac{1}{(r+1)!}(a+c)^{-(r+1)\alpha}.
\end{aligned}$$

□

Remark 3.2. When $c > 0$, the coefficients $\mathcal{C}^L_{r+1}$ and $\mathcal{C}^E_{r+1}$ can be equivalently expressed in terms of finite sums involving the Gamma function. More precisely, by applying r times integration by parts, we calculate for $k > r+1$ that

$$\begin{aligned}
\frac{1}{r!} \int t^r (a+ct)^{-k} dt &= \sum_{m=1}^{r+1} \frac{(-1)^{m+1} t^{r+1-m}}{(r+1-m)!} \frac{(a+ct)^{m-k}}{c^m \prod_{j=1}^m (j-k)} \\
&= \sum_{m=1}^{r+1} \frac{(-1)^{m+1} t^{r+1-m}}{(r+1-m)!} \frac{(a+ct)^{m-k}}{c^m (1-k)_m}
\end{aligned}$$

$$= -\sum_{m=1}^{r+1} \frac{t^{r+1-m}}{(r+1-m)!} \frac{(a+ct)^{m-k}}{c^m (k-m)_m},$$

and hence,

$$\frac{1}{r!}\int_0^1 t^r (a+ct)^{-k} dt = \frac{a^{r+1-k}}{c^{r+1}(k-r-1)_{r+1}} - \sum_{m=1}^{r+1} \frac{(a+c)^{m-k}}{(r+1-m)! c^m (k-m)_m},$$

where $(b)_k = \Gamma(b+k)/\Gamma(b)$ is again the Pochhammer symbol. Thus,

$$\mathcal{C}^L_{r+1} = \frac{a^{-(r+1)(\alpha-1)}\Gamma\big((r+1)\alpha\big)}{c^{r+1}\Gamma\big((r+1)(\alpha-1)\big)} - \sum_{m=1}^{r+1} \frac{(a+c)^{m-(r+1)\alpha}\Gamma\big((r+1)\alpha\big)}{(r+1-m)! c^m \Gamma\big((r+1)\alpha - m\big)},$$

$$\mathcal{C}^E_{r+1} = (r+1)^{(r+1)\alpha}\left(\frac{a^{-(r+1)(\alpha-1)}\Gamma\big((r+1)\alpha\big)}{c^{r+1}\Gamma\big((r+1)(\alpha-1)\big)} - \sum_{m=1}^{r+1} \frac{(a+c)^{m-(r+1)\alpha}\Gamma\big((r+1)\alpha\big)}{(r+1-m)! c^m \Gamma\big((r+1)\alpha - m\big)}\right).$$

Remark 3.3. In the absence of reinsurance ($r = 0$), the pre-constant simplifies to

$$\frac{a^{-\alpha+1} - (a+c)^{-\alpha+1}}{c(\alpha-1)},$$

which can also be derived from existing results, see e.g. [6, 53].

Finally, we know from [84] that the compound Poisson aggregate claim process $S(t) = \sum_{i=1}^{N(t)} X_i$ is a special Lévy process with Lévy measure $\nu(dx) = \lambda F(dx)$, which means that $n \cdot \nu[n, \infty) = \lambda n \bar{F}(n) = \lambda L(n) n^{-\alpha+1}$, $n \in \mathbb{N}$. We conclude the proof of Theorem 3.3.1 by combining this result with Lemma 3.3.6 to obtain the expression (3.3.1).

3.4 Numerical implementations

Our primary goal in this section is to verify our asymptotic approximations in Theorem 3.3.1 via numerical illustration. For this purpose, we employ the importance sampling scheme developed in Chapter 2 that is proved to be strongly efficient in the current setting.

We use a shifted Pareto distribution for the claim sizes, i.e. $\bar{F}(x) = (x+1)^{-\alpha}$, $x \geq 0$, and $\mathbf{E}X = 1/(\alpha-1)$. In addition, we calculate the net premiums $p_D^L = p - p_R^L$ and $p_D^E = p - p_R^E$ of the insurer after purchasing an LCR or ECOMOR reinsurance for a premium p_R^L and p_R^E, respectively.

We assume here that the reinsurance premiums are determined according to an *expected value principle*, see e.g. [1]. Hence, we need to determine $\mathbf{E}R(t)$. As the Pareto claims arrive according to a Poisson process with rate λ, we follow [13] to obtain

$$
\begin{aligned}
\mathbf{E}L_r(t) &= (\lambda t)^{1/\alpha} \sum_{i=1}^{r} \frac{\gamma(i-1/\alpha, \lambda t)}{\Gamma(i)} - \sum_{i=1}^{r} \frac{\gamma(i, \lambda t)}{\Gamma(i)}, \\
\mathbf{E}E_r(t) &= (\lambda t)^{1/\alpha} \left(\sum_{i=1}^{r} \frac{\gamma(i-1/\alpha, \lambda t)}{\Gamma(i)} - r \frac{\gamma(r+1-1/\alpha, \lambda t)}{\Gamma(r+1)} \right) \\
&\quad - \left(\sum_{i=1}^{r} \frac{\gamma(i, \lambda t)}{\Gamma(i)} - r \frac{\gamma(r+1, \lambda t)}{\Gamma(r+1)} \right),
\end{aligned}
$$

where $\gamma(k,s) = \int_0^s e^{-u} u^{k-1} du$ is the lower incomplete gamma function. Thus, if $\theta, \eta > 0$ are the relative safety loadings imposed by the insurer and reinsurer, respectively, we calculate the annual retained premium p_D over a period of n years via the formula $p_D = (1+\theta)\mathbf{E}S(1) - (1+\eta)\mathbf{E}R(n)/n$. Correspondingly,

$$
\begin{aligned}
p_D^L =& \frac{\lambda(1+\theta)}{\alpha-1} - \frac{1+\eta}{n} \left((\lambda n)^{1/\alpha} \sum_{i=1}^{r} \frac{\gamma(i-1/\alpha, \lambda n)}{\Gamma(i)} - \sum_{i=1}^{r} \frac{\gamma(i, \lambda n)}{\Gamma(i)} \right), \\
p_D^E =& \frac{(1+\theta)\lambda}{\alpha-1} - \frac{(1+\eta)(\lambda n)^{1/\alpha}}{n} \left(\sum_{i=1}^{r} \frac{\gamma(i-1/\alpha, \lambda n)}{\Gamma(i)} - r \frac{\gamma(r+1-1/\alpha, \lambda n)}{\Gamma(r+1)} \right) \\
&+ \frac{1+\eta}{n} \left(\sum_{i=1}^{r} \frac{\gamma(i, \lambda n)}{\Gamma(i)} - r \frac{\gamma(r+1, \lambda n)}{\Gamma(r+1)} \right).
\end{aligned}
$$

We fix now $n = 20$, $\alpha = 1.5$, $\lambda = 10$, $\theta = 0.2$, $\eta = 0.3$ (safety loadings for reinsurance are typically larger than for primary insurance, see [1]) to obtain the following figures:

Finally, we choose the values of a such that the asymptotic approximations for LCR and ECOMOR are simultaneously defined. In other words, it should hold that $a > \max\{-c_L, -c_E, 0\}$, where $c_i = p_D^i - \lambda/(\alpha-1)$, $i \in \{L, E\}$. It is clear

r	p_R^L	p_R^E	p_D^L	p_D^E	c_L	c_E
0	0	0	24	24	4	4
1	4.5309	3.0539	18.1098	20.0299	-1.8902	0.0298
2	6.0078	4.0719	16.1897	18.7065	-3.8102	-1.2935
3	6.9758	4.7505	14.9314	17.8242	-5.0686	-2.1757

Table 3.1: Premiums for LCR and ECOMOR treaties for varying r for $n = 20$, $\lambda = 10$, $\alpha = 1.5$, $\theta = 0.2$, and $\eta = 0.3$.

from Table 3.1 that $c_L < c_E$. Therefore, both approximations are simultaneously valid for $a > \max\{-c_L, 0\}$.

The results under both LCR and ECOMOR treaties for different combinations of r and a are presented in Figures 3.1–3.3. We plot the simulation estimates (circles) together with the large deviation approximation (line) of the rare event probabilities as a function of n. Note that the results for $r = 0$ can be considered as a sanity check for our simulation study.

We observe that the large deviation results become accurate as n grows, in line with Theorem 3.3.1. It is quite remarkable that in most cases the resulting approximation is already excellent for $n = 20$. This corresponds to a time horizon of 20 years for the present insurance application. For fixed n, the quality of the asymptotic approximation improves as a increases. Finally, we recognize that LCR always leads to lower ruin probabilities than ECOMOR, which is intuitively expected. However, the explicit expression given in Theorem 3.3.1, allows for the first time to quantitatively assess the effects of the model parameters on the resulting ruin probabilities.

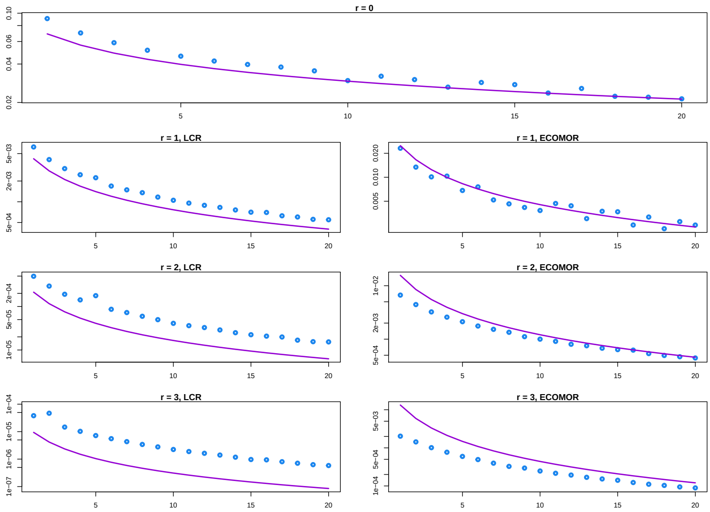

Figure 3.1: Numerical results for both LCR and ECOMOR treaties, for $a = 20$.

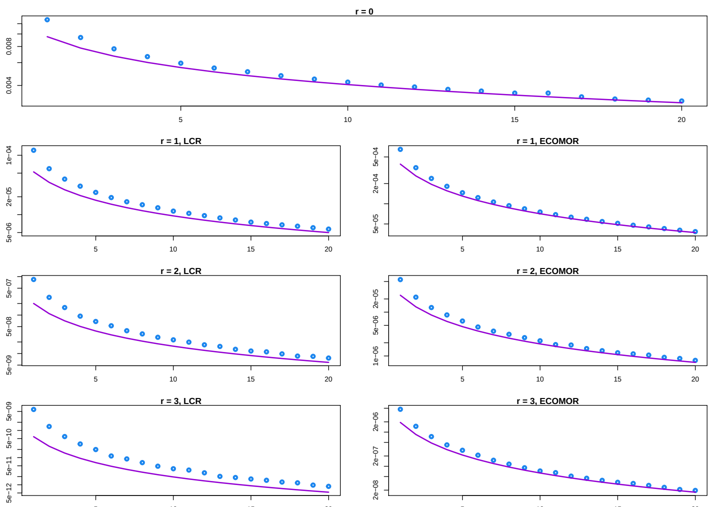

Figure 3.2: Numerical results for both LCR and ECOMOR treaties, for $a = 80$.

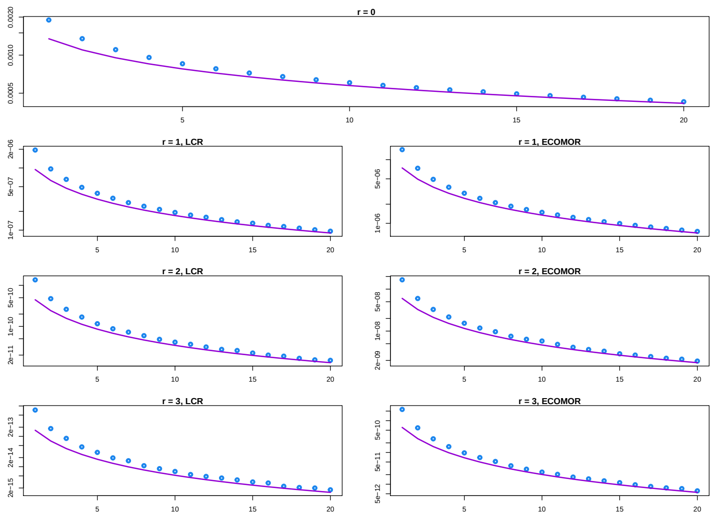

Figure 3.3: Numerical results for both LCR and ECOMOR treaties, for $a = 300$.

Chapter 4

Large deviations for Markov additive processes

4.1 Introduction

In this chapter, we develop sample-path large deviations for random walks with increments that are driven by an i.i.d. sequence of iterated random functions. To be precise, let X_n, $n \geq 0$, be such that $X_{n+1} = f_{n+1}(X_n)$, where f_n, $n \geq 1$, is sequence of i.i.d. random functions satisfying certain regularity conditions (see Assumption 4.3.1 below). Define $\bar{X}_n = \{\bar{X}_n(t), t \in [0,1]\}$, with

$$\bar{X}_n(t) = \sum_{i=0}^{\lfloor nt \rfloor - 1} X_i/n. \tag{4.1.1}$$

We are interested in large deviations of $\bar{X}_n$.

The starting point for our analysis is to consider the random difference equations, where

$$f_n(z) = A_n z + B_n \tag{4.1.2}$$

for a sequence of i.i.d. $\mathbb{R}^2$-valued random vectors (A_n, B_n). Such equations can be found, for example, in the context of ruin problems with investments, in the study of extremes of financial time series such as ARCH-type processes (see e.g. Section 8.4 of [54]), in tail asymptotics for exponential functionals of Lévy processes (see e.g. [88]), etc.

Here we consider a set of classical assumptions (see Assumption 4.2.1 below), which can be found in the Kesten-Goldie theorem (see [81] and [66]). Note that, under these assumptions, the Markov chain X_n, $n \geq 0$—regardless of its initial state X_0—has a unique stationary distribution π, for which we have

$$\pi(x, \infty) \sim c_+ x^{-\alpha} \quad \text{and} \quad \pi(-\infty, -x) \sim c_- x^{-\alpha}, \quad \text{as } x \to \infty, \tag{4.1.3}$$

for some c_-, c_+ satisfying $c_- + c_+ > 0$. We aim to develop asymptotic estimates of $\mathbf{P}(\bar{X}_n \in E)$ for a sufficiently general collection of sets E. In view of (4.1.3), it is natural to expect that our sample-path large deviations results should possess some heavy-tailed components. This conjecture turns out to be true. Next, we briefly describe our main results, as well as the methodologies that are used to derive the results.

To relate our problem with the existing theory of sample-path large deviations for stochastic processes, we first identify a sequence of regeneration times r_n, $n \geq 1$ (see [12]), and split the Markov chain into i.i.d. cycles. By aggregating the trajectory of $\bar{X}_n$ over regeneration cycles, we obtain a regenerative process with i.i.d. jump distributions and r_n, $n \geq 1$ as renewals. Under a set of assumptions originating from [81] and [66], we adopt a large deviation change of measure argument and further establish that the "area" under a typical regeneration cycle, denoted by $\mathfrak{R}$ (see (4.3.2) below), has an asymptotic power law. To be precise, we have

$$\mathbf{P}(\mathfrak{R} > x) \sim C_+ x^{-\alpha} \quad \text{and} \quad \mathbf{P}(\mathfrak{R} < -x) \sim C_- x^{-\alpha}, \quad \text{as } x \to \infty, \tag{4.1.4}$$

for some C_-, C_+ satisfying $C_- C_+ > 0$.

This approach brings us close to the framework studied in [105]. By analyzing the associated renewal process, we are able to utilize the idea from [105] and derive a sample-path large deviations result for the aggregated process under the Skorokhod J_1 metric. However, showing that the residual process is negligible in the sense of contributing to $\mathbf{P}(\bar{X}_n \in E)$ is not straightforward, especially when the increments of $\bar{X}_n$ are dependent in the current setting. To overcome this, we switch to a slightly weaker topology, the Skorokhod M_1'-topology (see Section 4.3.2 below), and derive asymptotic estimates of events involved with the "area" under the last ongoing cycle.

This paves the way for our main sample-path large deviations results, which are presented in Section 4.3. For the case where B_n as in (4.1.2) is nonnegative, our result establishes that

$$C_{\mathcal{J}^*}(E^\circ) \leq \varliminf_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in E)}{(n\mathbf{P}(\mathfrak{R} > n))^{\mathcal{J}^*}} \leq \varlimsup_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in E)}{(n\mathbf{P}(\mathfrak{R} > n))^{\mathcal{J}^*}} \leq C_{\mathcal{J}^*}(E^-). \tag{4.1.5}$$

Precise details can be found in Section 4.3.2 below. At this moment, we just mention that C_j is a measure on the Skorokhod space, and $\mathcal{J}^*$ denotes the minimum number of jumps that are required for a nondecreasing, piecewise linear function with drfit $\mathbf{E}B_1/(1-\mathbf{E}A_1)$ to be in the set E. In Sections 4.3.3 and 4.3.4, we consider two generalizations of the result presented in (4.1.5).

In order to contextualize our contribution, we provide a literature review. The investigation of tail estimates of one-dimensional random walks with heavy-tailed step size was initiated in [95], [96], where the author studied the sequence x_n for which

$$\mathbf{P}(\hat{S}_n/n > x_n) = n\mathbf{P}(\hat{S}_1 > x_n)(1 + o(1)), \quad \text{as } n \to \infty, \tag{4.1.6}$$

holds, where $\hat{S}_n$ is a random walk with i.i.d. heavy-tailed, not necessarily regularly varying, increments. For a detailed description of the large deviations regime we refer to e.g. [24], [43], and [60]. When (4.1.6) is valid, the so-called principle of a single big jump is said to hold. As a generalization of (4.1.6), a functional form has been derived in [75], where random walks with i.i.d. multi-dimensional regularly varying (cf. Definition 1.1 of [75]) step sizes are considered.

On the other hand, a significant number of studies try to answer the question of if and how the principle of a single big jump can be extended to the case where there is a certain dependence structure in the increments. Key references are [57], [74], [91, 92], [90], where stable processes, modulated processes, and stochastic differential equations are considered. As mentioned above, deriving large deviations results for autoregressive processes (cf. (4.1.2)) is one of the focus points of this chapter. Existing literature on this topic can be distinguished by the assumptions that need to be made. As a well-known insight, the regular variation in large deviations of autoregressive processes comes mainly from two types of sources. The first is by assuming that B_n is regularly varying, and A_n is sufficiently light-tailed. Regular variation occurs in this case due to the insight that the sum of independent regularly varying random variables is again regularly varying (cf. [83]). On the other hand, under assumptions as in [81, 66], a large value of the sample mean is not due to a single large value of the A_n or B_n but to large values of the products $A_1 \cdots A_n$ (cf. [35], [27]).

So far, all results are stated within the framework of the principle of a single big jump. However, not all rare events are caused by a single jump (for examples, see [29]). Various studies investigate rare events that are caused by multiple jumps using ad-hoc approaches, see [59], [120]. As mentioned in Chapter 1, [105] provides sample-path large deviations results for Lévy processes/random walks with regularly varying increments, which deal with a general class of rare events that can especially be caused by multiple jumps.

In view of the literature review above, our mathematical contribution can be described as follows. We start considering the case where $X_{n+1} = A_{n+1}X_n + B_{n+1}$. In Section 4.2, we first derive tail estimates on the "areas" under both the first return time and the regeneration cycle. The technique of splitting Markov chains into regeneration cycles can be found in many existing works, see e.g. [99], [18]. The one that comes closest to our analysis is [35], where the authors consider the AR(1) process as the underlying process and derive estimates on the right tail of the "area" under a regeneration cycle. In this chapter, we use a more delicate proof technique and provide estimates on both tails. Using the tail estimates in Theorem 4.3.1, we present in Sections 4.3.2 and 4.3.3 large deviations results for $\bar{X}_n$ as in (4.1.1). These results complement the approaches in [83], where the authors study the case of B_n being regularly varying, and hence, regular variation occurs due to large values in B_n. To establish Theorems 4.3.2 and 4.3.3, we introduce a notion of equivalence (see Lemma 4.2.3 below), which is proved to be very useful in connecting our problem with [105]. In Section 4.3.4, we extend our results to the case where $X_n = f_n(X_{n-1})$ for general f_n.

This chapter is organized as follows. In Section 4.2, we introduce some useful tools for future purposes. We present our main results in Section 4.3. In Section 4.4, we consider an example where our large deviations result is applicable. Sections 4.5–4.8 are devoted to additional proofs.

4.2 Preliminaries

In this chapter, we study the sample-path large deviations behavior of

$$\bar{X}_n = \left\{ \frac{1}{n} \sum_{i=0}^{\lfloor nt \rfloor - 1} X_i,\ t \in [0,1] \right\}, \tag{4.2.1}$$

where the underlying stochastic process $\{X_n\}_{n\geq 0}$ is defined by

$$X_{n+1} = A_{n+1}X_n + B_{n+1}, \quad n \geq 0, \tag{4.2.2}$$

and (A_n, B_n), $n \geq 1$, is a sequence of i.i.d. $\mathbb{R}^2$-valued random vectors, independent of X_0. Note that $\{X_n\}_{n\geq 0}$ is called a first-order autoregressive (AR(1)) process. We start with introducing a regularity condition.

Assumption 4.2.1. Assume that the following conditions hold.

1. $A_1 \geq 0$ a.s. and the law of $\log A_1$ conditioned on $\{A_1 > 0\}$ is nonarithmetic.

2. There exists an $\alpha \in (1, \infty)$ such that $\mathbf{E}A_1^\alpha = 1$, $\mathbf{E}|B_1|^\alpha < \infty$, and $\mathbf{E}A_1^\alpha \log^+ A_1 < \infty$, where $\log^+ x = \max\{\log x, 0\}$.

3. $\mathbf{P}(A_1 x + B_1 = x) < 1$ for every $x \in \mathbb{R}$.

Remark 4.1. The first two conditions in Assumption 4.2.1 imply that $\mathbf{E} \log A_1 < 0$ and $\mathbf{E} \log^+ |B_1| < \infty$, and hence (see e.g. Theorem 2.1.3 of [26]), the Markov chain has a unique stationary distribution, denoted by π. By assuming additionally the third condition, it was shown (see e.g. [81] and [66]) that

$$\pi(x, \infty) \sim c_+ x^{-\alpha} \quad \text{and} \quad \pi(-\infty, -x) \sim c_- x^{-\alpha}, \qquad \text{as } x \to \infty,$$

for some c_+, c_- satisfying $c_+ + c_- > 0$.

4.2.1 Background from Markov chain theory

In this section we review some preliminaries from Markov chain theory. We begin by introducing two conditions on general Markov chains.

We say that a Markov chain on some general state space $(\mathbb{S}, \mathcal{S})$ with transition kernel P satisfies a drift condition $(\mathcal{D})$ if

$$\int_{\mathbb{S}} h(y) P(x, dy) \le \gamma h(x) + \rho \mathbb{1}_{\mathcal{C}}(x), \qquad \text{for some } \gamma \in (0,1), \tag{$\mathcal{D}$}$$

where h is a function taking values in $[1, \infty)$, ρ is a positive constant, and $\mathcal{C}$ is a Borel subset of $\mathbb{R}$. Moreover, we say that a φ-irreducible Markov chain on $(\mathbb{S}, \mathcal{S})$ with transition kernel P satisfies a minorization condition $(\mathcal{M})$ if

$$\theta \mathbb{1}_{\mathcal{C}_0}(x) \phi(E \cap E_0) \le P(x, E), \qquad x \in \mathbb{S}, E \in \mathcal{S}, \tag{$\mathcal{M}$}$$

for some set $E_0 \subseteq \mathbb{S}$, some set $\mathcal{C}_0$ with $\varphi(\mathcal{C}_0) > 0$, some constant $\theta > 0$, and some probability measure ϕ on $(\mathbb{S}, \mathcal{S})$.

Remark 4.2. If the minorization condition $(\mathcal{M})$ holds for some general Markov chain $\{X_n\}_{n \ge 0}$ with transition kernel P, then (see e.g. [12]) there exists a sequence of strictly increasing random times r_n, $n \ge 1$, such that $\{X_n\}_{n \ge 0}$ regenerates at each r_n, i.e.,

- r_1, $r_2 - r_1$, $r_3 - r_2, \ldots$ are finite a.s. and mutually independent;
- the sequence $\{r_{i+1} - r_i\}_{i \ge 0}$ is i.i.d.;
- the random blocks $\{X_{r_{i-1}}, X_{r_{i-1}+1}, \ldots, X_{r_i - 1}\}_{i \ge 1}$ are independent, where we make the convention that $r_0 = 0$; and

- $\mathbf{P}(X_{r_i} \in E) = \phi(E \cap E_0)$.

To be precise, consider the augmented chain $\{(X_n, \eta_n)\}_{n\geq 0}$ with the transition kernel P' that is defined by

$$
\begin{aligned}
&P'((x,0), E \times \{\eta\}) \\
&= \begin{cases} [\theta\eta + (1-\theta)(1-\eta)]P(x,E), & \text{for } x \notin \mathcal{C}_0, \\ [\theta\eta + (1-\theta)(1-\eta)](P(x,E) - \theta\phi(E\cap E_0))/(1-\theta), & \text{for } x \in \mathcal{C}_0, \end{cases} \\
&P'((x,1), E \times \{\eta\}) \\
&= \begin{cases} [\theta\eta + (1-\theta)(1-\eta)]P(x,E), & \text{for } x \notin \mathcal{C}_0, \\ [\theta\eta + (1-\theta)(1-\eta)]\phi(E\cap E_0), & \text{for } x \in \mathcal{C}_0, \end{cases}
\end{aligned}
$$

where $\eta \in \{0, 1\}$, and θ is w.l.o.g. assumed to be in $(0, 1]$. Note that

- $\{\eta_n\}_{n\geq 0}$ is a sequence of i.i.d. Bernoulli random variables with $\mathbf{P}(\eta_n = 1) = \theta$;
- $\{X_n\}_{n\geq 0}$ is a Markov chain with transition kernel P; and
- $\{X_n\}_{n\geq 0}$ and $\{\eta_n\}_{n\geq 0}$ are independent.

Moreover, $r_i - 1$ is identified as the i-th return time of the Markov chain $\{(X_n, \eta_n)\}_{n\geq 0}$ to the set $\mathcal{C}_0 \times \{1\}$, and hence, is a stopping time w.r.t. (X_n, η_n).

Consider

$$X_{n+1} = A_{n+1}X_n + B_{n+1}, \quad n \geq 0,$$

which was defined in (4.2.2). We need the following result that describes path properties of X_n. Let π denote the stationary distribution of the AR(1) process $\{X_n\}_{n\geq 0}$ as in (4.2.2). Define $\mathscr{B}_r(x) = \{x' \colon |x - x'| < r\}$ for $x \in \mathbb{R}$ and $r > 0$.

Result 4.2.1 (Lemma 2.2.3 and Proposition 2.2.4 of [26])**.** *Let $\{X_n\}_{n\geq 0}$ be such that $X_{n+1} = A_{n+1}X_n + B_{n+1}$. Supposing that Assumption 4.2.1 holds, we have that:*

1. *$\{X_n\}_{n\geq 0}$ satisfies the drift condition $(\mathcal{D})$ with $\mathcal{C} = [-M, M]$ for some constant $M \geq 0$.*
2. *$\{X_n\}_{n\geq 0}$ is π-irreducible.*
3. *$\{X_n\}_{n\geq 0}$ is geometrically ergodic.*

As we will see in Section 4.3, the regeneration scheme described in Remark 4.2 plays an important role in our analysis. Hence, we need Assumption 4.2.2 below, which guarantees the existence of the regeneration times. In Proposition 4.2.1, we give sufficient conditions for Assumption 4.2.2. Note that the first two conditions (up to a slight modification) can also be found in Lemma 2.2.3 of [26].

Assumption 4.2.2. The minorization condition $(\mathcal{M})$ is satisfied with $\mathcal{C}_0 = [-d, d]$ for some $d > 0$.

Proposition 4.2.1. *Assume that at least one of the following conditions hold.*

1. *Let $B_1 \geq b$ a.s. for some $b > 0$. Moreover, there exist intervals $I_1 = (a_1, a_2)$, $I_2 = (b_0 - \delta, b_0 + \delta)$ for some $a_1 < a_2$, b_0, $\delta > 0$, a σ-finite measure ν_0 with b_0 in the support of ν_0, a measure ϕ and a constant $c_0 > 0$ such that for any Borel sets D_1, $D_2 \subseteq \mathbb{R}$,*

$$\mathbf{P}((A_1, B_1) \in (D_1 \times D_2)) \geq c_0 |D_1 \cap I_1| \nu_0(D_2 \cap I_2),$$

   where $|\cdot|$ denotes the Lebesgue measure on $\mathbb{R}$.

2. *There exist intervals $I_1 = (a_0 - \delta, a_0 + \delta)$, $I_2 = (b_1, b_2)$ for some a_0, $b_1 < b_2$, $\delta > 0$, a σ-finite measure ν_0 with a_0 in the support of ν_0, a measure ϕ and a constant $c_0 > 0$ such that for any Borel sets D_1, $D_2 \subseteq \mathbb{R}$,*

$$\mathbf{P}((A_1, B_1) \in (D_1 \times D_2)) \geq c_0 \nu_0(D_1 \cap I_1) |D_2 \cap I_2|. \tag{4.2.3}$$

3. *Let $A_1 = cB_1$ for some c. A_1 has a density which is bounded from below by some $c_0 > 0$ on some interval $I_1 = (a_0 - \delta, a_0 + \delta)$.*

Then, for any $x_0 \in \mathbb{R}$, there exists $\epsilon = \epsilon(x_0)$, $\theta > 0$ such that

$$\theta |E \cap E_0| \leq P(x, E), \qquad x \in \mathscr{B}_\epsilon(x_0), E \in \mathcal{B}(\mathbb{R}). \tag{4.2.4}$$

Let $\{r_n\}_{n \geq 0}$ be the sequence of regeneration times of $\{X_n\}_{n \geq 0}$, cf. Remark 4.2 above. Next we state the existence of the moment generating function of r_1 in a neighborhood of the origin, which particularly implies the geometric decay of $\mathbf{P}(r_1 > k)$ as $k \to \infty$.

Lemma 4.2.1. *Let $\{X_n\}_{n \geq 0}$ be such that $X_{n+1} = A_{n+1} X_n + B_{n+1}$. Suppose that Assumption 4.2.1 and 4.2.2 hold. Let $\{r_n\}_{n \geq 0}$ be the sequence of regeneration times associated with $\mathcal{C}_0$. Let E_1 be a bounded set. There exists $t > 1$ such that*

$$\sup_{x \in E_1} \mathbf{E}[t^{r_1} \mid X_0 = x] < \infty.$$

4.2.2 A useful change of measure

Another helpful tool in our context is the so-called α-shifted change of measure (see e.g. [37, 36]). To begin with, let ν denote the distribution of $(\log A_n, B_n)$ with (A_n, B_n) being as in (4.2.2), and let ν^α denote the α-shifted distribution w.r.t. $\log A_n$, i.e,

$$\nu^\alpha(E) = \int_E e^{\alpha x} d\nu(x,y), \qquad E \in \mathfrak{B}(\mathbb{R}^2).$$

Let $\mathcal{L}(\log A_n, B_n)$ denote the probability law of $(\log A_n, B_n)$. Let $\mathscr{D}$ be the dual change of measure such that, under $\mathscr{D}$,

$$\mathcal{L}(\log A_n, B_n) = \begin{cases} \nu^\alpha, & \text{for } n \le T(u^\beta), \\ \nu, & \text{for } n > T(u^\beta). \end{cases} \tag{4.2.5}$$

Let $\mathbf{E}^\alpha$ and $\mathbf{E}^{\mathscr{D}}$ denote taking expectation w.r.t. the α-shifted measure and the dual change of measure $\mathscr{D}$, respectively. Defining

$$S_n = \sum_{i=1}^n \log A_i, \tag{4.2.6}$$

we have the following result.

Result 4.2.2 (Lemma 5.3 of [37]). *Let τ be a stopping time w.r.t. $\{X_n\}_{n\ge 0}$, let $g\colon \mathbb{R}^\infty \to [0,\infty]$ be a deterministic function, and let g_n denote its projection onto the first $n+1$ coordinates, i.e., $g_n(x_0,\ldots,x_n) = g(x_0,\ldots,x_n,0,0,\ldots)$. Then*

$$\begin{aligned} \mathbf{E}[g_{\tau-1}(X_0,\ldots,X_{\tau-1})] = \mathbf{E}^{\mathscr{D}} &\left[g_{\tau-1}(X_0,\ldots,X_{\tau-1}) e^{-\alpha S_{T(u^\beta)}} \mathbb{1}_{\{T(u^\beta)<\tau\}} \right] \\ &+ \mathbf{E}^{\mathscr{D}} \left[g_{\tau-1}(X_0,\ldots,X_{\tau-1}) e^{-\alpha S_\tau} \mathbb{1}_{\{T(u^\beta)\ge\tau\}} \right]. \end{aligned}$$

Our analysis relies on the fact that the Markov chain X_n is closely related to a multiplicative random walk, that is,

$$X_{n+1} \approx A_{n+1} X_n, \quad \text{for large } n.$$

Roughly speaking, the process X_n resembles a perturbation of a multiplicative random walk, in an asymptotic sense (for details see [37, 36]). Hence, it is natural to consider the "discrepancy" process between X_n and $\prod_{i=1}^n A_i$, which is defined as

$$Z_n = X_n e^{-S_n} = X_0 + \sum_{k=1}^n B_k e^{-S_k}, \quad n \ge 0, \tag{4.2.7}$$

where S_n is as in (4.2.6). On the other hand, under the α-shifted measure, we have $\mathbf{E}^\alpha \log A_1 = \mathbf{E} A_1^\alpha \log A_1 > 0$ by Assumption 4.2.1 and the convexity of the moment generating function of $\log A_1$. As a consequence, we have the following result.

Lemma 4.2.2. *Suppose that Assumption 4.2.1 holds. Under the α-shifted measure, the following holds.*

1. $|X_n| \uparrow \infty$ *a.s. as* $n \to \infty$.
2. $Z_n \xrightarrow{a.s.} Z$ *as* $n \to \infty$, *where* $Z = X_0 + \sum_{k=1}^\infty B_k e^{-S_k}$.

4.2.3 $\mathbb{M}$-convergence

In this section we briefly review the notion of $\mathbb{M}$-convergence, which was introduced in [87, Section 2] and turns out to be very useful in deriving our large deviations results. The rest of the section is based on [105, Section 2].

Let $(\mathbb{S}, d)$ be a complete separable metric space, and $\mathscr{S}$ be the Borel σ-algebra on $\mathbb{S}$. Given a closed subset $\mathbb{C}$ of $\mathbb{S}$, let $\mathbb{S} \setminus \mathbb{C}$ be equipped with the relative topology as a subspace of $\mathbb{S}$, and consider the associated sub σ-algebra $\mathscr{S}_{\mathbb{S}\setminus\mathbb{C}} = \{E\colon E \subseteq \mathbb{S} \setminus \mathbb{C}, A \in \mathscr{S}\}$ on it. Define $\mathbb{C}^r = \{x \in \mathbb{S}\colon d(x, \mathbb{C}) < r\}$ for $r > 0$, and let $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ be the class of measures defined on $\mathscr{S}_{\mathbb{S}\setminus\mathbb{C}}$ whose restrictions to $\mathbb{S} \setminus \mathbb{C}^r$ are finite for all $r > 0$. Topologize $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ with a subbasis $\{\{\nu \in \mathbb{M}(\mathbb{S} \setminus \mathbb{C})\colon \nu(f) \in G\}\colon f \in \mathcal{C}_{\mathbb{S}\setminus\mathbb{C}}, G \text{ open in } \mathbb{R}_+\}$, where $\mathcal{C}_{\mathbb{S}\setminus\mathbb{C}}$ is the set of real-valued, non-negative, bounded, continuous functions whose support is bounded away from $\mathbb{C}$ (i.e., $f(\mathbb{C}^r) = \{0\}$ for some $r > 0$). A sequence of measures $\nu_n \in \mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ converges to $\nu \in \mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ if $\nu_n(f) \to \nu(f)$ for each $f \in \mathcal{C}_{\mathbb{S}\setminus\mathbb{C}}$. We say that a set $E_1 \subseteq \mathbb{S}$ is bounded away from another set $E_2 \subseteq \mathbb{S}$ if $\inf_{x \in E_1, y \in E_2} d(x, y) > 0$. The following characterization of $\mathbb{M}$-convergence can be considered as a generalization of the classical notion of weak convergence of measures, see e.g. [14].

Result 4.2.3 (Theorem 2.1 of [87])**.** *Let ν, $\nu_n \in \mathbb{M}(\mathbb{S} \setminus \mathbb{C})$. We have $\nu_n \to \nu$ in $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ as $n \to \infty$ if and only if*

$$\overline{\lim_{n\to\infty}}\, \nu_n(F) \le \nu(F)$$

for all closed $F \in \mathscr{S}_{\mathbb{S}\setminus\mathbb{C}}$ bounded away from $\mathbb{C}$ and

$$\underline{\lim_{n\to\infty}}\, \nu_n(G) \ge \nu(G)$$

for all open $G \in \mathscr{S}_{\mathbb{S}\setminus\mathbb{C}}$ bounded away from $\mathbb{C}$.

We now introduce a new notion of equivalence between two families of random objects, which will prove to be useful in Section 4.7. Let $F_\delta = \{x \in \mathbb{S}\colon d(x, F) \leq \delta\}$ and $G^{-\delta} = ((G^c)_\delta)^c$.

Definition 4.2.1. Suppose that X_n and Y_n are random elements taking values in a complete separable metric space $(\mathbb{S}, d)$. Y_n is said to be asymptotically equivalent to X_n with respect to ϵ_n and $\mathbb{C}$, if, for each $\delta > 0$ and $\gamma > 0$,

$$\begin{aligned}&\overline{\lim_{n\to\infty}}\, \epsilon_n^{-1}\mathbf{P}(X_n \in (\mathbb{S}\setminus\mathbb{C})^{-\gamma}, d(X_n, Y_n) \geq \delta)\\ &= \overline{\lim_{n\to\infty}}\, \epsilon_n^{-1}\mathbf{P}(Y_n \in (\mathbb{S}\setminus\mathbb{C})^{-\gamma}, d(X_n, Y_n) \geq \delta) = 0.\end{aligned}$$

Remark 4.3. Note that the asymptotic equivalence w.r.t. $\mathbb{C}$ implies the asymptotic equivalence w.r.t. $\mathbb{C}'$ if $\mathbb{C} \subseteq \mathbb{C}'$. In view of this, the strongest notion of asymptotic equivalence w.r.t. a given sequence ϵ_n is the one w.r.t. an empty set. In this case, the conditions for the asymptotic equivalence reduce to a simple condition: $\mathbf{P}(d(X_n, Y_n) \geq \delta) = o(\epsilon_n)$ for any $\delta > 0$. In our context, this simple condition suffices for the case of $B_1 \geq 0$ in Section 4.3.2; however, we have to work with the case that $\mathbb{C}$ is not an empty set when we deal with general B_1 in Section 4.3.3.

The usefulness of this notion of equivalence comes from the following result.

Lemma 4.2.3. *Suppose that $\epsilon_n^{-1}\mathbf{P}(X_n \in \cdot) \to \nu(\cdot)$ in $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$ for some sequence ϵ_n and a closed set $\mathbb{C}$. If Y_n is asymptotically equivalent to X_n with respect to ϵ_n and $\mathbb{C}$, then the law of Y_n has the same normalized limit, i.e., $\epsilon_n^{-1}\mathbf{P}(Y_n \in \cdot) \to \nu(\cdot)$ in $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$.*

4.3 Main results

In this section, we state the main results of the chapter. In Section 4.2, we analyze the tail estimates of the area under the first return time/regeneration cycle (see (4.3.2) below), which are proved to be useful in deriving the sample-path large deviations of $\bar{X}_n$. While Section 4.3.2 deals with the case where $B_1 \geq 0$ almost surely, we consider in Section 4.3.3 the case where B_1 is a general $\mathbb{R}$-valued random variable. In Section 4.3.4, we extend the results in Sections 4.2–4.3.3 to the case where X_n is driven by more general recursions.

4.3.1 Tail estimates on the area under the first return time/regeneration cycle

The goal here is to provide tail asymptotics for the area under the first return time and the regeneration cycle. To be precise, let

$$\tau_d = \inf\{n \geq 1 \colon |X_n| \leq d\} \tag{4.3.1}$$

denote the first return time of X_n to the set $[-d, d]$, where d is such that $[-d, d] \cap \mathrm{supp}(\pi) \neq \emptyset$. Moreover, let $\{r_n\}_{n \geq 0}$ be the sequence regeneration times of $\{X_n\}_{n\geq 0}$. We denote the area under the first return time and the regeneration cycle by

$$\mathfrak{B} = \sum_{n=0}^{\tau_d - 1} X_n \quad \text{and} \quad \mathfrak{R} = \sum_{n=0}^{r_1 - 1} X_n, \tag{4.3.2}$$

respectively. As a preparatory result for our main theorems, we derive the tail asymptotics of $\mathfrak{B}$ and $\mathfrak{R}$. Let $Z = X_0 + \sum_{k=1}^{\infty} B_k e^{-S_k}$. Finally, let $C_{+,\infty}$ and $C_{-,\infty}$ be the constants satisfying

$$\mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_k} > u\right) \sim C_{+,\infty} u^{-\alpha}, \quad \text{and} \quad \mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_k} < -u\right) \sim C_{-,\infty} u^{-\alpha}, \tag{4.3.3}$$

respectively.

Theorem 4.3.1. *Suppose that Assumption 4.2.1 holds.*

1. *First*

$$\lim_{u\to\infty} u^{\alpha} \mathbf{P}(\mathfrak{B} > u) = C_{+,\infty} \mathbf{E}^{\alpha}[(Z^+)^{\alpha} \mathbb{1}_{\{\tau_d = \infty\}}]$$
$$\text{and} \quad \lim_{u\to\infty} u^{\alpha} \mathbf{P}(\mathfrak{B} < -u) = C_{-,\infty} \mathbf{E}^{\alpha}[(Z^-)^{\alpha} \mathbb{1}_{\{\tau_d = \infty\}}].$$

2. *If Assumption 4.2.2 holds additionally, then*

$$\lim_{u\to\infty} u^{\alpha} \mathbf{P}(\mathfrak{R} > u) = C_+ \quad \text{and} \quad \lim_{u\to\infty} u^{\alpha} \mathbf{P}(\mathfrak{R} < -u) = C_-,$$

where $C_+ = C_{+,\infty} \mathbf{E}^{\alpha}[(Z^+)^{\alpha} \mathbb{1}_{\{r_1 = \infty\}}]$ *and* $C_- = C_{-,\infty} \mathbf{E}^{\alpha}[(Z^-)^{\alpha} \mathbb{1}_{\{r_1 = \infty\}}]$.

4.3.2 One-sided large deviations

The aim of Sections 4.3.2 and 4.3.3 is to establish limit theorems for $\mathbf{P}(\bar{X}_n \in A)$ as $n \to \infty$ under two different settings, for some general measurable set A.

In this section, we consider the case where B_1 is nonnegative and C_+ as in Theorem 4.3.1 is strictly positive. To deal with the dependence structure of the Markov chain within the regeneration cycle, we consider in this section the M_1' metric space. To be precise, define the extended completed graph Γ'_ξ of ξ by

$$\Gamma'_\xi = \{(x,t) \in \mathbb{R} \times [0,1] \colon x \in [\xi(t^-) \wedge \xi(t), \xi(t^-) \vee \xi(t)]\},$$

where $\xi(0^-) = 0$. Define an order on the graph Γ'_ξ by saying that $(x_1,t_1) \le (x_2,t_2)$ if either (i) $t_1 < t_2$ or (ii) $t_1 = t_2$ and $|\xi(t_1^-) - x_1| \le |\xi(t_2^-) - x_2|$. Let $\Pi'(\xi)$ be the set of parametric representations of $\xi \in \mathbb{D}$, i.e., $(u,v) \in \Pi'(\xi)$ if (u,v) is a continuous nondecreasing function mapping $[0,1]$ onto Γ'_ξ. For any $\xi_1, \xi_2 \in \mathbb{D}$, the M_1' metric is defined by

$$d_{M_1'}(\xi_1,\xi_2) = \inf_{\substack{(u_i,v_i)\in\Pi'(\xi_i)\\ i\in\{1,2\}}} \|u_1 - u_2\|_\infty \vee \|v_1 - v_2\|_\infty.$$

From now on, in case it is not mentioned specifically, we consider the metric space $(\mathbb{D}, d_{M_1'})$ by default.

For the one-sided large deviations result, we need the following elements. We say that a function $\xi \in \mathbb{D}$ is *piecewise constant*, if there exist finitely many time points t_i such that $0 = t_0 < t_1 < \cdots < t_m = 1$ and ξ is constant on the intervals $[t_{i-1}, t_i)$ for all $1 \le i \le m$. For $\xi \in \mathbb{D}$, define the set of discontinuities of ξ by

$$\mathrm{Disc}(\xi) = \{t \in [0,1] \colon \xi(t) \ne \xi(t^-)\}, \tag{4.3.4}$$

where $\xi(0^-) = 0$. Define, for $j \ge 0$,

$$\underline{\mathbb{D}}_{\leqslant j} = \{\xi \in \mathbb{D} \colon \xi \text{ piecewise constant and nondecreasing, } |\mathrm{Disc}(\xi)| \le j\}.$$

For $z \in \mathbb{R}$, define

$$\underline{\mathbb{D}}^z_{\leqslant j} = \{\xi \in \mathbb{D} \colon \xi = z \cdot id + \xi', \xi' \in \underline{\mathbb{D}}_{\leqslant j}\}, \quad \text{for } j \ge 0. \tag{4.3.5}$$

For each constant $\gamma > 1$, let $\nu_\gamma(x,\infty) = x^{-\alpha}$, and let ν_γ^j denote the restriction (to $\mathbb{R}_+^{j\downarrow} = \{x \in \mathbb{R}^j \colon x_1 \ge \cdots \ge x_j > 0\}$) of the j-fold product measure of ν_γ. Let C_0^z be the Dirac measure concentrated on the linear function zt. For $j \ge 1$, define a sequence of Borel measures $C_j^z \in \mathbb{M}(\mathbb{D} \setminus \mathbb{D}_{\leqslant j-1})$ concentrated on $\mathbb{D}_{\leqslant j-1}$ as

$$C_j^z(\,\cdot\,) = \mathbf{E}\left[\nu_\alpha^j\left\{x \in (0,\infty)^j : z \cdot id + \sum_{i=1}^{j} x_i \mathbb{1}_{U_i} \in \,\cdot\,\right\}\right], \tag{4.3.6}$$

where α is as in Assumption 4.2.1 and the random variables U_i, $i \geq 1$, are i.i.d. uniform distributed on $[0,1]$. For $E \subseteq \mathbb{D}$ and $z \in \mathbb{R}$, define

$$\mathcal{J}_z^{\uparrow}(E) = \inf\{j : E \cap \mathbb{D}_{\leqslant j}^z \neq \emptyset\}. \tag{4.3.7}$$

Setting $\mu = \mathbf{E}B_1/(1 - \mathbf{E}A_1)$, we state below the main theorem for the one-sided case.

Theorem 4.3.2. *Suppose that Assumptions 4.2.1 and 4.2.2 hold. Moreover, let $B_1 \geq 0$ and C_+ as in Theorem 4.3.1 be strictly positive.*

1. *For each $j \geq 0$,*

$$n^{j(\alpha-1)}\mathbf{P}(\bar{X}_n \in \,\cdot\,) \to (C_+\mathbf{E}r_1)^j C_j^\mu(\,\cdot\,),$$

in $\mathbb{M}(\mathbb{D} \setminus \mathbb{D}_{\leqslant j-1}^\mu)$ as $n \to \infty$.

2. *Suppose that E is a measurable set. If $\mathcal{J}_\mu^{\uparrow}(E) < \infty$ and E is bounded away from $\mathbb{D}_{\leqslant \mathcal{J}_\mu^{\uparrow}(E)-1}$, then*

$$\begin{aligned} &\varliminf_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in E)}{n^{-\mathcal{J}_\mu^{\uparrow}(E)(\alpha-1)}} \geq (C_+\mathbf{E}r_1)^{\mathcal{J}_\mu^{\uparrow}(E)} C^\mu_{\mathcal{J}_\mu^{\uparrow}(E)}(E^\circ) \\ \text{and} \quad &\varlimsup_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in E)}{n^{-\mathcal{J}_\mu^{\uparrow}(E)(\alpha-1)}} \leq (C_+\mathbf{E}r_1)^{\mathcal{J}_\mu^{\uparrow}(E)} C^\mu_{\mathcal{J}_\mu^{\uparrow}(E)}(E^-). \end{aligned}$$

4.3.3 Two-sided large deviations

In this section, we consider the case where B_1 is a general $\mathbb{R}$-valued random variable, and hence, can take negative values as well. Similarly as in Section 4.3.2, we need the following elements. Define the set of step functions with less than j discontinuities by

$$\mathbb{D}_{\ll j} = \{\xi \in \mathbb{D} : \xi \text{ piecewise constant, } |\mathrm{Disc}(\xi)| < j\}, \quad \text{for } j \geq 0.$$

For $z \in \mathbb{R}$, define

$$\mathbb{D}_{\ll j}^z = \{\xi \in \mathbb{D} : \xi = z \cdot id + \xi', \xi' \in \mathbb{D}_{\ll j}\}, \quad \text{for } j \geq 0. \tag{4.3.8}$$

Let $C^z_{0,0}$ be the Dirac measure concentrated on the linear function zt. For each $(j,k) \in \mathbb{Z}^2_+ \setminus \{(0,0)\}$, define a measure $C^z_{j,k}$ as

$$C^z_{j,k}(\,\cdot\,) = \mathbf{E}\left[\nu_\alpha^{j+k}\left\{(x,y) \in (0,\infty)^{j+k} \colon z \cdot id + \sum_{i=1}^{j} x_i \mathbb{1}_{U_i} - \sum_{i=1}^{j} x_i \mathbb{1}_{V_i} \in \,\cdot\,\right\}\right], \tag{4.3.9}$$

where the random variables U_i, V_i are i.i.d. uniform distributed on $[0,1]$. For $E \subseteq \mathbb{D}$ and $z \in \mathbb{R}$, define

$$\mathcal{J}_z(E) = \inf\{j \colon E \cap \mathbb{D}^z_{\ll j} \neq \emptyset\}. \tag{4.3.10}$$

Recalling $\mu = \mathbf{E}B_1/(1-\mathbf{E}A_1)$, we state below the main theorem for the two-sided case.

Theorem 4.3.3. *Suppose that Assumptions 4.2.1 and 4.2.2 hold. Let $\mathbf{E}|B_1|^m < \infty$ for every $m \in \mathbb{Z}_+$. Moreover, let C_+, C_- be as in Theorem 4.3.1 such that $C_+C_- > 0$.*

1. *For each $j \geq 0$,*

$$n^{j(\alpha-1)}\mathbf{P}(\bar{X}_n \in \,\cdot\,) \to (\mathbf{E}r_1)^j \sum\nolimits_{(l,m)\in I_{=j}} (C_+)^l (C_-)^m C^\mu_{l,m}(\,\cdot\,),$$

   in $\mathbb{M}(\mathbb{D} \setminus \mathbb{D}^\mu_{\ll j})$ as $n \to \infty$, where $I_{=j} = \{(l,m) \in \mathbb{Z}^2_+ \colon l + m = j\}$.

2. *Suppose that E is a measurable set. If $\mathcal{J}_\mu(E) < \infty$ and E is bounded away from $\mathbb{D}_{\ll \mathcal{J}_\mu(E)}$, then*

$$\varliminf_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in E)}{n^{-\mathcal{J}_\mu(E)(\alpha-1)}} \geq (\mathbf{E}r_1)^{\mathcal{J}_\mu(E)} \sum\nolimits_{(l,m)} (C_+)^l (C_-)^m C^\mu_{l,m}(E^\circ)$$

$$\text{and} \quad \varlimsup_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in E)}{n^{-\mathcal{J}_\mu(E)(\alpha-1)}} \leq (\mathbf{E}r_1)^{\mathcal{J}_\mu(E)} \sum\nolimits_{(l,m)} (C_+)^l (C_-)^m C^\mu_{l,m}(E^-),$$

   where the summations are over all (l,m) that belong to the set $I_{=\mathcal{J}_\mu(E)}$.

4.3.4 Extension to general recursions

In this section we extend the results in Sections 4.2–4.3.3 to the case where X_n, $n \geq 0$, is defined by more general recursions. To be precise, consider the nondegenerate Markov chain X_n, $n \geq 0$, satisfying

$$X_{n+1} = f_{n+1}(X_n), \quad n \geq 0, \tag{4.3.11}$$

where $\{f_n\}_{n\geq 1}$ are i.i.d. copies of a random function f, which we make precise below. Let π denote the stationary solution to (4.3.11). We say that f satisfies the *Lipschitz condition* ($\mathcal{L}$), if there exists a nonnegative random variable L with $\mathbf{E}\log L < 0$ such that

$$\sup_{x_1\neq x_2} \frac{|f(x_1) - f(x_2)|}{|x_1 - x_2|} = L, \tag{$\mathcal{L}$}$$

where $\mathbf{E}[|\log L| + \log^+ |f(x_0)|] < \infty$, for some $x_0 \in \mathrm{supp}(\pi)$. Moreover, we say that f satisfies a *cancellation condition* ($\mathfrak{C}$), if there exists a nonarithmetic random variable $A \in [0,\infty)$ and a random vector $(\underline{B}, \overline{B}) \in \mathbb{R}^2$ such that, for all $x \in \mathrm{supp}(\pi)$,

$$Ax + \underline{B} \leq f(x) \leq Ax + \overline{B}. \tag{$\mathfrak{C}$}$$

We need the following assumption on f.

Assumption 4.3.1. Let f satisfy the conditions $\mathcal{L}$ and $\mathfrak{C}$ and be monotone increasing, i.e., $f(x_1) \leq f(x_2)$ almost surely for $x_1 \leq x_2$. Moreover, for the random vector $(A, \underline{B}, \overline{B})$ as in ($\mathfrak{C}$) we have the following.

1. $A \geq 0$ a.s. and the law of $\log A$ conditioned on $\{A > 0\}$ is nonarithmetic.
2. There exists $\alpha \in (1,\infty)$ such that $\mathbf{E}A^\alpha = 1$, $\mathbf{E}|\underline{B}|^\alpha + \mathbf{E}|\overline{B}|^\alpha < \infty$, and $\mathbf{E}A^\alpha \log^+ A < \infty$.

Without introducing any new notations, let τ_d denote the first return time of X_n to the set $[-d,d]$, let r_n, $n \geq 1$, denote the sequence of regeneration times of X_n associated with $\mathcal{C}_0$ as in Assumption 4.2.2. Let

$$\mathfrak{B} = \sum_{n=0}^{\tau_d - 1} X_n \quad \text{and} \quad \mathfrak{R} = \sum_{n=0}^{r_1 - 1} X_n$$

be the area under the first return time and the regeneration cycle, respectively. Let $Z_n = e^{-S_n} X_n$ with $S_n = \sum_{i=1}^n \log A_i$ and let Z be the almost sure limit—whose existence is shown in the proof of Theorem 4.3.4 below—of Z_n. Finally, let $C_{+,\infty}$ and $C_{-,\infty}$ be as in (4.3.3).

Theorem 4.3.4. *Suppose that Assumption 4.3.1 holds.*

1. *First*

$$\lim_{u\to\infty} u^\alpha \mathbf{P}(\mathfrak{B} > u) = C_{+,\infty} \mathbf{E}^\alpha[(Z^+)^\alpha \mathbb{1}_{\{\tau_d = \infty\}}]$$
$$\textit{and} \quad \lim_{u\to\infty} u^\alpha \mathbf{P}(\mathfrak{B} < -u) = C_{-,\infty} \mathbf{E}^\alpha[(Z^-)^\alpha \mathbb{1}_{\{\tau_d = \infty\}}].$$

2. If Assumption 4.2.2 holds additionally, then

$$\lim_{u\to\infty} u^\alpha \mathbf{P}(\mathfrak{R} > u) = C_+ \quad \text{and} \quad \lim_{u\to\infty} u^\alpha \mathbf{P}(\mathfrak{R} < -u) = C_-,$$

where $C_+ = C_{+,\infty}\mathbf{E}^\alpha[(Z^+)^\alpha \mathbb{1}_{\{r_1=\infty\}}]$ *and* $C_- = C_{-,\infty}\mathbf{E}^\alpha[(Z^-)^\alpha \mathbb{1}_{\{r_1=\infty\}}]$.

Let $\mu = \int_{\mathbb{R}} x\pi(dx)$ denote the expectation of the stationary distribution of X_n, $n \geq 0$. Recall that $\mathbb{D}^z_{\leqslant j}$, $C^z_j(\cdot)$, $\mathcal{J}^\uparrow_z(\cdot)$, $\mathbb{D}^z_{\ll j}$, $C^z_{j,k}(\cdot)$ and $\mathcal{J}_z(\cdot)$ were defined in (4.3.5), (4.3.6), (4.3.7), (4.3.8), (4.3.9), and (4.3.10), respectively. The following theorems are extensions of Theorems 4.3.2 and 4.3.3 above.

Theorem 4.3.5. *Suppose that Assumptions 4.3.1 and 4.2.2 hold. Moreover, assume that* $f_n([0,\infty)) \subseteq [0,\infty)$ *for* f_n *as in* (4.3.11) *and* C_+ *as in Theorem 4.3.4 is strictly positive.*

1. For each $j \geq 0$,

$$n^{j(\alpha-1)}\mathbf{P}(\bar{X}_n \in \cdot\,) \to (C_+)^j C^\mu_j(\,\cdot\,),$$

in $\mathbb{M}(\mathbb{D}\setminus\mathbb{D}^\mu_{\leqslant j-1})$ *as* $n\to\infty$.

2. Suppose that E *is a measurable set. If* $\mathcal{J}_\mu(E) < \infty$ *and* E *is bounded away from* $\mathbb{D}_{\leqslant \mathcal{J}_\mu(E)-1}$, *then*

$$\begin{aligned} &\varliminf_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in E)}{n^{-\mathcal{J}_\mu(E)(\alpha-1)}} \geq (C_+)^j C^\mu_{\mathcal{J}_\mu(E)}(E^\circ) \\ \text{and} \quad &\varlimsup_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in E)}{n^{-\mathcal{J}_\mu(E)(\alpha-1)}} \leq (C_+)^j C^\mu_{\mathcal{J}_\mu(E)}(E^-). \end{aligned}$$

Theorem 4.3.6. *Suppose that Assumptions 4.3.1 and 4.2.2 hold. Let* $\mathbf{E}|\underline{B}_1|^m + \mathbf{E}|\overline{B}_1|^m < \infty$ *for every* $m \in \mathbb{Z}_+$. *Moreover, let* C_+, C_- *be as in Theorem 4.3.4 such that* $C_+C_- > 0$.

1. For each $j \geq 0$,

$$n^{j(\alpha-1)}\mathbf{P}(\bar{X}_n \in \cdot\,) \to \sum\nolimits_{(l,m)\in I_{=j}} (C_+)^l (C_-)^m C^\mu_{l,m}(\,\cdot\,),$$

in $\mathbb{M}(\mathbb{D}\setminus\mathbb{D}^\mu_{\ll j})$ *as* $n\to\infty$, *where* $I_{=j} = \{(l,m)\in\mathbb{Z}^2_+ : l+m=j\}$.

2. *Suppose that E is a measurable set. If $\mathcal{J}'_\mu(E) < \infty$ and E is bounded away from $\mathbb{D}_{\ll \mathcal{J}'_\mu(E)}$, then*

$$\begin{aligned} \varliminf_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in E)}{n^{-\mathcal{J}'_\mu(E)(\alpha-1)}} &\geq \sum\nolimits_{(l,m)} (C_+)^l (C_-)^m C^\mu_{l,m}(E^\circ) \\ \text{and} \quad \varlimsup_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in E)}{n^{-\mathcal{J}'_\mu(E)(\alpha-1)}} &\leq \sum\nolimits_{(l,m)} (C_+)^l (C_-)^m C^\mu_{l,m}(E^-), \end{aligned}$$

where the summations are over all (l,m) that belong to the set $I_{=\mathcal{J}'_\mu(E)}$.

4.4 An application in barrier option pricing

In this section we consider an application that arises in the context of financial mathematics; in particular we consider a down-in barrier option (see Section 11.3 of [111]).

Let the daily log return of some underlying asset be modelled by an AR(1) process X_n, $n \geq 0$, as in (4.2.2). Let Assumptions 4.2.1 and 4.2.2 hold. Let $\mathbf{E}|B_1|^m < \infty$ for every $m \in \mathbb{Z}_+$. For real numbers a_- and a_+, we are interested in providing sharp large-deviations estimates for $\mathbf{P}(E_n)$ as $n \to \infty$, where

$$E_n = \left\{ \bar{X}_n \geq a_+, \min_{0\leq k\leq n} \bar{X}_k \leq -a_- \right\},$$

$a_+ > \max\{\mu, 0\}$, and $a_- > \max\{-\mu, 0\}$. This choice of (a_-, a_+) leads to the case where the rare event is caused by two big jumps, and hence, is particularly interesting. Note that the probability of E_n can be interpreted as the chance of exercising a down-in barrier option. Defining

$$E = \left\{ \xi \in \mathbb{D} \colon \xi(1) \geq a_+, \inf_{t\in[0,1]} \xi(t) \leq -a_- \right\},$$

we obtain $\mathbf{P}(E_n) = \mathbf{P}(\bar{X}_n \in E)$. Obviously, we have $\mathcal{J}_\mu(E) = 2$, where $\mathcal{J}_\mu$ was defined in (4.3.10). Hence, to apply Theorem 4.3.3, we need to show $d_{M'_1}(E, \mathbb{D}^\mu_{\ll 2}) \geq r$ for some $r > 0$. To see this, we assume that $d_{M'_1}(E, \mathbb{D}^\mu_{\ll 2}) < r$ for all $r > 0$. Therefore, for any $\epsilon > 0$, there exists $\xi_1 \in E$ and ξ_2 with $\xi_2(t) = \mu t + x\mathbb{1}_{[y,1]}(t)$, $x \in \mathbb{R}$, and $y \in [0,1]$ such that $d_{M'_1}(\xi_1, \xi_2) < r + \epsilon$. By the definition of the M'_1 metric, for any $\delta_1 > 0$, there exists $(u_i, v_i) \in \Pi'(\xi_i)$, $i \in \{1,2\}$, such that

$$\|u_1 - u_2\|_\infty \vee \|v_1 - v_2\|_\infty < d_{M'_1}(\xi_1, \xi_2) + \delta_1 < r + \epsilon + \delta_1. \tag{4.4.1}$$

By (4.4.1), we have that

$$|a_+ - (\mu + x)| = |u_1(1) - u_2(1)| \le \|u_1 - u_2\|_\infty \vee \|v_1 - v_2\|_\infty < r + \epsilon + \delta_1.$$

Letting ϵ, $\delta_1 \to 0$, we obtain that $x \ge (a_+ - \mu) - r > 0$ for sufficiently small r. On th other hand, by the fact $\inf_{t\in[0,1]} \xi_1(t) \le -a_-$, for any $\delta_2 > 0$, there exsits $t' \in [0,1]$ such that $\xi_1(t') < -a_- + \delta_2$. Let s be such that $v_1(s) = t'$. Let $t'' = v_2(s)$. Again using (4.4.1), we obtain that

$$|\xi_1(t') - (\mu t'' + x\mathbb{1}_{[y,1]}(t''))| = |u_1(s) - u_2(s)| \le \|u_1 - u_2\|_\infty \vee \|v_1 - v_2\|_\infty < r + \epsilon + \delta_1,$$

and hence,

$$\mu t'' + x\mathbb{1}_{[y,1]}(t'') < \xi_1(t') + (r + \epsilon + \delta_1) < -a_- + r + \epsilon + \delta_1 + \delta_2. \qquad (4.4.2)$$

Combining (4.4.2) with the fact that $x > 0$, we obtain that

$$\mu\mathbb{1}_{(-\infty,0)}(\mu) \le \mu t'' \le \mu t'' + x\mathbb{1}_{[y,1]}(t'') < -a_- + r + \epsilon + \delta_1 + \delta_2. \qquad (4.4.3)$$

Letting ϵ, δ_1, $\delta_2 \to 0$, we see that (4.4.3) is contradictory to $a_- > \max\{-\mu, 0\} \ge 0$. In view of the above discussion, we proved $d_{M_1'}(E, \mathbb{D}^{\mu}_{\ll 2}) \ge r$ for some $r > 0$, and hence, we are in the framework of Theorem 4.3.3.

Next we determine the preconstant in the asymptotics. Define m, $\pi_1 : \mathbb{D} \to \mathbb{R}$ by $m(\xi) = \inf_{t\in[0,1]} \xi(t)$, and $\pi_1(\xi) = \xi(1)$. Note that π_1 and m (cf. [117, Lemma 13.4.1]) are continuous. Thus, $E = m^{-1}(-\infty, -a_-] \cap \pi_1^{-1}[a_+, \infty)$ is a closed set. Recall, for $z \in \mathbb{R}$, that $C^z_{j,k}$ was defined in (4.3.9). Since $C^{\mu}_{2,0}(E) = C^{\mu}_{0,2}(E) = 0$, it remains to consider $C^{\mu}_{1,1}(E^\circ)$ and $C^{\mu}_{1,1}(E)$. Combining the fact that $m^{-1}(-\infty, -a_-) \cap \pi_1^{-1}(a_+, \infty) \subseteq E$ with the discussion after [105, Theorem 3.2], we conclude that E is a $C^{\mu}_{1,1}$-continuous set. Therefore, applying Theorem 4.3.3 we obtain

$$\mathbf{P}(E_n) \sim C^{\mu}_{1,1}(E) C_+ C_- n^{-2(\alpha-1)}$$

as $n \to \infty$. In particular, the probability of interest is regularly varying of index $2 - 2\alpha$.

4.5 Proofs of Section 4.2

Proof of Proposition 4.2.1. Note that part 1) and part 2) are already proved in [26, page 22], for the case where $x_0 \ne 0$. Hence, we will concentrate on showing part 2) (for the case $x_0 = 0$) and part 3).

Part 2): We focus on the case where $x_0 = 0$. Fix a Borel set E. In view of (4.2.3) we observe that

$$\begin{aligned} P(x, E) = \mathbf{E}[\mathbb{1}_{A_1 x + B_1 \in E}] &\geq c_0 \int_{I_1} \int_{I_2} \mathbb{1}_{\{ax+b \in E\}} db \nu_0(da) \\ &= c_0 \int_{I_1} \int_{E} \mathbb{1}_{\{z-ax \in I_2\}} dz \nu_0(da). \end{aligned}$$

Let

$$E_0 = (b_1 + \epsilon(|a_0 - \delta| \vee |a_0 + \delta|), b_2 - \epsilon(|a_0 - \delta| \vee |a_0 + \delta|)).$$

The set E_0 is not empty if we choose $\epsilon < (b_2 - b_1)/(2(|a_0 - \delta| \vee |a_0 + \delta|))$. Note that if $x \in \mathscr{B}_\epsilon(0)$, $z \in E_0$, and $a \in I_1$, then $|ax| < \epsilon(|a_0 - \delta| \vee |a_0 + \delta|)$ and $z - ax \in I_2$. Hence, we have that

$$P(x, E) \geq c_0 \int_{I_1} \int_{E \cap E_0} dz \nu_0(da) \geq c_0 \nu_0(I_1) |E \cap E_0|.$$

The constant $c_0 \nu_0(I_1)$ is strictly positive since a_0 belongs to the support of ν_0.

Part 3): Pick ϵ so that $-1/c \notin \mathscr{B}_\epsilon(x_0)$. Suppose that $c > 0$ and $x_0 \geq 0$. Note that, for any $x \in \mathscr{B}_\epsilon(x_0)$

$$P(x, E) = \mathbf{E}[\mathbb{1}_{A_1 x + B_1 \in E}] \geq c_0 (1/c + x_0 + \epsilon)^{-1} \int_E \mathbb{1}_{\{z/(x+1/c) \in I_1\}} dz.$$

Let

$$E_0 = \begin{cases} ((a_0 - \delta)/(1 + (x_0 + \epsilon)c), (a_0 + \delta)/(1 + (x_0 - \epsilon)c)) & \text{for } a_0 \geq 0, \\ ((a_0 - \delta)/(1 + (x_0 - \epsilon)c), (a_0 + \delta)/(1 + (x_0 + \epsilon)c)) & \text{for } a_0 < 0. \end{cases}$$

Observe that if $x \in \mathscr{B}_\epsilon(x_0)$ and $z \in E_0$ then $z/(x + 1/c) \in I_1 = (a_0 - \delta, a_0 + \delta)$ for δ sufficiently small. Hence, we have that

$$P(x, E) \geq c_0 (1/c + x_0 + \epsilon)^{-1} \int_{E \cap E_0} dz \geq c_0 (1/c + x_0 + \epsilon)^{-1} |E \cap E_0|.$$

This proves the result for the case where $c > 0$ and $x_0 \geq 0$; the proofs for the other cases are analogous. □

Proof of Lemma 4.2.1. First we claim that $[-M, M]$ is a petite set (cf. [89, page 124]) for any $M > 0$. To see this, note that $[-M, M] \subseteq \bigcup_{x \in [-M,M]} \mathscr{B}_\epsilon(x)$, where

ϵ is as in (4.2.4). Combining this with the facts that $[-M, M]$ is compact and $B_\epsilon(x)$ is open, there exists a finite N such that $[-M, M] \subseteq \bigcup_{i=1}^N \mathscr{B}_\epsilon(x_i)$. By Theorem 5.2.2 of [89], $\mathscr{B}_\epsilon(x_i)$ is a small set, and hence, is petite. Therefore there exists a finite subcover of petite sets. By Proposition 5.5.5 of [89], the interval $[-M, M]$ is petite. Now we turn back to proving the statement of Lemma 4.2.1. By Theorem 15.2.6 of [89], any bounded set is h-geometrically regular with $h(x) = |x|^\epsilon + 1$, $\epsilon \in (0, 1]$. Thus, from the definition of h-geometrical regularity (cf. page 373 of [89]), there exists $t = t(h, \mathcal{C}_0)$ such that $\sup_{x \in E_1} \mathbf{E}[\sum_{k=0}^{\tau_{\mathcal{C}_0}-1} h(X_k) t^k \mid X_0 = x] < \infty$. In particular,

$$\chi_1(t) = \sup_{x \in E_1} \mathbf{E}[t^{\tau_{\mathcal{C}_0}} \mid X_0 = x] < \infty, \tag{4.5.1}$$

since $h \geq 1$. On the other hand, $\sup_{x \in \mathcal{C}_0} \mathbf{E}[t^{\tau_{\mathcal{C}_0}} \mid X_0 = x] < \infty$. In particular,

$$\chi_2(t) = \sup_{x \in \mathcal{C}_0} \mathbf{E}[t^{\tau_{\mathcal{C}_0}} \mid X_0 = x, X_1 \sim (P(x, \cdot) - \phi(\,\cdot\,))/(1 - \theta)] < \infty, \tag{4.5.2}$$

where θ and ϕ are as in $(\mathcal{M})$. From the regeneration scheme as described in Remark 4.2.1, we obtain that

$$\sup_{x \in E_1} \mathbf{E}[t^{r_1} \mid X_0 = x] \leq \chi_1(t) \left(\theta + \sum_{n=1}^{\infty} \theta (1 - \theta)^n (\chi_2(t))^n \right). \tag{4.5.3}$$

By (4.5.2) and the dominated convergence theorem, $\chi_2(t) \downarrow 1$ as $t \downarrow 1$. Thus, we have that $\chi_2(t) < (1 - \theta)^{-1}$ for sufficiently small $t > 1$. For this choice of t, the r.h.s. of (4.5.3) converges by (4.5.1). $\square$

Proof of Lemma 4.2.2. By Assumption 4.2.1, the set $[M, \infty)$ is attainable for the process $\{|X_n|\}_{n \geq 0}$ for sufficiently large M. Hence, by Theorem 8.3.6 of [89], Lemma 4.2.2 is proved once we show

$$\mathbf{P}^\alpha(|X_n| \geq M, \text{ for all } n \geq 1 \,|\, |X_0| \geq 2M) > 0.$$

Note that

$$\begin{aligned} |X_n| = e^{S_n} \left| X_0 + \sum_{i=1}^n B_i e^{-S_i} \right| &\geq e^{S_n} \left(|X_0| - \sum_{i=1}^n |B_i| e^{-S_i} \right) \\ &\geq e^{S_n} \left(|X_0| - \sum_{i=1}^\infty |B_i| e^{-S_i} \right). \end{aligned}$$

Combining this with the fact that $\mathbf{E}^\alpha \log A_1 > 0$, we conclude that $\mathbf{P}(\exp(S_n) \geq 1,\ \text{for all } n \geq 1) = \mathbf{P}(S_n \geq 0,\ \text{for all } n \geq 1) > 0$, and hence, the first statement is proved. The second statement follows from the fact that the random walk $-S_n$ has a negative drift under the α-shifted measure. $\square$

Proof of Lemma 4.2.3. Let G be an open set bounded away from $\mathbb{C}$ so that $G \subseteq (\mathbb{S}\setminus\mathbb{C})^{-\gamma}$ for some $\gamma > 0$. For a given $\delta > 0$, due to the assumed asymptotic equivalence, $\mathbf{P}(X_n \in \mathbb{S}\setminus\mathbb{C})^{-\gamma}, d(X_n, Y_n) \geq \delta) = o(\epsilon_n)$. Therefore,

$$\begin{aligned}
&\underline{\lim}_{n\to\infty} \epsilon_n^{-1}\mathbf{P}(Y_n \in G) \\
&\geq \underline{\lim}_{n\to\infty} \epsilon_n^{-1}\mathbf{P}(X_n \in G^{-\delta}, d(X_n, Y_n) < \delta) \\
&= \underline{\lim}_{n\to\infty} \epsilon_n^{-1}\{\mathbf{P}(X_n \in G^{-\delta}) - \mathbf{P}(X_n \in G^{-\delta}, d(X_n, Y_n) \geq \delta)\} \\
&\geq \underline{\lim}_{n\to\infty} \epsilon_n^{-1}\{\mathbf{P}(X_n \in G^{-\delta}) - \mathbf{P}(X_n \in (\mathbb{S}\setminus\mathbb{C})^{-\gamma}, d(X_n, Y_n) \geq \delta)\} \\
&= \underline{\lim}_{n\to\infty} \epsilon_n^{-1}\mathbf{P}(X_n \in G^{-\delta}) \geq \nu(G^{-\delta}).
\end{aligned}$$

Since G is an open set, $G = \bigcup_{\delta>0} G^{-\delta}$. Due to the continuity of measures, $\lim_{\delta\to 0}\nu(G^{-\delta}) = \nu(G)$, and hence, we arrive at the lower bound

$$\underline{\lim}_{n\to\infty} \epsilon_n^{-1}\mathbf{P}(Y_n \in G) \geq \nu(G)$$

by taking $\delta \to 0$. Now, turning to the upper bound, consider a closed set F bounded away from $\mathbb{C}$ so that $F \subseteq (\mathbb{S}\setminus\mathbb{C})^{-\gamma}$ for some $\gamma > 0$. Given a $\delta > 0$, by the equivalence assumption, $\mathbf{P}(Y_n \in \mathbb{S}\setminus\mathbb{C})^{-\gamma}, d(X_n, Y_n) \geq \delta) = o(\epsilon_n)$. Therefore,

$$\begin{aligned}
&\overline{\lim}_{n\to\infty} \epsilon_n^{-1}\mathbf{P}(Y_n \in F) \\
&= \overline{\lim}_{n\to\infty} \epsilon_n^{-1}\{\mathbf{P}(Y_n \in F, d(X_n, Y_n) < \delta) + \mathbf{P}(Y_n \in F, d(X_n, Y_n) \geq \delta)\} \\
&\leq \overline{\lim}_{n\to\infty} \epsilon_n^{-1}\{\mathbf{P}(X_n \in F_\delta) + \mathbf{P}(Y_n \in F, d(X_n, Y_n) \geq \delta)\} \\
&= \overline{\lim}_{n\to\infty} \epsilon_n^{-1}\mathbf{P}(X_n \in F_\delta) \leq \nu(F_\delta)
\end{aligned}$$

as long as δ is small enough so that F_δ is bounded away from $\mathbb{C}$. Note that $\{F_\delta\}$ is a decreasing sequence of sets, $F = \bigcup_{\delta>0} F_\delta$ (since F is closed), and $\nu \in \mathbb{M}(\mathbb{S}\setminus\mathbb{C})$ (and hence ν is a finite measure on $\mathbb{S}\setminus\mathbb{C}^r$ for some $r > 0$ such that $F_\delta \subseteq \mathbb{S}\setminus\mathbb{C}^r$ for some $\delta > 0$). Due to the continuity (from above) of finite measures, $\lim_{\delta\to 0}\nu(F_\delta) = \nu(F)$. Therefore, we arrive at the upper bound $\overline{\lim}_{n\to\infty} \epsilon_n^{-1}\mathbf{P}(Y_n \in F) \leq \nu(F)$ by taking $\delta \to 0$. $\square$

4.6 Proofs of Section 4.3.1

This section provides the proof of Theorem 4.3.1. Before turning to technical details, we briefly describe our strategy for proving the tail asymptotics of $\mathfrak{B}$, while a similar idea is behind the proof for $\mathfrak{R}$. Defining

$$T(u) = \inf\{n \geq 0 \colon |X_n| > u\},\ K_\beta^\gamma(u) = \inf\{n > T(u^\beta) \colon |X_n| \leq u^\gamma\}, \quad (4.6.1)$$

we write

$$\mathfrak{B} = \sum_{n=0}^{T(u^\beta)-1} X_n + \sum_{n=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_n + \sum_{n=K_\beta^\gamma(u)}^{\tau_d-1} X_n. \quad (4.6.2)$$

where $0 < \gamma < \beta < 1$. The proof of Theorem 4.3.1 (1) is based on the following fact.

- On the event $\{T(u^\beta) < \tau_d\}$, the first and the last term on the right hand side (r.h.s.) of (4.6.2) are negligible in contributing to the tail asymptotics. Proposition 4.6.1 proves such properties. Lemma 4.6.1 is useful in showing Proposition 4.6.1.
- In view of the last bullet, the second term on the r.h.s. of (4.6.2) plays the key role in $\mathbf{P}(\mathfrak{B} > u)$. Our analysis relies on the fact that the Markov chain X_n resembles a multiplicative random walk in the corresponding regime. Proposition 4.6.2 proves such asymptotics. Lemmas 4.6.2, 4.6.6, and 4.6.3 are helpful for proving Proposition 4.6.2.

As mentioned above, the following propositions are useful in proving Theorem 4.3.1.

Proposition 4.6.1. *Suppose that Assumption 4.2.1 holds. Then there exists* $0 < \gamma < \beta < 1$ *such that*

$$\mathbf{P}\left(\left|\sum_{n=0}^{T(u^\beta)-1} X_n\right| > u, T(u^\beta) < \tau_d\right) \text{ and } \mathbf{P}\left(\left|\sum_{n=K_\beta^\gamma(u)}^{\tau_d-1} X_n\right| > u, T(u^\beta) < \tau_d\right)$$

are of order $o(u^{-\alpha})$ *as* $u \to \infty$.

Proposition 4.6.2. *Suppose that Assumption 4.2.1 holds. Then there exists* $0 < \gamma < \beta < 1$ *such that*

$$\lim_{u\to\infty} u^\alpha \mathbf{P}\left(\sum_{n=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_n > u, T(u^\beta) < \tau_d\right) = C_{+,\infty}\mathbf{E}^\alpha[(Z^+)^\alpha \mathbb{1}_{\{\tau_d=\infty\}}]$$

$$\textit{and} \quad \lim_{u\to\infty} u^{\alpha}\mathbf{P}\left(\sum_{n=T(u^{\beta})}^{K_{\beta}^{\gamma}(u)-1} X_n < -u, T(u^{\beta}) < \tau_d\right) = C_{-,\infty}\mathbf{E}^{\alpha}[(Z^{-})^{\alpha}\mathbb{1}_{\{\tau_d=\infty\}}].$$

Proof of Theorem 4.3.1 (1). Recalling $T(u^{\beta}) = \inf\{n \geq 0 \colon |X_n| > u^{\beta}\}$ for $\beta \in (0,1)$, we can write

$$\mathbf{P}(\pm\mathfrak{B} > u) \leq \mathbf{P}(\pm\mathfrak{B} > u, T(u^{\beta}) < \tau_d) + \mathbf{P}(|\mathfrak{B}| > u, T(u^{\beta}) \geq \tau_d). \tag{4.6.3}$$

Since $\mathbf{P}(\tau_d > n)$ decays geometrically in n, we have that

$$\begin{aligned}\mathbf{P}(|\mathfrak{B}| > u, T(u^{\beta}) \geq \tau_d) &\leq \mathbf{P}\left(\sum_{n=0}^{\tau_d-1}|X_n| > u, T(u^{\beta}) \geq \tau_d\right)\\ &\leq \mathbf{P}(u^{\beta}\tau_d \geq u) = \mathbf{P}(\tau_d \geq u^{1-\beta}) = o(u^{-\alpha}).\end{aligned} \tag{4.6.4}$$

Using (4.6.3) and (4.6.4), we can focus on analyzing the first term on the r.h.s. of (4.6.3). For $0 < \gamma < \beta < 1$, recall $K_{\beta}^{\gamma}(u) = \inf\{n \geq T(u^{\beta}) \colon |X_n| \leq u^{\gamma}\}$. Using the decomposition in (4.6.2), we obtain that, for $\epsilon \in (0,1)$,

$$\begin{aligned}\mathbf{P}(\mathfrak{B} > u, T(u^{\beta}) < \tau_d) \leq \mathbf{P}&\left(\left|\sum_{n=0}^{T(u^{\beta})-1} X_n\right| > \frac{\epsilon u}{2}, T(u^{\beta}) < \tau_d\right)\\ &+\mathbf{P}\left(\sum_{n=T(u^{\beta})}^{K_{\beta}^{\gamma}(u)-1} X_n > (1-\epsilon)u, T(u^{\beta}) < \tau_d\right)\\ &+\mathbf{P}\left(\left|\sum_{n=K_{\beta}^{\gamma}(u)}^{\tau_d-1} X_n\right| > \frac{\epsilon u}{2}, T(u^{\beta}) < \tau_d\right),\end{aligned} \tag{4.6.5}$$

and

$$\begin{aligned}\mathbf{P}(\mathfrak{B} > u, T(u^{\beta}) < \tau_d) \geq -\mathbf{P}&\left(\left|\sum_{n=0}^{T(u^{\beta})-1} X_n\right| > \frac{\epsilon u}{2}, T(u^{\beta}) < \tau_d\right)\\ &+\mathbf{P}\left(\sum_{n=T(u^{\beta})}^{K_{\beta}^{\gamma}(u)-1} X_n > (1+\epsilon)u, T(u^{\beta}) < \tau_d\right)\\ &-\mathbf{P}\left(\left|\sum_{n=K_{\beta}^{\gamma}(u)}^{\tau_d-1} X_n\right| > \frac{\epsilon u}{2}, T(u^{\beta}) < \tau_d\right).\end{aligned} \tag{4.6.6}$$

Moreover, we can use similar estimates to "sandwich" the quantity $\mathbf{P}(\mathfrak{B} < -u, T(u^\beta) < \tau_d)$. Thus, using Propositions 4.6.1 and 4.6.2 above, we prove Theorem 4.3.1 (1). $\square$

We need the following propositions to prove Theorem 4.3.1 (2).

Proposition 4.6.3. *Suppose that Assumptions 4.2.1 and 4.2.2 hold. Then there exists* $0 < \gamma < \beta < 1$ *such that*

$$\mathbf{P}\left(\left|\sum_{n=0}^{T(u^\beta)-1} X_n\right| > u, T(u^\beta) < r_1\right) \text{ and } \mathbf{P}\left(\left|\sum_{n=K_\beta^\gamma(u)}^{\tau_d-1} X_n\right| > u, T(u^\beta) < r_1\right)$$

are of order $o(u^{-\alpha})$ *as* $u \to \infty$.

Proposition 4.6.4. *Suppose that Assumptions 4.2.1 and 4.2.2 hold. Then there exists* $0 < \gamma < \beta < 1$ *such that*

$$\lim_{u\to\infty} u^\alpha \mathbf{P}\left(\sum_{n=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_n > u, T(u^\beta) < r_1\right) = C_+$$

$$\text{and} \quad \lim_{u\to\infty} u^\alpha \mathbf{P}\left(\sum_{n=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_n < -u, T(u^\beta) < r_1\right) = C_-,$$

where $C_+ = C_{+,\infty}\mathbf{E}^\alpha[(Z^+)^\alpha \mathbb{1}_{\{r_1=\infty\}}]$ *and* $C_+ = C_{+,\infty}\mathbf{E}^\alpha[(Z^-)^\alpha \mathbb{1}_{\{r_1=\infty\}}]$.

Proof of Theorem 4.3.1 (2). Using similar arguments as in (4.6.3) and (4.6.4), we can focus on $\mathbf{P}(\pm\mathfrak{R} > u, T(u^\beta) < r_1)$. Combining the similar "sandwich" technique as in (4.6.5)–(4.6.6) with Proposition 4.6.3, it remains to analyze

$$u^\alpha \mathbf{P}\left(\sum_{n=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_n > u, T(u^\beta) < r_1\right).$$

Using Proposition 4.6.4, we conclude the proof. $\square$

Next we prove Proposition 4.6.1. For this, we need the following lemma. Let $\{Y_n\}_{n\ge0}$ be the $\mathbb{R}_+$-valued Markov chain defined by $Y_{n+1} = A_{n+1}Y_n + |B_{n+1}|$, for $n \ge 0$, and $\tau = \inf\{n \ge 1\colon Y_n \le d\}$.

Lemma 4.6.1. *Suppose that Assumption 4.2.1 holds. Let $L > 0$, and let $\epsilon > 0$ be such that $\lfloor \alpha - \epsilon \rfloor \geq 1$. Then there exists a positive constant c such that, for sufficiently large x,*

$$\mathbf{E}[\tau^{\alpha+L} \mid Y_0 = x] \leq c x^{\lfloor \alpha - \epsilon \rfloor}.$$

In particular $\mathbf{E}[\tau^{\alpha+L} \mid Y_0 = x] = \mathcal{O}(x)$.

Proof of Proposition 4.6.1. To begin with, note that

$$\mathbf{P}\left(\left|\sum_{n=0}^{T(u^\beta)-1} X_n\right| > u, T(u^\beta) < \tau_d\right) \leq \mathbf{P}\left(\sum_{n=0}^{T(u^\beta)-1} |X_n| > u, T(u^\beta) < \tau_d\right)$$
$$\leq \mathbf{P}(u^\beta \tau_d > u) = \mathbf{P}(\tau_d > u^{1-\beta}),$$

which decays exponentially. It remains to show the second claim. Define

$$\mathfrak{E}_1(u) = \{\exists\, n \in \{K_\beta^\gamma(u), K_\beta^\gamma(u)+1, \ldots, \tau_d\} \colon |X_n| \geq u^\rho\}.$$

Note that

$$\mathbf{P}\left(\left|\sum_{n=K_\beta^\gamma(u)}^{\tau_d-1} X_n\right| > u, T(u^\beta) < \tau_d\right)$$
$$\leq \mathbf{P}\left(\sum_{n=K_\beta^\gamma(u)}^{\tau_d-1} |X_n| > u, T(u^\beta) < \tau_d, \mathfrak{E}_1(u)\right)$$
$$+ \mathbf{P}\left(\sum_{n=K_\beta^\gamma(u)}^{\tau_d-1} |X_n| > u, T(u^\beta) < \tau_d, (\mathfrak{E}_1(u))^c\right),$$

where the second term in the last equation is bounded by $\mathbf{P}(\tau_d > u^{1-\rho})$, and hence is of order $o(u^{-\alpha})$. It remains to analyze the first term, which is bounded by $\mathbf{P}(T(u^\beta) < \tau_d, \mathfrak{E}_1(u))$. Our goal here is to show that

$$\mathbf{P}(T(u^\beta) < \tau_d, \mathfrak{E}_1(u)) = o(u^{-\alpha}), \quad \text{as } u \to \infty. \tag{4.6.7}$$

To begin with, note that, under the dual change of measure $\mathscr{D}$ we have $K_\beta^\gamma(u) < \infty$ almost surely. Moreover, $|X_{K_\beta^\gamma(u)+n}| \leq Y'_n$, for all $n \geq 0$, where $\{Y'_n\}_{n \geq 0}$ is the AR(1) process that is defined by

$$Y'_0 = u^\gamma, \qquad Y'_{n+1} = A_{K_\beta^\gamma(u)+n+1} Y'_n + |B_{K_\beta^\gamma(u)+n+1}|, \qquad \text{for } n \geq 0.$$

Hence, by defining $\tau' = \inf\{n \geq 1 \colon Y_n' \leq d\}$, we have that

$$\begin{aligned}\mathbf{P}(T(u^\beta) < \tau_d, \mathfrak{E}_1(u)) &= \mathbf{E}^{\mathscr{D}}[e^{-S_{T(u^\beta)}} \mathbb{1}_{\{T(u^\beta)<\tau_d\}} \mathbb{1}_{\mathfrak{E}_1(u)}] \\ &= \mathbf{E}^{\mathscr{D}}[e^{-S_{T(u^\beta)}} \mathbb{1}_{\{|X_n|>d, \forall n \leq T(u^\beta)\}} \mathbb{1}_{\mathfrak{E}_1(u)}] \\ &\leq \mathbf{E}^{\mathscr{D}}[e^{-S_{T(u^\beta)}} \mathbb{1}_{\{|X_n|>d, \forall n \leq T(u^\beta)\}} \mathbb{1}_{\{\exists\, n \leq \tau' \colon Y_n' \geq u^\rho\}}].\end{aligned}$$

Now using the strong Markov property we obtain that

$$\begin{aligned}\mathbf{P}(T(u^\beta) < \tau_d, \mathfrak{E}_1(u)) &\leq \mathbf{E}^{\mathscr{D}}[e^{-S_{T(u^\beta)}} \mathbb{1}_{\{|X_n|>d, \forall n \leq T(u^\beta)\}}] \mathbf{P}(\exists\, n \leq \tau' \colon Y_n' \geq u^\rho) \\ &= \mathbf{P}(T(u^\beta) < \tau_d) \mathbf{P}(\exists\, n \leq \tau' \colon Y_n' \geq u^\rho),\end{aligned}$$

where $\mathbf{P}(T(u^\beta) < \tau_d) \sim c u^{-\alpha\beta}$ (cf. Corollary 4.2 of [36]). It remains to analyze the asymptotic behavior of

$$\mathbf{P}(\exists\, n \leq \tau' \colon Y_n' \geq u^\rho) = \mathbf{P}(\exists\, n \leq \tau \colon Y_n \geq u^\rho \,|\, Y_0 = u^\gamma), \qquad \text{as } u \to \infty,$$

where $Y_{n+1} = A_{n+1} Y_n + |B_{n+1}|$, for $n \geq 0$, and $\tau = \inf\{n \geq 1 \colon Y_n \leq d\}$. Once again we adopt the idea of dual change of measure. To be precise, setting $T = \inf\{n \geq 1 \colon Y_n \geq u^\rho\}$, we apply the α-shifted change of measure over the time interval $[1, T]$. By doing this we obtain that

$$\begin{aligned}u^{\alpha(\rho-\gamma)} \mathbf{P}(T < \tau \,|\, Y_0 = u^\gamma) &= u^{\alpha(\rho-\gamma)} \mathbf{E}^\alpha \left[e^{-\alpha S_T} \mathbb{1}_{\{T<\tau\}} \,\middle|\, Y_0 = u^\gamma \right] \\ &= \mathbf{E}^\alpha \left[\left(\frac{Y_T}{u^\rho} \right)^{-\alpha} \left(\frac{Y_T}{e^{S_T} u^\gamma} \right)^{\alpha} \mathbb{1}_{\{T<\tau\}} \,\middle|\, Y_0 = u^\gamma \right] \\ &\leq \mathbf{E}^\alpha \left[\left(\frac{Y_T}{e^{S_T} u^\gamma} \right)^{\alpha} \mathbb{1}_{\{T<\tau\}} \,\middle|\, Y_0 = u^\gamma \right].\end{aligned}$$

Now it is sufficient to show that

$$\varlimsup_{u \to \infty} \mathbf{E}^\alpha \left[\left(\frac{Y_T}{e^{S_T} u^\gamma} \right)^{\alpha} \mathbb{1}_{\{T<\tau\}} \,\middle|\, Y_0 = u^\gamma \right] < \infty, \tag{4.6.8}$$

since once it is proved we can set $\beta + \rho - \gamma > 1$ so that $\mathbf{P}(T(u^\beta) < \tau_d, \mathfrak{E}(u)) = o(u^{-\alpha})$. Note that

$$\frac{Y_T}{e^{S_T} u^\gamma} = e^{-S_T} u^{-\gamma} \left(e^{S_T} u^\gamma + e^{S_T} \sum_{k=1}^{T} |B_k| e^{-S_k} \right) = 1 + u^{-\gamma} \sum_{k=1}^{T} |B_k| e^{-S_k}.$$

Thus, we have that

$$\frac{Y_T}{e^{S_T}u^\gamma}\mathbb{1}_{\{T<\tau\}} \le 1 + u^{-\gamma}\sum_{k=1}^{T}|B_k|e^{-S_k}\mathbb{1}_{\{T<\tau\}} \le 1 + u^{-\gamma}\sum_{k=1}^{\infty}|B_k|e^{-S_k}\mathbb{1}_{\{k<\tau\}},$$

and hence,

$$\begin{aligned}
&\mathbf{E}^\alpha\left[\left(\frac{Y_T}{e^{S_T}u^\gamma}\mathbb{1}_{\{T<\tau\}}\right)^\alpha \,\middle|\, Y_0 = u^\gamma\right]^{1/\alpha} \\
&\le \mathbf{E}^\alpha\left[\left(1 + u^{-\gamma}\sum_{k=1}^{\infty}|B_k|e^{-S_k}\mathbb{1}_{\{k<\tau\}}\right)^\alpha \,\middle|\, Y_0 = u^\gamma\right]^{1/\alpha} \\
&\le 1 + \sum_{k=1}^{\infty}\mathbf{E}^\alpha\left[u^{-\alpha\gamma}|B_k|^\alpha e^{-\alpha S_k}\mathbb{1}_{\{k<\tau\}} \,\middle|\, Y_0 = u^\gamma\right]^{1/\alpha} \qquad (4.6.9)\\
&= 1 + u^{-\gamma}\sum_{k=1}^{\infty}\mathbf{E}^\alpha\left[e^{-\alpha S_k}|B_k|^\alpha\mathbb{1}_{\{k<\tau\}} \,\middle|\, Y_0 = u^\gamma\right]^{1/\alpha} \\
&= 1 + u^{-\gamma}\sum_{k=1}^{\infty}\mathbf{E}\left[|B_k|^\alpha\mathbb{1}_{\{k<\tau\}} \,\middle|\, Y_0 = u^\gamma\right]^{1/\alpha} \\
&= 1 + u^{-\gamma}\sum_{k=1}^{\infty}(\mathbf{E}|B_k|^\alpha)^{1/\alpha}\mathbf{P}(\tau > k \,|\, Y_0 = u^\gamma)^{1/\alpha} \\
&\le 1 + u^{-\gamma}(\mathbf{E}|B_1|^\alpha)^{1/\alpha}\mathbf{E}[\tau^{\alpha+L} \,|\, Y_0 = u^\gamma]\sum_{k=1}^{\infty}k^{-(\alpha+\epsilon)/\alpha},
\end{aligned}$$

for some $L > 0$, where in (4.6.9) we used the Minkowski inequality. Using Lemma 4.6.1 above, we prove (4.6.7), (4.6.8), and hence, Proposition 4.6.1. □

The following lemmas are useful in proving Proposition 4.6.2. Let $C_{+,\infty}$ be as in (4.3.3). Set

$$\mathscr{G}_+(u) = u^{(1-\beta)\alpha}\mathbf{P}^{\mathscr{D}}\left(\sum_{n=T(u^\beta)}^{K_\beta^\gamma(u)-1}X_n > u \,\middle|\, \mathcal{F}_{T(u^\beta)}\right)\left(\frac{X_{T(u^\beta)}}{u^\beta}\right)^{-\alpha}\mathbb{1}_{\{Z_{T(u^\beta)}>0\}}, \qquad (4.6.10)$$

and

$$\mathscr{G}_-(u) = u^{(1-\beta)\alpha}\mathbf{P}^{\mathscr{D}}\left(\sum_{n=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_n > u \,\middle|\, \mathcal{F}_{T(u^\beta)}\right)\left|\frac{X_{T(u^\beta)}}{u^\beta}\right|^{-\alpha}\mathbb{1}_{\{Z_{T(u^\beta)}\leq 0\}}. \tag{4.6.11}$$

Lemma 4.6.2. *Suppose that Assumption 4.2.1 holds. Under the measure* $\mathbf{P}^\alpha$*,*

$$\mathscr{G}_+(u) \xrightarrow{a.s.} C_{+,\infty}\mathbb{1}_{\{Z>0\}} \quad \textit{and} \quad \mathscr{G}_-(u) \xrightarrow{a.s.} 0, \quad \textit{as } u\to\infty.$$

Moreover, $\mathscr{G}_+(u)$ *and* $\mathscr{G}_-(u)$ *are bounded in* u *by some constants almost surely.*

Recall that Z_n, τ_d, and $T(u)$ are defined in (4.2.7), (4.3.1), and (4.6.1), respectively.

Lemma 4.6.3. *Suppose that Assumption 4.2.1 holds. The random variables* $Z^+_{T(u^\beta)}\mathbb{1}_{\{T(u^\beta)<\tau_d\}}$ *and* $Z^-_{T(u^\beta)}\mathbb{1}_{\{T(u^\beta)<\tau_d\}}$ *are bounded by*

$$\bar{Z} = |X_0| + \sum_{n=1}^{\infty}|B_n|e^{-S_n}\mathbb{1}_{\{n<\tau_d\}}.$$

Moreover, $\mathbf{E}^\alpha[\bar{Z}^\alpha] < \infty$.

Proof of Proposition 4.6.2. We focus on deriving the first asymptotics, since the second one follows using similar arguments. Note that

$$\begin{aligned}
&u^\alpha\mathbf{P}\left(\sum_{n=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_n > u, T(u^\beta)<\tau_d\right)\\
&= u^\alpha\mathbf{P}\left(\sum_{n=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_n > u, X_{T(u^\beta)}>0, T(u^\beta)<\tau_d\right)\\
&\quad + u^\alpha\mathbf{P}\left(\sum_{n=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_n > u, X_{T(u^\beta)}<0, T(u^\beta)<\tau_d\right)\\
&= \mathbf{(I.1)} + \mathbf{(I.2)}.
\end{aligned} \tag{4.6.12}$$

We start considering the first term on the r.h.s. of (4.6.12). Applying the dual change of measure $\mathscr{D}$ together with Result 4.2.2, we obtain that

$$\mathbf{(I.1)} = \mathbf{E}^{\mathscr{D}}[g_{\tau_d-1}(X_0,\ldots,X_{\tau_d-1})\mathbb{1}_{\{X_{T(u^\beta)}>0\}}e^{-\alpha S_{T(u^\beta)}}\mathbb{1}_{\{T(u^\beta)<\tau_d\}}],$$

where we recall that g_{τ_d-1} is the projection of the function

$$g(X_0, X_1, \ldots) = 1, \quad \text{if } \sum\nolimits_{n=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_n > u,$$

onto its first $\tau_d - 1$ coordinates. Recall $Z_n = X_n e^{-S_n} = X_0 + \sum_{k=1}^n e^{-S_k}$ was defined in (4.2.7). Note that

$$\begin{aligned}
\textbf{(I.1)} &= u^\alpha \mathbf{E}^{\mathscr{D}} \left[g_{\tau_d-1}(X_0, \ldots, X_{\tau_d-1}) \mathbb{1}_{\{X_{T(u^\beta)}>0\}} e^{-\alpha S_{T(u^\beta)}} \mathbb{1}_{\{T(u^\beta)<\tau_d\}} \right] \\
&= u^\alpha \mathbf{E}^{\mathscr{D}} \Big[g_{\tau_d-1}(X_0, \ldots, X_{\tau_d-1}) |X_{T(u^\beta)}|^{-\alpha} \\
&\qquad\qquad |X_{T(u^\beta)}|^\alpha \mathbb{1}_{\{X_{T(u^\beta)}>0\}} e^{-\alpha S_{T(u^\beta)}} \mathbb{1}_{\{T(u^\beta)<\tau_d\}} \Big] \\
&= \mathbf{E}^{\mathscr{D}} \left[(Z^+_{T(u^\beta)})^\alpha \mathbb{1}_{\{T(u^\beta)<\tau_d\}} \mathscr{G}_+(u) \right],
\end{aligned} \tag{4.6.13}$$

for all $n \geq 0$. Using Lemma 4.6.2, Lemma 4.6.3, the dominated convergence theorem and the fact that $T(u^\beta) \to \infty$ as $u \to \infty$, we obtain that

$$\begin{aligned}
\lim_{u\to\infty} \textbf{(I.1)} &= \lim_{u\to\infty} \mathbf{E}^{\mathscr{D}} \left[(Z^+_{T(u^\beta)})^\alpha \mathbb{1}_{\{T(u^\beta)<\tau_d\}} \mathscr{G}_+(u) \right] \\
&= \lim_{u\to\infty} \mathbf{E}^{\alpha} \left[(Z^+_{T(u^\beta)})^\alpha \mathbb{1}_{\{T(u^\beta)<\tau_d\}} \mathscr{G}_+(u) \right] \\
&= \mathbf{E}^{\alpha} \left[\lim_{u\to\infty} (Z^+_{T(u^\beta)})^\alpha \mathbb{1}_{\{T(u^\beta)<\tau_d\}} \mathscr{G}_+(u) \right] \\
&= \mathbf{E}^{\alpha} \left[(Z^+)^\alpha \mathbb{1}_{\{\tau_d=\infty\}} C_{+,\infty} \right] \\
&= C_{+,\infty} \mathbf{E}^{\alpha} \left[(Z^+)^\alpha \mathbb{1}_{\{\tau_d=\infty\}} \right].
\end{aligned}$$

Analogously, we have that

$$\textbf{(I.2)} = \mathbf{E}^{\mathscr{D}} \left[(Z^-_{T(u^\beta)})^\alpha \mathbb{1}_{\{T(u^\beta)<\tau_d\}} \mathscr{G}_-(u) \right] \to 0, \quad \text{as } u \to \infty, \tag{4.6.14}$$

where $\mathscr{G}_-(u)$ was defined in (4.6.11). Using (4.6.12), (4.6.13), and (4.6.14), we prove the first asymptotics in Proposition 4.6.2. The second one can be shown analogously. $\square$

We need the following lemmas to prove Proposition 4.6.3. Let $Y_{n+1} = A_{n+1}Y_n + |B_{n+1}|$ and let $r - 1$ be the first time that (Y_n, η_n) returns to the set $[-d, d] \times \{1\}$.

Lemma 4.6.4. *Suppose that Assumptions 4.2.1 and 4.2.2 hold. Let $\epsilon > 0$, and let $L > 0$ be such that $\lfloor \alpha - \epsilon \rfloor \geq 1$. Then there exists a positive constant c such that, for sufficiently large x,*

$$\mathbf{E}[r^{\alpha+L} \mid Y_0 = x] \leq cx^{\lfloor \alpha - \epsilon \rfloor}.$$

In particular, $\mathbf{E}[r^{\alpha+L} \mid Y_0 = x] = \mathcal{O}(x)$.

Lemma 4.6.5. *Suppose that Assumptions 4.2.1 and 4.2.2 hold. We have that*

$$\lim_{u\to\infty} u^{\alpha}\mathbf{P}(T(u) < r_1) = \mathbf{E}\left[e^{-\alpha\mathfrak{X}}\right] \mathbf{E}^{\alpha}\left[|Z|^{\alpha}\mathbb{1}_{\{r_1=\infty\}}\right],$$

where $\mathfrak{X}$ is the positive random variable such that $\log X_{T(u)} - \log u$ converges in distribution to $\mathfrak{X}$ as $u \to \infty$ under $\mathbf{P}^{\alpha}$.

Proof of Proposition 4.6.3. By replacing τ_d with r_1, the proposition can be shown using almost identical arguments as in the proof of Proposition 4.6.1. Nonetheless, we need to show that

- $\mathbf{P}(T(u^{\beta}) < r_1) \sim cu^{-\alpha\beta}$ for some constant c, and that
- $\mathbf{E}[r^{\alpha+\epsilon}|Y_0 = x] = \mathcal{O}(x)$, where $Y_{n+1} = A_{n+1}Y_n + |B_{n+1}|$ and $r - 1$ is the first time that (Y_n, η_n) returns to the set $[-d, d] \times \{1\}$.

For this, we use Lemmas 4.6.4 and 4.6.5 above. □

Proof of Proposition 4.6.4. Using Lemma 4.6.2, Lemma 4.6.3, the dominated convergence theorem and the fact that $T(u^{\beta}) \to \infty$ as $u \to \infty$, one can prove the first asymptotics. The second one follows by a similar analysis. □

Next we provide the proofs of all lemmas in this section. To show Lemma 4.6.1, we introduce a result on bounding functionals of passage times for Markov chains. Let $\{V_n\}_{n\geq 0}$ be an $\{\mathcal{F}_n\}$-adapted stochastic process taking values in an unbounded subset of $\mathbb{R}_+$. Let $\{U_n\}_{n\geq 0}$ be another $\{\mathcal{F}_n\}$-adapted stochastic process taking values in an unbounded subset of $\mathbb{R}_+$ such that for any $n \geq 0$, U_n is integrable. Let $\tau_d = \inf\{n \geq 0 \colon V_n \leq d\}$ be the first time V_n returning to the set $(-\infty, d]$.

Result 4.6.1 (Theorem 2.2′ of [10])**.** *Suppose there exists a positive real number d and positive on (d, ∞) functions g, h such that for any $n \geq 0$, $U_n \leq h(V_n)$ and*

$$\mathbf{E}[U_{n+1} - U_n \mid \mathcal{F}_n] \leq -g(V_n) \qquad \text{on } \{\tau_d > n\}.$$

Then for any convex in a neighborhood of ∞ function $f \in \mathcal{G}$ satisfying

$$\overline{\lim_{y\to\infty}} \frac{f(2y)}{f(y)} \le c_f,$$

for some positive constant c_f, and

$$\liminf_{y\to\infty} \frac{g(y)}{f' \circ f^{-1} \circ h(y)} > 0,$$

there exists a positive constant c such that, for all $x \ge d$

$$\mathbf{E}[f(\tau_d) \,|\, V_0 = x] \le ch(x).$$

Morever, the following lemma is useful in proving Lemma 4.6.1. Define

$$\mathfrak{E}_2^\gamma(u) = \{|B_n| \le u^\gamma, \forall 1 \le n < K_\beta^\gamma(u)\}. \tag{4.6.15}$$

Lemma 4.6.6. *Suppose that Assumption 4.2.1 holds. Let v be fixed such that $|v| > 1$. For any $\beta + \gamma > 1$ and any $\epsilon > 0$ there exists an u_0 sufficiently large so that, for all $u \ge u_0$,*

$$\mathbf{P}((\mathfrak{E}_2^\gamma(u))^c \,|\, X_0 = vu^\beta) \le \epsilon |v| u^{-(1-\beta)\alpha}.$$

Proof of Lemma 4.6.1. We want to apply Result 4.6.1. Set $f(y) = y^{\alpha+L}$, $h(y) = y^{\underline{\alpha}}$ with $\underline{\alpha} = \lfloor \alpha - \epsilon \rfloor$, and $U_n = h(Y_n)$. Using the binomial formula, we have that

$$\mathbf{E}[U_{n+1} - U_n \,|\, \mathcal{F}_n] \le (\mathbf{E}[A_1^{\underline{\alpha}}] - 1)Y_n^{\underline{\alpha}} + c_1 Y_n^{\underline{\alpha}-1}, \qquad \text{on } \{Y_n \ge 1\},$$

for some positive constant c_1 depending on the first $(\underline{\alpha} - 1)$-st moments of both A_1 and B_1. Using the fact that $\underline{\alpha} < \alpha$ and the moment generating function of $\log A_1$ is convex, we have $\mathbf{E}[A_1^{\underline{\alpha}}] < 1$. Thus, there exists a sufficiently large d' such that, on $\{Y_n > d'\}$,

$$\mathbf{E}[U_{n+1} - U_n \,|\, \mathcal{F}_n] \le (\mathbf{E}[A_1^{\underline{\alpha}}] - 1)Y_n^{\underline{\alpha}} + c_1 Y_n^{\underline{\alpha}-1} \le -c_2 Y_n^{\underline{\alpha}} = -g(Y_n),$$

where $g(y) = c_2 y^{\underline{\alpha}} = c_2 h(y)$, and c_2 is a positive constant depending on d'. Obviously, $f \in \mathcal{G}$ is convex, and $f(2y)/f(y) \le 2^{\alpha+L}$. Moreover, setting $\bar{\alpha} = \alpha + L$ we have that

$$\frac{g(y)}{f' \circ f^{-1} \circ h(y)} = \frac{c_2 h(y)}{\bar{\alpha} h(y)^{(\bar{\alpha}-1)/\bar{\alpha}}} = \frac{c_2}{\bar{\alpha}} h(y)^{1-(\bar{\alpha}-1)/\bar{\alpha}} \to \infty,$$

since $\bar{\alpha} > \alpha > 1$. In view of these, we can apply Result 4.6.1 and obtain that

$$\mathbf{E}[\tilde{\tau}^{\alpha+L} \,|\, Y_0 = x] \le c_3 x^{\lfloor \alpha-\epsilon \rfloor}, \qquad \text{for all } x \ge d', \tag{4.6.16}$$

for some positive constant c_3, where $\tilde{\tau} = \inf\{n \ge 1\colon Y_n \le d'\}$. W.l.o.g. we assume that $d' \ge d$. Using Minkowski's inequality we obtain that

$$\begin{aligned}
&\mathbf{E}[\tau^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} \\
&= \mathbf{E}[(\tilde{\tau} + \tau - \tilde{\tau})^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} \\
&\le \mathbf{E}[\tilde{\tau}^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} + \mathbf{E}[(\tau - \tilde{\tau})^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} \\
&\le \mathbf{E}[\tilde{\tau}^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} + \sup_{y \in [0,d']} \mathbf{E}[\tau^{\alpha+L} \,|\, Y_0 = y]^{1/(\alpha+L)} \\
&\le \mathbf{E}[\tilde{\tau}^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} + \sup_{y \in [0,d']} \mathbf{E}[t^{\tau} \,|\, Y_0 = y]^{1/(\alpha+L)} + \mathcal{O}(1),
\end{aligned}$$

as $x \to \infty$, where, by following the arguments as in the proof of Lemma 4.2.1 above, t can be chosen such that

$$\sup_{y \in [0,d']} \mathbf{E}[t^{\tau} \,|\, Y_0 = y]^{1/(\alpha+L)} < \infty.$$

For this choice of t, we have that

$$\mathbf{E}[\tau^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} \le \mathbf{E}[\tilde{\tau}^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} + \mathcal{O}(1), \qquad \text{as } x \to \infty. \tag{4.6.17}$$

Using (4.6.16) and (4.6.17), there exists a $c > 0$ such that $\mathbf{E}[\tau^{\alpha+L} \,|\, Y_0 = x] \le c x^{\lfloor \alpha-\epsilon \rfloor}$ for sufficiently large x. $\square$

Proof of Lemma 4.6.2. We prove first the statements associated with $\mathscr{G}_+(u)$. As $\mathbb{1}_{\{Z_{T(u^\beta)} > 0\}} \xrightarrow{\text{a.s.}} \mathbb{1}_{\{Z > 0\}}$ under $\mathbf{P}^\alpha$, it is sufficient to show that

$$\lim_{u \to \infty} u^{(1-\beta)\alpha} \mathbf{P}\left(\sum_{k=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_k > u \;\middle|\; \frac{X_{T(u^\beta)}}{u^\beta} = v \right) = C_{+,\infty} v^\alpha, \qquad \text{for } v > 1.$$

Noting

$$\left| \frac{X_n}{X_{n-1}} \right| \le A_n + \frac{|B_n|}{|X_{n-1}|} < A_n + |B_n| u^{-\gamma}, \qquad \text{for } T(u^\beta) < n < K_\beta^\gamma(u),$$

we obtain that, for $\delta > 0$ and $v \geq 1$

$$\begin{aligned}
&\mathbf{P}\left(\sum_{k=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_k > u \,\middle|\, \frac{X_{T(u^\beta)}}{u^\beta} = v\right) \\
&\leq \mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_k^{(u)}} > \frac{u^{1-\beta}}{v}\right) = \mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_k} + \sum_{k=0}^{\infty} e^{S_k^{(u)}} - \sum_{k=0}^{\infty} e^{S_k} > \frac{u^{1-\beta}}{v}\right) \\
&\leq \mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_k} > \frac{u^{1-\beta}}{v} - \delta\right) + \mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_k^{(u)}} - \sum_{k=0}^{\infty} e^{S_k} > \delta\right), \qquad (4.6.18)
\end{aligned}$$

where $S_n^{(u)} = \sum_{i=1}^n \log(A_i + |B_i| u^{-\gamma})$. Note that

$$\mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_k} > \frac{u^{1-\beta}}{v} - \delta\right) \sim C_{+,\infty} \left(\frac{u^{1-\beta}}{v}\right)^{-\alpha}. \qquad (4.6.19)$$

Moreover, using the Markov's inequality and the fact that $S_n^{(u)} \geq S_n$ we obtain that

$$\begin{aligned}
&u^{(1-\beta)\alpha}\mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_n^{(u)}} - \sum_{k=0}^{\infty} e^{S_n} > \delta\right) \\
&\leq \delta^{-1} u^{(1-\beta)\alpha} \mathbf{E}\left[\sum_{k=0}^{\infty} e^{S_n^{(u)}} - \sum_{k=0}^{\infty} e^{S_n}\right] \\
&= \delta^{-1} u^{(1-\beta)\alpha} \left(\sum_{k=0}^{\infty} \mathbf{E}[A_1 + |B_1| u^{-\gamma}]^k - \sum_{k=0}^{\infty} \mathbf{E}[A_1]^k\right) \\
&= \delta^{-1} u^{(1-\beta)\alpha} \left(\frac{1}{1 - \mathbf{E}A_1 - u^{-\gamma}\mathbf{E}|B_1|} - \frac{1}{1 - \mathbf{E}A_1}\right) \\
&= \delta^{-1} u^{(1-\beta)\alpha} \left(\frac{u^{-\gamma}\mathbf{E}|B_1|}{(1 - \mathbf{E}A_1 - u^{-\gamma}\mathbf{E}|B_1|)(1 - \mathbf{E}A_1)}\right) \\
&= \mathcal{O}(u^{(1-\beta)\alpha - \gamma}),
\end{aligned}$$

where in the final step we use the fact that $\mathbf{E}A_1 < 1$. By choosing β sufficiently close to 1 so that $(1-\beta)\alpha < \gamma$, we have that

$$\mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_k^{(u)}} - \sum_{k=0}^{\infty} e^{S_k} > \delta\right) = o(u^{-(1-\beta)\alpha}). \qquad (4.6.20)$$

Using (4.6.18)–(4.6.20), an upper bound is given by, for $v > 1$

$$\overline{\lim_{u\to\infty}} u^{(1-\beta)\alpha}\mathbf{P}\left(\sum_{k=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_k > u \,\middle|\, \frac{X_{T(u^\beta)}}{u^\beta} = v\right) \le C_{+,\infty} v^\alpha. \tag{4.6.21}$$

Next we show the corresponding lower bound. By the Markov property we obtain that

$$\mathbf{P}\left(\sum_{k=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_k > u \,\middle|\, \frac{X_{T(u^\beta)}}{u^\beta} = v\right) = \mathbf{P}\left(\sum_{k=0}^{K_\beta^\gamma(u)-1} X_k > u \,\middle|\, \frac{X_0}{u^\beta} = v\right). \tag{4.6.22}$$

Note that, on the event $\{X_0 \ge u^\beta\}$

$$\left|\frac{X_n}{X_{n-1}}\right| \ge \left(A_n - \frac{|B_n|}{|X_{n-1}|}\right)^+ > (A_n - u^{-\gamma}|B_n|)^+, \tag{4.6.23}$$

for all $n < K_\beta^\gamma(u)$. This implies that

$$e^{\underline{S}^{(u)}_{K_\beta^\gamma(u)}} \le \left|\frac{X_{K_\beta^\gamma(u)}}{X_0}\right| \le \frac{u^{\gamma-\beta}}{v}, \qquad \text{where } \underline{S}_n^{(u)} = \sum_{i=1}^n \log(A_i - u^{-\gamma}|B_i^*|)^+,$$

and hence

$$K_\beta^\gamma(u) \ge \inf\{n \ge 1 \colon \underline{S}_n^{(u)} \le -\log v - (\beta-\gamma)\log u\} = K'(u). \tag{4.6.24}$$

Recall that $\mathfrak{E}_2^\gamma(u) = \{|B_n| \le u^\gamma, \forall 1 \le n < K_\beta^\gamma(u)\}$. In view of (4.6.22)–(4.6.24), we have that, for $\delta > 0$,

$$\begin{aligned}
&\mathbf{P}\left(\sum_{k=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_k > u \,\middle|\, \frac{X_{T(u^\beta)}}{u^\beta} = v\right) \\
&= \mathbf{P}\left(\sum_{k=0}^{K_\beta^\gamma(u)-1} X_k > u \,\middle|\, \frac{X_0}{u^\beta} = v\right) \\
&\ge \mathbf{P}\left(\sum_{k=0}^{K_\beta^\gamma(u)-1} X_k > u,\ \mathfrak{E}_2^\gamma(u) \,\middle|\, \frac{X_0}{u^\beta} = v\right)
\end{aligned}$$

$$\begin{aligned}
&\geq \mathbf{P}\left(\sum_{k=0}^{K_\beta^\gamma(u)-1} e^{\underline{S}_k^{(u)}} > \frac{u^{1-\beta}}{v},\ \mathfrak{E}_2^\gamma(u) \,\middle|\, \frac{X_0}{u^\beta} = v\right) \\
&\geq \mathbf{P}\left(\sum_{k=0}^{K_\beta^\gamma(u)-1} e^{\underline{S}_k^{(u)}} > \frac{u^{1-\beta}}{v} \,\middle|\, \frac{X_0}{u^\beta} = v\right) \\
&\quad - \mathbf{P}\left(\mathfrak{E}_2^\gamma(u)^c \,\middle|\, \frac{X_0}{u^\beta} = v\right) \\
&= \mathbf{P}\left(\sum_{k=0}^{K_\beta^\gamma(u)-1} e^{\underline{S}_k^{(u)}} > \frac{u^{1-\beta}}{v} \,\middle|\, \frac{X_0}{u^\beta} = v\right) + o(u^{\alpha(1-\beta)})v^\alpha,
\end{aligned}$$

where in the last inequality we have used Lemma 4.6.6 above. Thus, it is sufficient to consider

$$\begin{aligned}
&\mathbf{P}\left(\sum_{k=0}^{K_\beta^\gamma(u)-1} e^{\underline{S}_k^{(u)}} > \frac{u^{1-\beta}}{v}\right) \geq \mathbf{P}\left(\sum_{k=0}^{K'(u)-1} e^{\underline{S}_k^{(u)}} > \frac{u^{1-\beta}}{v}\right) \\
&\geq \mathbf{P}\left(\sum_{k=0}^{\infty} e^{\underline{S}_k^{(u)}} > \frac{u^{1-\beta}}{v} + \delta\right) - \mathbf{P}\left(\sum_{k=K'(u)}^{\infty} e^{\underline{S}_k^{(u)}} > \delta\right) \\
&\geq \mathbf{P}\left(\sum_{k=0}^{\infty} e^{\underline{S}_k^{(u)}} > \frac{u^{1-\beta}}{v} + \delta\right) - \delta^{-1}\mathbf{E}\left[\sum_{k=K'(u)}^{\infty} e^{\underline{S}_k^{(u)}}\right] \\
&= \mathbf{P}\left(\sum_{k=0}^{\infty} e^{\underline{S}_k^{(u)}} > \frac{u^{1-\beta}}{v} + \delta\right) - \delta^{-1}\mathbf{E}\left[e^{S_{K'(u)}} \sum_{k=0}^{\infty} e^{\underline{S}_{k+K'(u)}^{(u)} - \underline{S}_{K'(u)}^{(u)}}\right] \\
&\geq \mathbf{P}\left(\sum_{k=0}^{\infty} e^{\underline{S}_k^{(u)}} > \frac{u^{1-\beta}}{v} + \delta\right) - \frac{u^{\gamma-\beta}}{\delta v}\mathbf{E}\left[\sum_{k=0}^{\infty} e^{\underline{S}_k^{(u)}}\right] \\
&\geq \mathbf{P}\left(\sum_{k=0}^{\infty} e^{\underline{S}_k^{(u)}} > \frac{u^{1-\beta}}{v} + \delta\right) - \frac{u^{\gamma-\beta}}{\delta}\mathbf{E}\left[\sum_{k=0}^{\infty} e^{\underline{S}_k^{(u)}}\right], \qquad (4.6.25)
\end{aligned}$$

where in the last inequality we use the fact that $1 < v = X_0/u^\beta$. For the second term in (4.6.25), we have that

$$\mathbf{E}\left[\sum_{k=0}^{\infty} e^{\underline{S}_k^{(u)}}\right] \leq \mathbf{E}\left[\sum_{k=0}^{\infty} e^{S_k}\right] < \infty,$$

and hence,

$$u^{(1-\beta)\alpha}\delta^{-1}u^{\gamma-\beta}\mathbf{E}\left[\sum_{k=0}^{\infty} e^{\underline{S}_k^{(u)}}\right] = o(1), \qquad \text{for } \beta > (\alpha+\gamma)/(\alpha+1). \quad (4.6.26)$$

Therefore, it remains to consider the first term in (4.6.25). Note that

$$\begin{aligned}
&\mathbf{P}\left(\sum_{k=0}^{\infty} e^{\underline{S}_k^{(u)}} > \frac{u^{1-\beta}}{v} + \delta\right) \\
&\geq \mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_k} > \frac{u^{1-\beta}}{v} + 2\delta\right) - \mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_k} - \sum_{k=0}^{\infty} e^{\underline{S}_k^{(u)}} > \delta\right) \\
&\geq \mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_k} > \frac{u^{1-\beta}}{v} + 2\delta\right) - \delta^{-1}\mathbf{E}\left[\sum_{k=0}^{\infty} e^{S_n} - \sum_{k=0}^{\infty} e^{\underline{S}_n^{(u)}}\right] \\
&= \mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_k} > \frac{u^{1-\beta}}{v} + 2\delta\right) \\
&\quad - \delta^{-1}\left(\sum_{k=0}^{\infty}(\mathbf{E}A_1)^k - \sum_{k=0}^{\infty}(\mathbf{E}(A_1 - u^{-\gamma}|B_1^*|)^+)^k\right) \\
&= \mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_k} > \frac{u^{1-\beta}}{v} + 2\delta\right) \\
&\quad - \delta^{-1}\left(\frac{1}{1-\mathbf{E}A_1} - \frac{1}{1-\mathbf{E}(A_1 - u^{-\gamma}|B_1|)^+}\right) \\
&\geq \mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_k} > \frac{u^{1-\beta}}{v} + 2\delta\right) - \delta^{-1}\frac{\mathbf{E}A_1 - \mathbf{E}(A_1 - u^{-\gamma}|B_1|)^+}{(1-\mathbf{E}A_1)(1-\mathbf{E}(A_1 - u^{-\gamma}|B_1|)^+)} \\
&\geq \mathbf{P}\left(\sum_{k=0}^{\infty} e^{S_k} > \frac{u^{1-\beta}}{v} + 2\delta\right) - \delta^{-1}\frac{u^{-\gamma}\mathbf{E}|B_1|}{(1-\mathbf{E}A_1)(1-\mathbf{E}(A_1 - u^{-\gamma}|B_1|)^+)}.
\end{aligned} \quad (4.6.27)$$

In view of (4.6.25)–(4.6.27), we have that

$$\liminf_{u\to\infty} u^{(1-\beta)\alpha}\mathbf{P}\left(\sum_{k=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_k > u \,\middle|\, \frac{X_{T(u^\beta)}}{u^\beta} = v\right) \geq C_{+,\infty}v^\alpha.$$

Combining this with (4.6.21) we have that

$$\lim_{u\to\infty} u^{(1-\beta)\alpha}\mathbf{P}\left(\sum_{k=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_k > u \,\middle|\, \frac{X_{T(u^\beta)}}{u^\beta}\right)\left(\frac{X_{T(u^\beta)}}{u^\beta}\right)^{-\alpha} = C_{+,\infty},$$

$\mathbf{P}^\alpha$-almost surely.

Next we show the boundedness of $\mathscr{G}_+(u)$. Using (4.6.21), for $\epsilon > 0$, there exists $U(\epsilon)$ (independent of v) such that

$$\left(\frac{u^{(1-\beta)}}{v}\right)^\alpha \mathbf{P}\left(\sum_{k=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_k > u \,\middle|\, \frac{X_{T(u^\beta)}}{u^\beta} = v\right) \le C_{+,\infty} + \epsilon,$$

for all $u^{(1-\beta)} \ge vU(\epsilon)$. Moreover, for all $0 < u^{(1-\beta)} < vU(\epsilon)$

$$u^{(1-\beta)\alpha}\mathbf{P}\left(\sum_{k=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_k > u \,\middle|\, \frac{X_{T(u^\beta)}}{u^\beta} = v\right) \le u^{(1-\beta)\alpha} \le v^\alpha U(\epsilon)^\alpha. \quad (4.6.28)$$

Thus

$$u^{(1-\beta)\alpha}\mathbf{P}\left(\sum_{k=T(u^\beta)}^{K_\beta^\gamma(u)-1} X_k > u \,\middle|\, \frac{X_{T(u^\beta)}}{u^\beta} = v\right) \le \max\{C_{+,\infty} + \epsilon, U(\epsilon)^\alpha\}v^\alpha,$$

for all $u > 0$. This implies that $\mathscr{G}_+(u) \le \max\{C_{+,\infty} + \epsilon, U(\epsilon)^\alpha\} < \infty$.

Finally, we show the statements involved with $\mathscr{G}_-$. By the Markov property, it is sufficient to show that, for any arbitrary $\epsilon > 0$ and $v < -1$

$$\overline{\lim_{u\to\infty}}\, u^{(1-\beta)\alpha}\mathbf{P}\left(\sum_{k=0}^{K_\beta^\gamma(u)-1} X_k > u \,\middle|\, \frac{X_0}{u^\beta} = v\right) \le \epsilon|v|^\alpha.$$

Recall

$$\mathfrak{E}_2^\gamma(u) = \{|B_n| \le u^\gamma, \forall 1 \le n < K_\beta^\gamma(u)\},$$

was defined in (4.6.15). We have that

$$\mathbf{P}\left(\sum_{k=0}^{K_\beta^\gamma(u)-1} X_k > u \,\middle|\, \frac{X_0}{u^\beta} = v\right)$$

$$\leq \mathbf{P}\left(\sum_{k=0}^{K_\beta^\gamma(u)-1} X_k > u,\, \mathfrak{E}_2^\gamma(u) \,\middle|\, \frac{X_0}{u^\beta} = v\right) + \mathbf{P}\left(\mathfrak{E}_2^\gamma(u)^c \,\middle|\, X_0 = vu^\beta\right)$$
$$= \mathbf{P}\left(\mathfrak{E}_2^\gamma(u)^c \,\middle|\, X_0 = vu^\beta\right) = o(u^{-(1-\beta)\alpha})|v|,$$

thanks to Lemma 4.6.6. The boundedness of $\mathscr{G}_u^-$ follows using similar arguments as in (4.6.28). $\square$

Remark 4.4. Using similar arguments as in the proof of Lemma 4.6.2, one can show that

$$\lim_{u\to\infty} u^{(1-\beta)\alpha}\mathbf{P}^{\mathscr{D}}\left(\sum_{n=T(u^\beta)}^{K_\beta^\gamma(u)-1} |X_n| > u \,\middle|\, \mathcal{F}_{T(u^\beta)}\right)\left|\frac{X_{T(u^\beta)}}{u^\beta}\right|^{-\alpha} = C_{+,\infty}.$$

As a consequence of this result, we have that

$$u^\alpha \mathbf{P}(\sum_{n=0}^{r_1-1} |X_n| > u) \to C_{+,\infty}\mathbf{E}^\alpha[|Z|^\alpha \mathbb{1}_{\{r_1=\infty\}}$$

as $u \to \infty$.

Proof of Lemma 4.6.3. Note that $Z^+_{T(u^\beta)}\mathbb{1}_{\{T(u^\beta)<\tau_d\}}$ and $Z^-_{T(u^\beta)}\mathbb{1}_{\{T(u^\beta)<\tau_d\}}$ are bounded by $|Z_{T(u^\beta)}\mathbb{1}_{\{T(u^\beta)<\tau_d\}}|$, for which we have that

$$|Z_{T(u^\beta)}\mathbb{1}_{\{T(u^\beta)<\tau_d\}}| \leq |X_0| + \sum_{n=1}^{T(u^\beta)} |B_n|e^{-S_n}\mathbb{1}_{\{T(u^\beta)<\tau_d\}}$$
$$\leq |X_0| + \sum_{n=1}^{\infty} |B_n|e^{-S_n}\mathbb{1}_{\{n<\tau_d\}} = \bar{Z}.$$

Moreover, using Minkowski's inequality we have that

$$\mathbf{E}^\alpha[\bar{Z}^\alpha]^{1/\alpha} \leq |X_0| + \sum_{n=1}^{\infty} \mathbf{E}^\alpha\left[|B_n|e^{-\alpha S_n}\mathbb{1}_{\{n<\tau_d\}}\right]^{1/\alpha}$$
$$= |X_0| + \sum_{n=1}^{\infty} \mathbf{E}\left[|B_n|\mathbb{1}_{\{n<\tau_d\}}\right]^{1/\alpha}$$
$$= |X_0| + (\mathbf{E}|B_1|)^{1/\alpha}\sum_{n=1}^{\infty} \mathbf{P}(\tau_d > n)^{1/\alpha} < \infty,$$

where in the last inequality we use the fact that $\mathbf{P}(\tau_d > n)$ decays exponentially in n. $\square$

Proof of Lemma 4.6.4. Recall that $\tau = \inf\{n \geq 1 \colon Y_n \leq d\}$. Using Minkowski's inequality we obtain that

$$\begin{aligned}
&\mathbf{E}[r^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} = \mathbf{E}[(\tau + r - \tau)^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} \\
&\leq \mathbf{E}[\tau^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} + \mathbf{E}[(r-\tau)^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} \\
&\leq \mathbf{E}[\tau^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} + \sup_{y \in [0,d]} \mathbf{E}[r^{\alpha+L} \,|\, Y_0 = y]^{1/(\alpha+L)} \\
&\leq \mathbf{E}[\tau^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} + \sup_{y \in [0,d]} \mathbf{E}[t^{r} \,|\, Y_0 = y]^{1/(\alpha+L)} + \mathcal{O}(1),
\end{aligned}$$

as $x \to \infty$, where, by following the arguments as in the proof of Lemma 4.2.1, t can be chosen such that $\sup_{y \in [0,d]} \mathbf{E}[t^r \,|\, Y_0 = y] < \infty$. For this choice of t, we have that

$$\mathbf{E}[r^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} \leq \mathbf{E}[\tau^{\alpha+L} \,|\, Y_0 = x]^{1/(\alpha+L)} + \mathcal{O}(1), \qquad \text{as } x \to \infty.$$

Finally, using Lemma 4.6.1 above we have $\mathbf{E}[r^{\alpha+L} \,|\, Y_0 = x] \leq c x^{\lfloor \alpha - \epsilon \rfloor}$ for sufficiently large x. $\square$

Proof of Lemma 4.6.5. Note that both $|Z^{\alpha}_{T(u)}| \mathbb{1}_{\{T(u) < r_1\}}$ and $|Z^{\alpha}_n| \mathbb{1}_{\{n \leq r_1\}}$ are bounded by

$$\bar{Z} = |X_0| + \sum_{n=1}^{\infty} |B_n| e^{-S_n} \mathbb{1}_{\{n < r_1\}},$$

whose α-th moment is finite thanks to Lemma 4.6.3. Moreover, note that $\{X_n\}_{n \geq 0}$ is transient in the α-shifted measure (cf. Lemma 4.2.2 above), and hence, $T(u) < \infty$ a.s. Applying a change of measure argument, we obtain that

$$\begin{aligned}
&u^{\alpha} \mathbf{P}(T(u) < r_1) \\
&= u^{\alpha} \mathbf{E}^{\alpha}[e^{-\alpha S_{T(u)}} \mathbb{1}_{\{T(u) < r_1\}}] = \mathbf{E}^{\alpha}\left[|Z_{T(u)}|^{\alpha} \left| \frac{X_{T(u)}}{u} \right|^{-\alpha} \mathbb{1}_{\{T(u) < r_1\}} \right] \\
&= \mathbf{E}^{\alpha}\left[|Z_n|^{\alpha} \mathbb{1}_{\{n \leq T(u)\}} \mathbb{1}_{\{n \leq r_1\}} \mathbf{E}^{\alpha}\left[\left| \frac{X_{T(u)}}{u} \right|^{-\alpha} \,\middle|\, \mathcal{F}_n \right] \right] \\
&\quad + \mathbf{E}^{\alpha}\left[\left(|Z_{T(u)}|^{\alpha} \mathbb{1}_{\{T(u) < r_1\}} - |Z_n|^{\alpha} \mathbb{1}_{\{n \leq T(u)\}} \mathbb{1}_{\{n \leq r_1\}} \right) \left| \frac{X_{T(u)}}{u} \right|^{-\alpha} \right]
\end{aligned}$$

$$= \textbf{(III.1)} + \textbf{(III.2)},$$

where $\{\mathcal{F}_n\}_{n\geq 0}$ is the natural filtration. Since $(X_{T(u)}/u)^{-\alpha} \leq 1$ and $T(u) \to \infty$ a.s. as $u \to \infty$,

$$\begin{aligned}
&\lim_{n\to\infty} \lim_{u\to\infty} \textbf{(III.2)} \\
&\leq \lim_{n\to\infty} \lim_{u\to\infty} \mathbf{E}^{\alpha} \left[|Z_{T(u)}|^{\alpha} \mathbb{1}_{\{T(u)<r_1\}} - |Z_n|^{\alpha} \mathbb{1}_{\{n\leq T(u)\}} \mathbb{1}_{\{n\leq r_1\}} \right] \\
&= \lim_{n\to\infty} \mathbf{E}^{\alpha} \left[\lim_{u\to\infty} \left(|Z_{T(u)}|^{\alpha} \mathbb{1}_{\{T(u)<r_1\}} - |Z_n|^{\alpha} \mathbb{1}_{\{n\leq T(u)\}} \mathbb{1}_{\{n\leq r_1\}} \right) \right] \\
&= \lim_{n\to\infty} \mathbf{E}^{\alpha} \left[|Z|^{\alpha} \mathbb{1}_{\{r_1=\infty\}} - |Z_n|^{\alpha} \mathbb{1}_{\{n\leq r_1\}} \right] = 0.
\end{aligned}$$

It remains to consider **(III.1)**. Note that, given $\mathcal{F}_n$, $n \leq T(u)$, the random variable $\log |X_{T(u)}| - \log u$ converges in distribution to some positive random variable $\mathfrak{X}$—which is independent of $\mathcal{F}_n$, $n \leq T(u)$—as $u \to \infty$, under the α-shifted measure (cf. e.g. Theorem 3.8 of [36]). Hence we have that

$$\lim_{u\to\infty} \mathbf{E}^{\alpha} \left[\left| \frac{X_{T(u)}}{u} \right|^{-\alpha} \middle| \mathcal{F}_n \right] \mathbb{1}_{\{n\leq T(u)\}} = \mathbf{E}\left[e^{-\alpha \mathfrak{X}} \right].$$

Moreover, using the dominated convergence theorem and the fact

$$\mathbb{1}_{\{n\leq T(u)\}} \mathbf{E}^{\alpha}[|X_{T(u)}/u|^{-\alpha} \,|\, \mathcal{F}_n] \leq 1,$$

we obtain that

$$\begin{aligned}
&\lim_{n\to\infty} \lim_{u\to\infty} \textbf{(III.1)} \\
&= \lim_{n\to\infty} \mathbf{E}^{\alpha} \left[|Z_n|^{\alpha} \mathbb{1}_{\{n\leq r_1\}} \lim_{u\to\infty} \mathbf{E}^{\alpha} \left[\left| \frac{X_{T(u)}}{u} \right|^{-\alpha} \middle| \mathcal{F}_n \right] \mathbb{1}_{\{n\leq T(u)\}} \right] \\
&= \mathbf{E}\left[e^{-\alpha \mathfrak{X}} \right] \lim_{n\to\infty} \mathbf{E}^{\alpha} \left[|Z_n|^{\alpha} \mathbb{1}_{\{n\leq r_1\}} \right] = \mathbf{E}\left[e^{-\alpha \mathfrak{X}} \right] \mathbf{E}^{\alpha} \left[|Z|^{\alpha} \mathbb{1}_{\{r_1=\infty\}} \right].
\end{aligned}$$

□

Proof of Lemma 4.6.6. To begin with, we write, for some $\delta > 0$,

$$\begin{aligned}
\mathbf{P}((\mathfrak{C}_2^{\gamma}(u))^c | X_0 = vu^{\beta}) &= \mathbf{P}(\exists n < K_{\beta}^{\gamma}(u) \colon |B_n| > u^{\gamma} | X_0 = vu^{\beta}) \\
&\leq \mathbf{P}(\exists n < \tau_d \colon |B_n| > u^{\gamma} | X_0 = vu^{\beta}) \\
&\leq \mathbf{P}(\exists n \leq u^{\delta} \colon |B_n| > u^{\gamma}) + \mathbf{P}(\tau_d \geq u^{\delta} | X_0 = vu^{\beta})
\end{aligned}$$

$$= \textbf{(II.1)} + \textbf{(II.2)}.$$

To bound **(II.1)**, we have that

$$\textbf{(II.1)} \le u^{\delta}\mathbf{P}(|B_1| > u^{\gamma}) \le u^{\delta-\alpha\gamma}\mathbf{E}|B_1|^{\alpha} = o(u^{-(1-\beta)\alpha}),$$

for $(1-\beta)\alpha + \delta - \alpha\gamma < 0$. Since $\textbf{(II.2)} \le u^{-(\alpha+L)\delta}\mathbf{E}[\tau_d^{\alpha+L}|X_0 = vu^{\beta}]$, it is sufficient to bound $\mathbf{E}[\tau_d^{\alpha+L}|X_0 = vu^{\beta}]$. Recall $\{Y_n\}_{n\ge 0}$ is the $\mathbb{R}_+$-valued Markov chain defined by $Y_{n+1} = A_{n+1}Y_n + |B_{n+1}|$, for $n \ge 0$, and $\tau = \inf\{n \ge 1 \colon Y_n \le d\}$. Note that $\mathbf{E}[\tau_d^{\alpha+L}|X_0 = vu^{\beta}] \le \mathbf{E}[\tau^{\alpha+L}|Y_0 = |v|u^{\beta}]$. Combining this with Proposition 4.6.1, we conclude that there exist c and u_0 such that

$$\textbf{(II.2)} \le u^{-(\alpha+L)\delta}\mathbf{E}[\tau^{\alpha+L}|Y_0 = |v|u^{\beta}] \le c|v|u^{\beta}u^{-(\alpha+L)\delta}, \qquad \forall u \ge u_0.$$

Thus, the proposition is proved by setting $L = L(\delta, \alpha, \beta)$ be sufficiently large. Combining the estimates above we conclude the proof. □

4.7 Proofs of Sections 4.3.2 and 4.3.3

Again, we briefly describe our strategy of proof before diving into the technicalities. Define $\bar{X}'_n = \{\bar{X}'_n(t), t \in [0,1]\}$, where

$$\bar{X}'_n(t) = \frac{1}{n}\sum_{i=1}^{N(nt)} X'_i \quad \text{and} \quad X'_i = \sum_{j=r_{i-1}}^{r_i - 1} X_j, \tag{4.7.1}$$

where $\{r_i\}_{i\ge 0}$ is the sequence of regeneration times as in Remark 4.2, and

$$N(s) = \sup\{j \ge 0 \colon r_j - 1 \le s\}. \tag{4.7.2}$$

Thanks to Theorem 4.1 of [105] and Theorem 4.3.1 above, we are able to establish an asymptotic equivalence between $\bar{X}'_n$ and some random walk $\bar{W}_n$ that will be specified below. This allows us to provide a large deviations result for $\bar{X}'_n$. In both the one-sided and the two-sided case, we will show that the residual process $\bar{X}_n - \bar{X}'_n$ is negligible in an asymptotic sense.

We state here three lemmata that will play key roles in the proofs of Theorems 4.3.2 and 4.3.3. Let $\bar{W}_n = \{\bar{W}_n(t), t \in [0,1]\}$ be such that

$$\bar{W}_n(t) = \frac{1}{n}\sum_{i=1}^{\lfloor nt/\mathbf{E}r_1 \rfloor} X'_i, \tag{4.7.3}$$

where X_i' is as in (4.7.1). We begin with stating an asymptotic equivalence between $\bar{X}_n'$ and $\bar{W}_n$, however, w.r.t. the J_1-topology, which is stronger than the M_1'-topology introduced in the beginning of Section 4.3.2. Let d_{J_1} denote the Skorokhod J_1 metric on $\mathbb{D}$, which is defined by

$$d_{J_1}(\xi_1, \xi_2) = \inf_{\lambda \in \Lambda} ||\lambda - id||_\infty \vee ||\xi_1 \circ \lambda - \xi_2||_\infty, \qquad \xi_1, \xi_2 \in \mathbb{D},$$

where id denotes the identity mapping, $||\cdot||_\infty$ denotes the uniform metric, that is, $\|x\|_\infty = \sup_{t\in[0,1]} |x(t)|$, and Λ denotes the set of all strictly increasing, continuous bijections from $[0,1]$ to itself. Moreover, for $j \geq 0$, define

$$\mathbb{D}^\mu_{\leqslant j} = \{\xi \in \underline{\mathbb{D}}^\mu_{\leqslant j} \colon \xi(0) = 0\} \quad \text{and} \quad \mathbb{D}^\mu_{\ll j} = \{\xi \in \underline{\mathbb{D}}^\mu_{\ll j} \colon \xi(0) = 0\}.$$

Lemma 4.7.1. *Consider the metric space $(\mathbb{D}, d_{J_1})$. Suppose that Assumptions 4.2.1 and 4.2.2 hold. For any $j \geq 0$, the following holds.*

1. *If $B_1 \geq 0$ and C_+ as in Theorem 4.3.1 is strictly positive, then the stochastic process $\bar{X}_n'$ is asymptotically equivalent to $\bar{W}_n$ w.r.t. $n^{-j(\alpha-1)}$ and $\mathbb{D}^\mu_{\leqslant j-1}$.*
2. *If C_+ and C_- as in Theorem 4.3.1 satisfy $C_+C_- > 0$, then the stochastic process $\bar{X}_n'$ is asymptotically equivalent to $\bar{W}_n$ w.r.t. $n^{-j(\alpha-1)}$ and $\mathbb{D}^\mu_{\ll j}$.*

Proof. We only show *part 2)*, since *part 1)* can be proved by a similar argument. By Definition 4.2.1, it is sufficient to show, for any $\delta > 0$ and $\gamma > 0$,

$$\begin{aligned}&\overline{\lim_{n\to\infty}}\, n^{j(\alpha-1)}\mathbf{P}(\bar{X}_n' \in (\mathbb{D}\setminus\mathbb{D}^\mu_{\ll j})^{-\gamma}, d_{J_1}(\bar{X}_n', \bar{W}_n) \geq \delta) \\ &= \overline{\lim_{n\to\infty}}\, n^{j(\alpha-1)}\mathbf{P}(\bar{W}_n \in (\mathbb{D}\setminus\mathbb{D}^\mu_{\ll j})^{-\gamma}, d_{J_1}(\bar{X}_n', \bar{W}_n) \geq \delta) = 0. \qquad (4.7.4)\end{aligned}$$

To prove (4.7.4), it is convenient to consider the space of paths on a longer time horizon $[0,2]$. Let $\bar{W}_n$ denote the stochastic process $\{\bar{W}_n(t), t \in [0,2]\}$ over the time horizon $[0,2]$, and $\mathbb{D}^{\mu;[0,2]}_{\ll j}$ denote the space of step functions on $[0,2]$ that corresponds to $\mathbb{D}^\mu_{\ll j}$. Let $d^{[0,2]}_{J_1}$ denote the Skorokhod J_1 metric on $\mathbb{D}^{[0,2]} = \mathbb{D}([0,2], \mathbb{R})$.

Note that $d_{J_1}(\bar{W}_n, \mathbb{D}^\mu_{\ll j}) \geq \gamma$ implies that $d^{[0,2]}_{J_1}(\bar{W}^{[0,2]}_n, \mathbb{D}^{\mu;[0,2]}_{\ll j}) \geq \gamma$, and $d_{J_1}(\bar{X}_n', \mathbb{D}^\mu_{\ll j}) \geq \gamma$ implies that either $d^{[0,2]}_{J_1}(\bar{W}^{[0,2]}_n, \mathbb{D}^{\mu;[0,2]}_{\ll j}) \geq \gamma$ or $2n/\mathbf{E}r_1 \leq N(n)$. Therefore, (4.7.4) is implied by

$$\overline{\lim_{n\to\infty}}\, n^{j(\alpha-1)}\mathbf{P}(d^{[0,2]}_{J_1}(\bar{W}^{[0,2]}_n, \mathbb{D}^{\mu;[0,2]}_{\ll j}) \geq \gamma, d_{J_1}(\bar{X}_n', \bar{W}_n) \geq \delta) = 0. \qquad (4.7.5)$$

To prove (4.7.5), we adopt the construction of a piecewise linear non-decreasing homeomorphism $\bar{\lambda}_n$ from [105, the proof of Theorem 4.1]. Let $t_0 = 0$ and t_i be the i-th jump time of $N(n\cdot)$ and t_L be the last jump time of $N(n\cdot)$. Let $L = (\lfloor n/\mathbf{E}r_1\rfloor - 1) \wedge N(n)$. Define $\bar{\lambda}_n$ in such a way that $\bar{\lambda}_n(t) = \mathbf{E}r_1 N(nt)/n$ on $t_0, \ldots, t_L$, $\bar{\lambda}_n(1) = 1$, and $\bar{\lambda}_n$ is a piecewise linear interpolation in between. For such $\bar{\lambda}_n$, $\bar{W}_n(\bar{\lambda}_n(t)) = \bar{X}'_n(t)$ for all $t \in [0, t_L]$, and hence, $\|\bar{W}_n \circ \bar{\lambda}_n - \bar{X}'_n\|_\infty = \sup_{t\in[t_L,1]} |\bar{W}_n \circ \bar{\lambda}_n(t) - \bar{X}'_n(t)|$. Therefore,

$$
\begin{aligned}
d_{J_1}(\bar{W}_n, \bar{X}'_n) &= \inf_{\lambda\in\Lambda} \|\lambda - id\|_\infty \vee \|\bar{W}_n \circ \lambda - \bar{X}'_n\|_\infty \\
&\le \|\bar{\lambda}_n - id\|_\infty \vee \|\bar{W}_n \circ \bar{\lambda}_n - \bar{X}'_n\|_\infty \\
&= \|\bar{\lambda}_n - id\|_\infty \vee \sup_{t\in[t_L,1]} |\bar{W}_n \circ \bar{\lambda}_n(t) - \bar{X}'_n(t)|. \qquad (4.7.6)
\end{aligned}
$$

Note that the second term can be bounded (with high probability) as follows. For an arbitrary $\epsilon > 0$, consider two cases: $\lfloor n/\mathbf{E}r_1 - n\epsilon\rfloor < N(n) < \lfloor n/\mathbf{E}r_1\rfloor$ and $\lfloor n/\mathbf{E}r_1\rfloor \le N(n) \le \lfloor n/\mathbf{E}r_1 + n\epsilon\rfloor$. Set

$$
W_n = \sum_{i=1}^{\lfloor n/\mathbf{E}r_1\rfloor} X'_i.
$$

If $\lfloor n/\mathbf{E}r_1 - n\epsilon\rfloor < N(n) < \lfloor n/\mathbf{E}r_1\rfloor$, by the construction of $\bar{\lambda}_n$,

$$
\sup_{t\in[t_L,1]} |\bar{W}_n \circ \bar{\lambda}_n(t) - \bar{X}'_n(t)| \le \sup_{s,t\in[1-\epsilon,1]} |\bar{W}_n(s) - \bar{W}_n(t)|. \qquad (4.7.7)
$$

On the other hand, if $\lfloor n/\mathbf{E}r_1\rfloor \le N(n) \le \lfloor n/\mathbf{E}r_1 + n\epsilon\rfloor$,

$$
\sup_{t\in[t_L,1]} |\bar{W}_n \circ \bar{\lambda}_n(t) - \bar{X}'_n(t)| \le \sup_{s,t\in[1,1+\epsilon]} |\bar{W}_n(s) - \bar{W}_n(t)|. \qquad (4.7.8)
$$

From (4.7.7) and (4.7.8), we see that on the event $\{\lfloor n/\mathbf{E}r_1 - n\epsilon\rfloor < N(n) \le \lfloor n/\mathbf{E}r_1 + n\epsilon\rfloor\}$,

$$
\sup_{t\in[t_L,1]} |\bar{W}_n \circ \bar{\lambda}_n(t) - \bar{X}'_n(t)| \le \sup_{s,t\in[1-\epsilon,1+\epsilon]} |\bar{W}_n(s) - \bar{W}_n(t)|. \qquad (4.7.9)
$$

Using (4.7.6) and (4.7.9), we obtain that

$$
\begin{aligned}
&\mathbf{P}(d_{J_1}^{[0,2]}(\bar{W}_n^{[0,2]}, \mathbb{D}_{\ll j}^{\mu;[0,2]}) \ge \gamma, d_{J_1}(\bar{X}'_n, \bar{W}_n) \ge \delta) \\
&\le \mathbf{P}\left(d_{J_1}^{[0,2]}(\bar{W}_n^{[0,2]}, \mathbb{D}_{\ll j}^{\mu;[0,2]}) \ge \gamma, \sup_{s,t\in[1-\epsilon,1+\epsilon]} |\bar{W}_n(s) - \bar{W}_n(t)| \ge \delta\right)
\end{aligned}
$$

$$+\,\mathbf{P}(\{\lfloor n/\mathbf{E}r_1 - n\epsilon\rfloor < N(n) \le \lfloor n/\mathbf{E}r_1 + n\epsilon\rfloor\}^c) + \mathbf{P}(\|\bar{\lambda}_n - id\|_\infty \ge \delta). \tag{4.7.10}$$

Thanks to Cramér's theorem, the second term in (4.7.10) decays geometrically. Moreover, for the last term in (4.7.10), we have that

$$\begin{aligned}
&\mathbf{P}(\|N(n\,\cdot\,)/n - \cdot/\mathbf{E}r_1\|_\infty > \delta)\\
&= \mathbf{P}\left(\sup_{t\in[0,1]} |N(nt)/n - t/\mathbf{E}r_1| > \delta\right)\\
&= 1 - \lim_{m\to\infty} \mathbf{P}\left(\sup_{0\le l\le 2^m}\left|\frac{N(nl/2^m)}{n} - \frac{l}{\mathbf{E}r_1 2^m}\right| \le \delta\right)\\
&= 1 - \lim_{m\to\infty} \mathbf{P}\left(\left|\frac{N(nl/2^m)}{n} - \frac{l}{\mathbf{E}r_1 2^m}\right| \le \delta, \forall 0 \le l \le 2^m\right)\\
&= 1 - \lim_{m\to\infty} \mathbf{P}\left(\frac{N(nl/2^m)}{n} \in \left[\frac{l}{\mathbf{E}r_1 2^m} - \delta, \frac{l}{\mathbf{E}r_1 2^m} + \delta\right], \forall l \le 2^m\right).
\end{aligned}$$

Let $\Delta_i = r_i - r_{i-1}$. Using the fact that $N(n) < k \iff \sum_{i=1}^k \Delta_i > n$, we obtain that

$$\begin{aligned}
&\mathbf{P}(\|N(nt)/n - t/\mathbf{E}r_1\|_\infty > \delta)\\
&= 1 - \lim_{m\to\infty}\\
&\quad \mathbf{P}\left(\sum_{i=1}^{\lfloor (l/(\mathbf{E}r_1 2^m) - \delta)n\rfloor + 1} \Delta_i \le nl/2^m < \sum_{i=1}^{\lfloor (l/(\mathbf{E}r_1 2^m) + \delta)n\rfloor + 1} \Delta_i, \forall 0 \le l \le 2^m\right)\\
&= 1 - \mathbf{P}\left(\sup_{t\in[0,1]} \sum_{i=1}^{\lfloor (t/\mathbf{E}r_1 - \delta)n\rfloor + 1} \Delta_i - nt \le 0, \inf_{t\in[0,1]} \sum_{i=1}^{\lfloor (t/\mathbf{E}r_1 + \delta)n\rfloor + 1} \Delta_i - nt > 0\right)\\
&\le 1 - \mathbf{P}\left(\sup_{t\in[0,1]} \sum_{i=1}^{\lfloor (t/\mathbf{E}r_1 - \delta/2)n\rfloor} \Delta_i - nt < 0, \inf_{t\in[0,1]} \sum_{i=1}^{\lfloor (t/\mathbf{E}r_1 + 3\delta/2)n\rfloor} \Delta_i - nt > 0\right)\\
&= 1 - \mathbf{P}\Bigg(\sup_{t\in[0,1/\mathbf{E}r_1 - \delta/2]} \frac{1}{n}\left(\sum_{i=1}^{\lfloor nt\rfloor} \Delta_i - n\mathbf{E}r_1 t - n\mathbf{E}r_1\delta\right) < 0,\\
&\qquad\qquad \inf_{t\in[\delta,1/\mathbf{E}r_1 + 3\delta/2]} \frac{1}{n}\left(\sum_{i=1}^{\lfloor nt\rfloor} \Delta_i - n\mathbf{E}r_1 t + n\mathbf{E}r_1\delta\right) > 0\Bigg)
\end{aligned}$$

$$
\begin{aligned}
&= 1 - \mathbf{P}\Bigg(\sup_{t\in[0,1/\mathbf{E}r_1-\delta/2]} \frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} \Delta_i - \mathbf{E}r_1 t < \mathbf{E}r_1\delta, \\
&\qquad \inf_{t\in[\delta,1/\mathbf{E}r_1+3\delta/2]} \frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} \Delta_i - \mathbf{E}r_1 t > -\mathbf{E}r_1\delta \Bigg) \\
&\to 0,
\end{aligned}
$$

as $n \to \infty$, at an exponential rate by Mogulskii's theorem.

For the first term in (4.7.10), we have that (see [105, page 21])

$$
\begin{aligned}
&\overline{\lim_{n\to\infty}}\, n^{j(\alpha-1)}\mathbf{P}\Bigg(d_{J_1}^{[0,2]}(\bar{W}_n^{[0,2]}, \mathbb{D}_{\ll j}^{\mu;[0,2]}) \geq \gamma, \\
&\qquad\qquad \sup_{s,t\in[1-\epsilon,1+\epsilon]} |\bar{W}_n(s) - \bar{W}_n(t)| \geq \delta \Bigg) \leq c\epsilon
\end{aligned}
$$

for some $c > 0$, where the intuition behind the asymptotics above is that, given the rare event takes place, the random walk $\bar{W}_n^{[0,2]}$ must have j big jumps and one of them has to occur in the time interval $[1-\epsilon, 1+\epsilon]$. Since the choice of $\epsilon > 0$ was arbitrary, (4.7.4) is proved by letting $\epsilon \to 0$. $\square$

The next two lemmata are useful for future purposes.

Lemma 4.7.2. *For $\xi, \zeta \in \mathbb{D}$, we have that $d_{M_1'}(\xi, \zeta) \leq d_{J_1}(\xi, \zeta)$.*

Recall that Disc(ξ) is the set of discontinuities of $\xi \in \mathbb{D}$ and was defined in (4.3.4).

Lemma 4.7.3. *If $d_{M_1'}(\xi_n, \xi) \to 0$ as $n \to \infty$, then, for each $t \in Disc(\xi)^c$*

$$
\lim_{\delta\downarrow 0} \overline{\lim_{n\to\infty}} \sup_{t_1\in\mathscr{B}_\delta(t)\cap[0,1]} |\xi_n(t_1) - \xi(t_1)| = 0.
$$

Proof. Let $t \in \text{Disc}(\xi)^c$. We first prove the statement for the case where $t \in (0,1)$. Let $\epsilon > 0$ be fixed. Choose $\delta = \delta(\epsilon) > 0$ such that

$$
|\xi(t_1) - \xi(t)| < \epsilon, \qquad \text{for } t_1 \in \mathscr{B}_\delta(t) \subseteq (0,1). \tag{4.7.11}
$$

By the definition of the M_1' convergence, for the given ϵ, there exists n_0, such that $d_{M_1'}(\xi_n, \xi) < (\delta \wedge \epsilon)/8$ for all $n \geq n_0$. Moreover, for each fixed $n \geq n_0$, one can find $(u_n, v_n) \in \Gamma'(\xi_n)$ and $(u, v) \in \Gamma'(\xi)$ such that

$$
\|u_n - u\|_\infty \vee \|v_n - v\|_\infty < (\delta \wedge \epsilon)/4. \tag{4.7.12}
$$

Let $\underline{s}$, s, $\bar{s}$ be such that $v(\underline{s}) = t - \delta/2$, $v(s) = t$ and $v(\bar{s}) = t + \delta/2$. Moreover, by (4.7.12) we have that $v_n(\underline{s}) < t - \delta/4$ and $v_n(\bar{s}) > t + \delta/4$. Thus, for all $t_1 \in (t - \delta/4, t + \delta/4)$ there exists $s_n \in (\underline{s}, \bar{s})$ such that $(u_n(s_n), v_n(s_n)) = (\xi_n(t_1), t_1)$. Combining this with (4.7.11) and (4.7.12), we obtain that

$$\begin{aligned}
|\xi_n(t_1) - \xi(t_1)| &\leq |\xi_n(t_1) - \xi(t)| + |\xi(t_1) - \xi(t)| \\
&= |u_n(s_n) - u(s)| + |\xi(t_1) - \xi(t)| \\
&\leq |u_n(s_n) - u(s_n)| + |u(s_n) - u(s)| + \epsilon \\
&\leq (\delta \wedge \epsilon)/2 + \epsilon + \epsilon < 3\epsilon.
\end{aligned}$$

Finally, the case where $t \in \{0, 1\}$ can be dealt with similarly. □

The reminder of this section is split into two parts that deal with Theorems 4.3.2 and 4.3.3 respectively.

4.7.1 Proofs of Theorem 4.3.2

We consider the case where B_1 is nonnegative. Let us give the "roadmap" of proving Theorem 4.3.2.

- In Corollary 4.7.1 below we establish a sample-path large deviations result for the aggregated process $\bar{X}'_n$ (see (4.7.1) above) by considering a suitably defined random walk together with utilizing Theorem 4.1 of [105]. For the $\mathbb{M}$-convergence in Corollary 4.7.1 we need Lemma 4.7.4 below.
- In Proposition 4.7.1 we show the asymptotic equivalence between the aggregated process $\bar{X}'_n$ and the original process $\bar{X}_n$. Again, one technical lemma, see Lemma 4.7.5 below, is needed.
- Part 1) of Theorem 4.3.2 follows by combining Corollary 4.7.1 with Proposition 4.7.1. Part 2) is a direct consequence of part 1).

Lemma 4.7.4. *For all $j \geq 0$ and all $z \in \mathbb{R}$, the set $\underline{\mathbb{D}}^z_{\leqslant j}$ is closed w.r.t. $(\mathbb{D}, d_{M'_1})$.*

Recall that C^z_j was defined in (4.3.6) for $z \in \mathbb{R}$.

Corollary 4.7.1. *Suppose that Assumptions 4.2.1 and 4.2.2 hold. Moreover, let $B_1 \geq 0$ and C_+ as in Theorem 4.3.1 be strictly positive. For any $j \geq 0$,*

$$n^{j(\alpha-1)}\mathbf{P}(\bar{X}'_n \in \cdot\,) \to (C_+\mathbf{E}r_1)^j C^\mu_j(\,\cdot\,),$$

in $\mathbb{M}(\mathbb{D} \setminus \underline{\mathbb{D}}^\mu_{\leqslant j-1})$ as $n \to \infty$.

Proposition 4.7.1. *Suppose that Assumptions 4.2.1 and 4.2.2 hold. If* $B_1 \geq 0$ *and* C_+ *as in Theorem 4.3.1 is strictly positive, then* $\bar{X}_n$ *is asymptotically equivalent to* $\bar{X}'_n$ *w.r.t.* $(n\mathbf{P}(X'_1 \geq n))^j$ *and* $\underline{\mathbb{D}}^{\mu}_{\leqslant j-1}$.

Proof of Theorem 4.3.2. Part 1) follows by combining Corollary 4.7.1 with Proposition 4.7.1. Part 2) is a direct consequence of part 1). □

Proof of Lemma 4.7.4. We give the proof for the case where $z = 0$, while the proof for $z \neq 0$ follows using the same arguments. The statement is trivial for $\underline{\mathbb{D}}_{\leqslant 0} = \{0\}$, we focus on the case where $j \geq 1$. Let $\xi_n, n \geq 1$, be a sequence such that $\xi_n \in \underline{\mathbb{D}}_{\leqslant j}$, for all $n \geq 1$, and $\lim_{n\to\infty} d_{M'_1}(\xi_n, \xi) = 0$ for some $\xi \in \mathbb{D}$. Our goal is to prove that $\xi \in \underline{\mathbb{D}}_{\leqslant j}$. Note that by Lemma 4.7.3 above, for every $t \in \mathrm{Disc}(\xi)^c \cup \{1\}$,

$$\lim_{n\to\infty} \xi_n(t) = \xi(t). \tag{4.7.13}$$

We first show that ξ has at most j discontinuity points. Assume that $|\mathrm{Disc}(\xi)| \geq j+1$. Then there exists $0 \leq t_{1,-} < t_{1,+} < \cdots < t_{j+1,-} < t_{j+1,+} \leq 1$ such that $t_{i,-}$, $t_{i,+} \in \mathrm{Disc}(\xi)^c \cup \{1\}$, and $|\xi(t_{i,-}) - \xi(t_{i,+})| > 0$, for all $i \in \{1, \dots, j+1\}$. By (4.7.13), there exists N' such that $|\xi_{N'}(t_{i,-}) - \xi_{N'}(t_{i,+})| > 0$ for all $i \in \{1, \dots, j+1\}$. This leads to the contradiction that $|\mathrm{Disc}(\xi_{N'})| \leq j$. Now let $\underline{t} < \bar{t}$ be two neighbouring discontinuity points of ξ. We claim that ξ is constant on $(\underline{t}, \bar{t})$. To see this, assume that the opposite statement holds. Then there exists $t_1 < t_{j+2}$ such that $\underline{t} < t_1 < t_{j+2} < \bar{t}$ and $\xi(t_1) \neq \xi(t_{j+2})$. W.l.o.g. we assume that $\xi(t_1) < \xi(t_{j+2})$. Since ξ is continuous on $(\underline{t}, \bar{t})$, there exists $t_1 < t_2 < \cdots < t_{j+2}$ such that

$$\xi(t_1) < \xi(t_2) < \cdots < \xi(t_{j+2}) \qquad \text{with} \quad \epsilon' = \min_{i\in\{1,\dots,j+1\}} \xi(t_{i+1}) - \xi(t_i). \tag{4.7.14}$$

On the other hand, for any $\epsilon > 0$, by (4.7.13) there exists $N = N(\epsilon)$ such that

$$\xi_N(t_i) \in (\xi(t_i) - \epsilon, \xi(t_i) + \epsilon), \qquad \text{for all } i \in \{1, \dots, j+2\}. \tag{4.7.15}$$

In view of (4.7.14) and (4.7.15), by choosing $\epsilon < \epsilon'$ we conclude that ξ_N has at least $j+1$ discontinuity points, which leads to the contradiction that $|\mathrm{Disc}(\xi_N)| \leq j$. Thus we conclude that ξ is constant between any two neighbouring discontinuity points. Similarly one can show that $\xi(t^+) - \xi(t^-) > 0$ for every $t \in \mathrm{Disc}(\xi)$. □

Proof of Corollary 4.7.1. Note that $\underline{\mathbb{D}}^{\mu}_{\leqslant j} = \mathbb{D}^{\mu}_{\leqslant j} \cup \{\xi \in \mathbb{D} \colon \xi(0) > 0,\ \xi - \xi(0) \in \mathbb{D}^{\mu}_{\leqslant j-1}\}$. In particular, $\mathbb{D}^{\mu}_{\leqslant j} \subseteq \underline{\mathbb{D}}^{\mu}_{\leqslant j}$. Using Lemma 4.7.2, Corollary 4.7.1 is a consequence of Lemma 4.7.1 and Theorem 4.1 in [105]. □

The following lemma is essential in the proving Proposition 4.7.1. Recall $\bar{X}'_n$ was defined in (4.7.1). Define

$$R_n = \{R_n(t), t \in [0,1]\}, \quad \text{where } R_n(t) = \frac{1}{n}\sum_{i=r_{N(n)}}^{\lfloor nt\rfloor -1} X_i. \tag{4.7.16}$$

Lemma 4.7.5. *Suppose that Assumptions 4.2.1 and 4.2.2 hold. Moreover, let $B_1 \geq 0$ and C_+ as in Theorem 4.3.1 be strictly positive. The following holds for any $\delta > 0$, $\gamma > 0$, and $j \geq 0$.*

1. *First we have that*

$$\mathbf{P}(\bar{X}'_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\leqslant j-1})^{-\gamma}, R_n(1) \geq \delta) = o((n\mathbf{P}(X'_1 \geq n))^{j+1}), \quad \text{as } n \to \infty.$$

2. *Moreover, we have that*

$$\mathbf{P}(R_n \in (\mathbb{D} \setminus \mathbb{D}_{\leqslant 1})^{-\gamma}) = o((n\mathbf{P}(X'_1 \geq n))^{j}), \quad \text{as } n \to \infty.$$

Proof of Proposition 4.7.1. To begin with, for $\epsilon > 0$, define

$$\mathfrak{E}^{\epsilon}_3(n) = \{N^-_\epsilon(n) < N(n) \leq N^+_\epsilon(n)\}, \tag{4.7.17}$$

where $N^-_\epsilon(n) = \lfloor n/\mathbf{E}r_1 - n\epsilon\rfloor$ and $N^+_\epsilon(n) = \lfloor n/\mathbf{E}r_1 + n\epsilon\rfloor$. Using Cramér's theorem, it is easy to see that $\mathbf{P}(\mathfrak{E}^{\epsilon}_3(n)^c)$ decays exponentially to 0 as $n \to \infty$. Defining $\Delta_i = r_i - r_{i-1}$, we have that

$$\{d_{M'_1}(\bar{X}_n, \bar{X}'_n) \geq 2\delta\} \subseteq \{\exists\, i \leq N(n) \ s.t.\ \Delta_i \geq n\delta\} \cup \{R_n(1) \geq \delta\}. \tag{4.7.18}$$

First we show that for any $j \geq 0$, $\delta > 0$, and $\gamma > 0$,

$$\lim_{n\to\infty} (n\mathbf{P}(X'_1 \geq n))^{-j}\mathbf{P}(\bar{X}'_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\leqslant j-1})^{-\gamma}, d_{M'_1}(\bar{X}_n, \bar{X}'_n) \geq 2\delta) = 0.$$

By (4.7.18) we have that

$$\begin{aligned}
&\mathbf{P}(\bar{X}'_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\leqslant j-1})^{-\gamma}, d_{M'_1}(\bar{X}_n, \bar{X}'_n) \geq 2\delta) \\
&\leq \mathbf{P}(\exists\, i \leq N(n) \ s.t.\ \Delta_i \geq n\delta) + \mathbf{P}(\bar{X}'_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\leqslant j-1})^{-\gamma},\ R_n(1) \geq \delta) \\
&= \mathbf{P}(\exists\, i \leq N(n) \ s.t.\ \Delta_i \geq n\delta) + o((n\mathbf{P}(X'_1 \geq n))^{j}), \qquad (4.7.19)
\end{aligned}$$

where in (4.7.19) we used Lemma 4.7.5 (1) above. It remains to analyze the first term in (4.7.19). Note that

$$\mathbf{P}(\exists\, i \leq N(n) \ s.t.\ \Delta_i \geq n\delta)$$

$$
\begin{aligned}
&\le \mathbf{P}(\exists\, i \le N(n)\ s.t.\ \Delta_i \ge n\delta, \mathfrak{E}_3^\epsilon(n)) + \mathbf{P}(\mathfrak{E}_3^\epsilon(n)^c) \\
&= \mathbf{P}(\exists\, i \le N(n)\ s.t.\ \Delta_i \ge n\delta, \mathfrak{E}_3^\epsilon(n)) + o((n\mathbf{P}(X_1' \ge n))^j) \\
&\le \mathbf{P}\left(\exists\, i \le \lfloor n/\mathbf{E}\tau_1 + n\epsilon \rfloor\ s.t.\ \Delta_i \ge n\delta\right) + o((n\mathbf{P}(X_1' \ge n))^j) \\
&\le \lfloor n/\mathbf{E}r_1 + n\epsilon \rfloor \mathbf{P}(r_1 \ge n\delta) + o((n\mathbf{P}(X_1' \ge n))^j) \\
&= o((n\mathbf{P}(X_1' \ge n))^j),
\end{aligned}
$$

for any $j \ge 0$. Next we show that

$$\lim_{n\to\infty} (n\mathbf{P}(X_1' \ge n))^{-j} \mathbf{P}(\bar{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\leqslant j-1})^{-\gamma}, d_{M_1'}(\bar{X}_n, \bar{X}_n') \ge 2\delta) = 0.$$

In view of the estimation right above, it is sufficient to show that

$$\lim_{n\to\infty} (n\mathbf{P}(X_1' \ge n))^{-j} \mathbf{P}(\bar{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\leqslant j-1})^{-\gamma},\ \bar{X}_n' \in (\underline{\mathbb{D}}^{\mu}_{\leqslant j-1})_\rho,\ R_n(1) \ge \delta) = 0,$$

for some $\rho > 0$. Note that

$$
\begin{aligned}
&\mathbf{P}(\bar{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\leqslant j-1})^{-\gamma},\ \bar{X}_n' \in (\underline{\mathbb{D}}^{\mu}_{\leqslant j-1})_\rho,\ R_n(1) \ge \delta) \\
&= \mathbf{P}(\bar{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\leqslant j-1})^{-\gamma},\ \bar{X}_n' \in (\underline{\mathbb{D}}^{\mu}_{\leqslant j-1})_\rho \cap (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\leqslant j-2})^{-\rho},\ R_n(1) \ge \delta) \\
&\quad + \mathbf{P}(\bar{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\leqslant j-1})^{-\gamma},\ \bar{X}_n' \in (\underline{\mathbb{D}}^{\mu}_{\leqslant j-1})_\rho \cap (\underline{\mathbb{D}}^{\mu}_{\leqslant j-2})_\rho,\ R_n(1) \ge \delta) \\
&\le \mathbf{P}(\bar{X}_n' \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\leqslant j-2})^{-\rho},\ R_n(1) \ge \delta) \\
&\quad + \mathbf{P}(\bar{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\leqslant j-1})^{-\gamma},\ \bar{X}_n' \in (\underline{\mathbb{D}}^{\mu}_{\leqslant j-2})_\rho) \\
&= \mathbf{P}(\bar{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\leqslant j-1})^{-\gamma},\ \bar{X}_n' \in (\underline{\mathbb{D}}^{\mu}_{\leqslant j-2})_\rho) + o(n^{-j(\alpha-1)}).
\end{aligned}
$$

Thus, it remains to consider the first term in the last equation. Combining Lemma 4.7.5 (2) above with the fact that

$$\mathbf{P}(\bar{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\leqslant j-1})^{-\gamma},\ \bar{X}_n' \in (\underline{\mathbb{D}}^{\mu}_{\leqslant j-2})_\rho) \le \mathbf{P}(R_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}_{\leqslant 1})^{-\rho}) + o(n^{-j(\alpha-1)}),$$

for ρ small enough, we conclude the proof. □

Proof of Lemma 4.7.5. Part 1): We start showing the first equivalence. Defining $\bar{X}'_{\leqslant k,n} = \{\bar{X}'_{\leqslant k,n}(t), t \in [0,1]\}$ by $\bar{X}'_{\leqslant k,n}(t) = 1/n \sum_{i=1}^{N(nt)\wedge k} X_i'$, we have that

$$
\begin{aligned}
&\mathbf{P}(\bar{X}_n' \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\leqslant j-1})^{-\gamma}, R_n(1) \ge \delta) \\
&\le \mathbf{P}\left(\bar{X}_n' \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\leqslant j-1})^{-\gamma},\ \sum_{i=r_{N(n)}}^{r_{N(n)+1}-1} X_i \ge n\delta,\ \mathfrak{E}_3^\epsilon(n)\right) + \mathbf{P}(\mathfrak{E}_3^\epsilon(n)^c)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{k=N_\epsilon^-(n)}^{N_\epsilon^+(n)} \mathbf{P}(\bar{X}'_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\leqslant j-1})^{-\gamma},\, X'_{N(n)+1} \geq n\delta,\, N(n) = k) \\
&\quad + o((n\mathbf{P}(X'_1 \geq n))^{j+1}) \\
&= \sum_{k=N_\epsilon^-(n)}^{N_\epsilon^+(n)} \mathbf{P}(\bar{X}'_{\leqslant k,n} \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\leqslant j-1})^{-\gamma},\, X'_{k+1} \geq n\delta,\, N(n) = k) \\
&\quad + o((n\mathbf{P}(X'_1 \geq n))^{j+1}) \\
&\leq \sum_{k=N_\epsilon^-(n)}^{N_\epsilon^+(n)} \mathbf{P}(\bar{X}'_{\leqslant k,n} \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\leqslant j-1})^{-\gamma},\, X'_{k+1} \geq n\delta) + o((n\mathbf{P}(X'_1 \geq n))^{j+1}) \\
&= \sum_{k=N_\epsilon^-(n)}^{N_\epsilon^+(n)} \mathbf{P}(\bar{X}'_{\leqslant k,n} \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\leqslant j-1})^{-\gamma})\mathbf{P}\left(X'_{k+1} \geq n\delta\right) + o((n\mathbf{P}(X'_1 \geq n))^{j+1}) \\
&\leq \mathbf{P}(X'_1 \geq n\delta) \sum_{k=N_\epsilon^-(n)}^{N_\epsilon^+(n)} \mathbf{P}(\bar{X}'_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\leqslant j-1})^{-\gamma/2}) + o((n\mathbf{P}(X'_1 \geq n))^{j+1}) \\
&\leq 2\epsilon n\mathbf{P}(X'_1 \geq n\delta)\mathbf{P}(\bar{X}'_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\leqslant j-1})^{-\gamma/2}) + o((n\mathbf{P}(X'_1 \geq n))^{j+1}). \quad (4.7.20)
\end{aligned}
$$

It remains to consider the first term in (4.7.20). Using Corollary 4.7.1, we have that

$$
\varlimsup_{n\to\infty} (n\mathbf{P}(X'_1 \geq n))^{-(j+1)}\, 2\epsilon n\mathbf{P}(X'_1 \geq n\delta)\mathbf{P}(\bar{X}'_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\leqslant j-1})^{-\gamma/2}) \leq c\epsilon, \tag{4.7.21}
$$

for some $c > 0$ independent of ϵ. Part (1) is proved using (4.7.20) and (4.7.21), together with letting $\epsilon \to 0$.

Part 2): Note that

$$
\begin{aligned}
&\mathbf{P}\left(R_n \in (\mathbb{D} \setminus \mathbb{D}_{\leqslant 1})^{-\gamma}\right) \\
&= \mathbf{P}\left(R_n \in (\mathbb{D} \setminus \mathbb{D}_{\leqslant 1})^{-\gamma},\, \frac{r_{N(n)}+1}{n} > \rho\right) \\
&\quad + \mathbf{P}\left(R_n \in (\mathbb{D} \setminus \mathbb{D}_{\leqslant 1})^{-\gamma},\, \frac{r_{N(n)}+1}{n} \leq \rho\right),
\end{aligned}
$$

where the first term equals zero for sufficiently large $\rho \in (0,1)$. Hence, it is sufficient to consider the second term which is bounded by

$$\begin{aligned}
\mathbf{P}\left(\frac{r_{N(n)}+1}{n} \le \rho\right) &\le \mathbf{P}\left(r_{N(n)} \le n\rho\right) \le \mathbf{P}\left(r_{N(n)} \le n\rho,\ \mathfrak{E}_3^\epsilon(n)\right) + \mathbf{P}\left(\mathfrak{E}_3^\epsilon(n)^c\right) \\
&= \mathbf{P}\left(\sum_{i=1}^{N(n)} \Delta_i \le n\rho,\ \mathfrak{E}_3^\epsilon(n)\right) + o((n\mathbf{P}(X_1' \ge n))^j) \\
&\le \mathbf{P}\left(\sum_{i=1}^{N_\epsilon^-(n)} \Delta_i \le n\rho\right) + o((n\mathbf{P}(X_1' \ge n))^j) \\
&\le \mathbf{P}\left(\sum_{i=1}^{N_\epsilon^-(n)} \frac{\Delta_i}{N_\epsilon^-(n)} \le \frac{\rho}{1/\mathbf{E}r_1 - \epsilon}\right) + o((n\mathbf{P}(X_1' \ge n))^j).
\end{aligned} \tag{4.7.22}$$

Note that, for every $\rho \in (0,1)$ there exists a sufficiently small $\epsilon > 0$ such that $\rho/(1/\mathbf{E}r_1 - \epsilon) < \mathbf{E}r_1$. For this choice of ϵ, the first term in (4.7.22) decays exponentially thanks to Cramér's theorem. □

4.7.2 Proofs of Theorem 4.3.3

We consider the case where B_1 is a general random variable taking values in $\mathbb{R}$. The idea behind the proof of Theorem 4.3.3 is similar to the one in the one-sided case.

- In Corollary 4.7.2 below we establish a sample-path large deviations result for the aggregated process $\bar{X}_n'$ (see (4.7.1) above).
- In Proposition 4.7.2 we show the asymptotic equivalence between the aggregated process $\bar{X}_n'$ and the original process $\bar{X}_n$. In Lemma 4.7.7 we deal with the technical issues appearing in Proposition 4.7.2.
- Part 1) of Theorem 4.3.3 follows by combining Corollary 4.7.2 with Proposition 4.7.2. Part 2) is a direct consequence of part 1).

Lemma 4.7.6. *For all $j \ge 0$ and all $z \in \mathbb{R}$, the set $\mathbb{D}^z_{\ll j}$ is closed w.r.t. $(\mathbb{D}, d_{M_1'})$.*

Recall $C^z_{j,k}$ was defined in (4.3.9). Let C_+, C_- be as in Theorem 4.3.1.

Corollary 4.7.2. *Suppose that Assumptions 4.2.1 and 4.2.2 hold. If* $C_+C_- > 0$, *then for any* $j \geq 1$

$$n^{j(\alpha-1)}\mathbf{P}(\bar{X}'_n \in \cdot\,) \to (\mathbf{E}r_1)^j \sum\nolimits_{(l,m)\in I_{=j}} (C_+)^l (C_-)^m C^{\mu}_{l,m}(\,\cdot\,),$$

in $\mathbb{M}(\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})$ *as* $n \to \infty$, *where* $I_{=j} = \{(l,m) \in \mathbb{Z}^2_+ : l+m = j\}$.

Proposition 4.7.2. *Suppose that Assumptions 4.2.1 and 4.2.2 hold. If* $C_+C_- > 0$, *then the following hold for all* $j \geq 0$:

1. *First*

$$\lim_{n\to\infty} n^{j(\alpha-1)}\mathbf{P}(\bar{X}'_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, d_{M'_1}(\bar{X}_n, \bar{X}'_n) > \delta) = 0.$$

2. *Assume additionally that* $\mathbf{E}|B_1|^m < \infty$ *for every* $m \in \mathbb{Z}_+$. *Then*

$$\lim_{n\to\infty} n^{j(\alpha-1)}\mathbf{P}(\bar{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, d_{M'_1}(\bar{X}_n, \bar{X}'_n) > \delta) = 0.$$

In particular, $\bar{X}_n$ *is asymptotically equivalent to* $\bar{X}'_n$ *w.r.t.* $n^{-j(\alpha-1)}$ *and* $\mathbb{D}^{\mu}_{\ll j}$.

We need the following lemma to prove Proposition 4.7.2. Set

$$R_{p,n}(t) = \frac{1}{n}\sum_{i=r_p}^{\lfloor r_{p+1}t\rfloor-1} X_i.$$

Let $T_1(u) = T(u) = \inf\{n \geq 0 \colon |X_n| > u\}$ and

$$T_{i+1}(u) = \inf\{n \geq T_i(u) \colon -\operatorname{sign}(X_{T_i}(u))X_n > u\}, \quad i \geq 1.$$

Define $\bar{X}_{i,n} = \{\bar{X}_{i,n}(t), t \in [0,1]\}$ and $\bar{X}'_{i,n} = \{\bar{X}'_{i,n}(t), t \in [0,1]\}$ by

$$\bar{X}_{i,n}(t) = \frac{1}{n}\sum_{l=r_{i-1}}^{\lfloor nt\rfloor \wedge r_i - 1} X_l, \quad \text{and} \quad \bar{X}'_{i,n}(t) = \frac{X'_i}{n}\mathbb{1}_{[r_i/n,1]}(t). \tag{4.7.23}$$

respectively.

Lemma 4.7.7. *Suppose that Assumptions 4.2.1 and 4.2.2 hold. Moreover, assume that* $\mathbf{E}|B_1|^m < \infty$ *for every* $m \in \mathbb{Z}_+$. *Let* C_+, C_- *be as in Theorem 4.3.1 such that* $C_+C_- > 0$.

1. *For any $i \geq 1$, $j \geq 2$, $\epsilon > 0$, and $\delta > 0$, there exists c_1, c_2 and n_1, n_2 (independent of i) respectively such that*

$$\begin{aligned} &\mathbf{P}(d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta) \leq c_1 n^{-(2-\epsilon)\alpha}, \qquad \textit{for all } n \geq n_1, \\ \textit{and} \quad &\mathbf{P}(\bar{X}_{i,n} \in (\mathbb{D} \setminus \mathbb{D}_{\ll j})^{-\delta}) \leq c_2 n^{-(j-\epsilon)\alpha}, \qquad \textit{for all } n \geq n_2. \end{aligned}$$

2. *For any $j \geq 1$, $\hat{X}_n$ is asymptotically equivalent to $\bar{X}'_n$ w.r.t. $n^{-j(\alpha-1)}$ and $\mathbb{D}^{\mu}_{\ll j}$.*

3. *For any $i \in \{N^-_\epsilon(n), \dots, N^+_\epsilon(n)\}$, $j \geq 1$, $\delta > 0$, and $\epsilon > 0$, there exists c and n_0 (independent of i) such that*

$$\mathbf{P}(R_{i,n} \in (\mathbb{D} \setminus \mathbb{D}_{\ll j})^{-\delta}) \leq c n^{-(j-\epsilon)\alpha}, \qquad \textit{for all } n \geq n_0.$$

Remark 4.5. Without the additional assumption $\mathbf{E}|B_1|^m < \infty$ for every $m \in \mathbb{Z}_+$, one can still show that $\mathbf{P}(T_2(n^\beta) < r_1) = o(n^{-\alpha})$, by following the arguments as in the proof of Lemma 4.7.7. Hence, under Assumptions 4.2.1 and 4.2.2, uniformly in i,

$$\lim_{n\to\infty} n^\alpha \mathbf{P}(d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta) = 0.$$

Proof of Proposition 4.7.2. To begin with, recall that, for $\epsilon > 0$

$$\mathfrak{E}^\epsilon_3(n) = \{N^-_\epsilon(n) \leq N(n) \leq N^+_\epsilon(n)\},$$

where $N^-_\epsilon(n) = n\lfloor 1/\mathbf{E}r_1 - \epsilon\rfloor$ and $N^+_\epsilon(n) = n\lfloor 1/\mathbf{E}r_1 + \epsilon\rfloor$. Moreover, $\mathbf{P}((\mathfrak{E}^\epsilon_3(n))^c)$ decays exponentially to 0 as $n \to \infty$. Let R_n be as in (4.7.16). Recalling $\Delta_i = r_i - r_{i-1}$, we have that

$$\{d_{M_1'}(\bar{X}_n, \bar{X}'_n) > \delta\} \subseteq \big\{\exists\, i \leq N(n) \text{ s.t. } d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta\big\} \cup \{\|R_n\|_\infty \geq \delta\}\,. \tag{4.7.24}$$

To see (4.7.24), we assume that the opposite statement holds. Given that the event $\{d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) < \delta\}$ takes place, there exist $(u_1^i, v_1^i) \in \Gamma'(\bar{X}_{i,n})$ and $(u_2^i, v_2^i) \in \Gamma'(\bar{X}'_{i,n})$ such that $\|u_1^i - u_2^i\|_\infty \vee \|v_1^i - v_2^i\|_\infty < \delta + \eta$. W.l.o.g. we assume that

$$\{s\colon v_1^i(s) = r_{i-1}/n,\, u_1^i(s) = 0\} \cap \{s\colon v_2^i(s) = r_{i-1}/n,\, u_2^i(s) = 0\} \neq \emptyset, \tag{4.7.25}$$

as well as

$$\{s\colon v_1^i(s) = r_i/n,\, u_1^i(s) = X'_i/n\} \cap \{s\colon v_2^i(s) = r_i/n,\, u_2^i(s) = X'_i/n\} \neq \emptyset.$$

We give here the reasoning for (4.7.25), where the second equation can be obtained by following same arguments. Let $s_1 \in \{s\colon v_1^i(s) = r_{i-1}/n,\, u_1^i(s) = 0\}$ and $s_2 \in \{s\colon v_2^i(s) = r_{i-1}/n,\, u_2^i(s) = 0\}$. When $s_1 = s_2$, we are done. We assume $s_1 < s_2$, otherwise one can change the role of s_1 and s_2. Define a new parametric representation $(\bar{u}_2^i, \bar{v}_2^i) \in \Gamma'(\bar{X}'_{i,n})$ by

$$\bar{v}_2^i(s) = \begin{cases} v_1(s), & \text{for } s \in [0, s_1], \\ v_1(s_1), & \text{for } s \in (s_1, s_2), \\ v_2(s), & \text{for } s \in [s_2, 1], \end{cases} \qquad \bar{u}_2^i(s) = \begin{cases} 0, & \text{for } s \in [0, s_1], \\ 0, & \text{for } s \in (s_1, s_2), \\ u_2(s), & \text{for } s \in [s_2, 1]. \end{cases}$$

It is easy to check that indeed $(\bar{u}_2^i, \bar{v}_2^i)$ is a parametric representation of $\Gamma'(\bar{X}'_{i,n})$. Moreover, $\|u_1^i - \bar{u}_2^i\|_\infty = \|u_1^i - u_2^i\|_\infty < \delta + \eta$,

$$|v_1^i(s) - \bar{v}_2^i(s)| = |v_1^i(s) - v_1^i(s_1)| \le v_1^i(s_2) - v_1^i(s_1) = v_1^i(s_2) - v_2^i(s_2) < \delta + \eta,$$

for $s \in (s_1, s_2)$, and hence, $\|v_1^i - \bar{v}_2^i\|_\infty < \delta + \eta$. In view of the construction above, we can replace v_2^i by $\bar{v}_2^i$, so that (4.7.25) holds. For the similar reasoning, on the event $\{\|R_n\|_\infty < \delta\} \subseteq \{d_{M_1'}(R_n, 0) < \delta\}$, there exist $(u_1^{N(n)+1}, v_1^{N(n)+1}) \in \Gamma'(R_n)$ and $(u_2^{N(n)+1}, v_2^{N(n)+1}) \in \Gamma'(0)$ such that

$$\|u_1^{N(n)+1} - u_2^{N(n)+1}\|_\infty \vee \|v_1^{N(n)+1} - v_2^{N(n)+1}\|_\infty < \delta + \eta,$$

and the intersection of

$$\{s\colon v_1^{N(n)+1}(s) = r_{N(n)}/n,\, u_1^{N(n)+1}(s) = 0\}$$

and

$$\{s\colon v_2^{N(n)+1}(s) = r_{N(n)}/n,\, u_2^{N(n)+1}(s) = 0\}$$

is an empty set. Now, we pick $s_-^1 = 0$, $s_+^{N(n)+1} = 1$,

$$s_+^i \in \{s\colon v_1^i(s) = r_i/n,\, u_1^i(s) = X_i'/n\} \cap \{s\colon v_2^i(s) = r_i/n,\, u_2^i(s) = X_i'/n\},$$

for $i \in \{1, \dots, N(n)\}$, and

$$s_-^i \in \{s\colon v_1^i(s) = r_i/n,\, u_1^i(s) = 0\} \cap \{s\colon v_2^i(s) = r_i/n,\, u_2^i(s) = 0\},$$

for $i \in \{2, \dots, N(n) + 1\}$. W.l.o.g. we assume that $s_+^i = s_-^{i+1}$, otherwise one can apply a strictly increasing, continuous bijection from $[0, 1]$ to itself to the corresponding parametric representation, which preserves the uniform distance between parametric representations. Finally, we define parametric representations $(u_1, v_1) \in \Gamma'(\bar{X}_n)$ and $(u_2, v_2) \in \Gamma'(\bar{X}'_n)$ by $v_i(s) = v_i^j(s)$, and $u_i(s) = u_i^j(s) + \sum_{k=1}^{j-1} X'_k$, for $s \in [s_-^j, s_+^j]$, $j \in \{1, \ldots, N(n)+1\}$, and $i \in \{1, 2\}$. It is easy to check that $\|u_1 - u_2\|_\infty \vee \|v_1 - v_2\|_\infty < \delta + \eta$, and hence, $d(\bar{X}_n, \bar{X}'_n) \leq \|u_1 - u_2\|_\infty \vee \|v_1 - v_2\|_\infty < \delta + \eta$. Letting $\eta \to 0$ leads to the contradiction of $d_{M_1'}(\bar{X}_n, \bar{X}'_n) > \delta$.

Part 1): For $\gamma > 0$ and $j \geq 1$, define

$$\mathcal{D}^{\gamma}_{\geqslant j} = \{\xi \in \mathbb{D} \colon |\mathrm{Disc}_\gamma(\xi)| \geq j\}, \quad \mathrm{Disc}_\gamma(\xi) = \{t \in \mathrm{Disc}(\xi) \colon |\xi(t) - \xi(t^-)| \geq \gamma\}. \tag{4.7.26}$$

Note that (cf. the proof of Lemma 2 in [29]), for any $L > 0$, there exists a $\bar{\gamma} = \bar{\gamma}(\gamma, L) > 0$ sufficiently small such that

$$\mathbf{P}(\bar{X}'_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\ll j})^{-\gamma} \cap (\mathcal{D}^{\bar{\gamma}}_{\geqslant j})^c) = o(n^{-L}). \tag{4.7.27}$$

Thus, it suffices to show that for any $j \geq 1$ and any $\delta > 0$

$$\lim_{n\to\infty} n^{j(\alpha-1)} \mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\bar{\gamma}}_{\geqslant j},\, d_{M_1'}(\bar{X}_n, \bar{X}'_n) \geq 2\delta) = 0.$$

By (4.7.24) we have that

$$\begin{aligned}
&\mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\bar{\gamma}}_{\geqslant j},\, d_{M_1'}(\bar{X}_n, \bar{X}'_n) \geq 2\delta) \\
&\leq \mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\bar{\gamma}}_{\geqslant j},\, \exists\, i \leq N(n) \text{ } s.t. \text{ } d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta) \\
&\quad + \mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\bar{\gamma}}_{\geqslant j},\, \|R_n\|_\infty \geq \delta) \\
&= \mathbf{(IV.1)} + \mathbf{(IV.2)},
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{(IV.1)} = &\mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\bar{\gamma}}_{\geqslant j},\, \mathfrak{E}_3^{\epsilon}(n),\, \exists\, i \leq N(n) \text{ } s.t. \text{ } d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta) \\
&+ o(n^{-j(\alpha-1)}).
\end{aligned}$$

For $p \in \mathbb{Z}_+$, let $\mathcal{P}(E, p)$ denote the set of all p-permutations of a discrete set E. Using Lemma 4.7.7 (1) and the fact that the blocks $\{X_{r_{i-1}}, \ldots, X_{r_i}\}$, $i \geq 1$ are mutually independent, we obtain that

$$\begin{aligned}
&\mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\bar{\gamma}}_{\geqslant j},\, \mathfrak{E}_3^{\epsilon}(n),\, \exists\, i \leq N(n) \text{ s.t.} d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta) \\
&\leq \mathbf{P}(\exists (i_1, \ldots, i_j) \in \mathcal{P}(\{1, \ldots, N_\epsilon^+(n)\}, j) \text{ s.t.} \\
&\qquad d_{M_1'}(\bar{X}_{i_1,n}, \bar{X}'_{i_1,n}) \geq \delta,\, |X'_{i_p}| \geq n\bar{\gamma}, \forall 2 \leq p \leq j)
\end{aligned}$$

$$
\begin{aligned}
&= \mathcal{O}(n^j)\mathbf{P}(d_{M_1'}(\bar{X}_{i_1,n}, \bar{X}'_{i_1,n}) \geq \delta)\mathbf{P}(|X'_{i_p}| \geq n\bar{\gamma})^{j-1} \\
&= \mathcal{O}(n^j)o(n^{-\alpha})\mathcal{O}(n^{-(j-1)\alpha}) = o(n^{-j(\alpha-1)}),
\end{aligned}
$$

where $\mathbf{P}(d_{M_1'}(\bar{X}_{i_1,n}, \bar{X}'_{i_1,n}) \geq \delta)$ is of order $o(n^{-\alpha})$ thanks to Remark 4.5. Recalling $\bar{X}'_{\leqslant m,n} = \{1/n \sum_{i=1}^{N(nt)\wedge m} X'_i, t \in [0,1]\}$, we have that

$$
\begin{aligned}
\textbf{(IV.2)} &\leq \mathbf{P}\left(\bar{X}'_n \in \mathcal{D}^{\bar{\gamma}}_{\geqslant j}, \sum_{i=r_{N(n)}}^{r_{N(n)+1}-1} |X_i| \geq n\delta,\, \mathfrak{E}^{\epsilon}_3(n)\right) + \mathbf{P}(\mathfrak{E}^{\epsilon}_3(n)^c) \\
&= \sum_{m=N^-_\epsilon(n)}^{N^+_\epsilon(n)} \mathbf{P}\left(\bar{X}'_n \in \mathcal{D}^{\bar{\gamma}}_{\geqslant j}, \sum_{i=r_{N(n)}}^{r_{N(n)+1}-1} |X_i| \geq n\delta,\, N(n) = m\right) \\
&\quad + o(n^{-j(\alpha-1)}) \\
&\leq \sum_{m=N^-_\epsilon(n)}^{N^+_\epsilon(n)} \mathbf{P}\left(\bar{X}'_{\leqslant m,n} \in \mathcal{D}^{\bar{\gamma}}_{\geqslant j}, \sum_{i=r_m}^{r_{m+1}-1} |X_i| \geq n\delta\right) + o(n^{-j(\alpha-1)}) \\
&= \mathbf{P}\left(\sum_{i=0}^{r_1-1} |X_i| \geq n\delta\right) \sum_{m=N^-_\epsilon(n)}^{N^+_\epsilon(n)} \mathbf{P}(\bar{X}'_{\leqslant m,n} \in \mathcal{D}^{\bar{\gamma}}_{\geqslant j}) + o(n^{-j(\alpha-1)}) \\
&\leq \mathbf{P}\left(\sum_{i=0}^{r_1-1} |X_i| \geq n\delta\right) \sum_{m=N^-_\epsilon(n)}^{N^+_\epsilon(n)} \mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\bar{\gamma}}_{\geqslant j}) + o(n^{-j(\alpha-1)}) \\
&\leq \mathbf{P}\left(\sum_{i=0}^{r_1-1} |X_i| \geq n\delta\right) 2\epsilon n \mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\bar{\gamma}}_{\geqslant j}) + o(n^{-j(\alpha-1)}) \\
&= 2\epsilon n \mathcal{O}(n^{-\alpha})\mathcal{O}(n^{-j(\alpha-1)}) = o(n^{-j(\alpha-1)}),
\end{aligned}
$$

where $\mathbf{P}(\sum_{i=0}^{r_1-1} |X_i| \geq n\delta)$ is of order $\mathcal{O}(n^{-\alpha})$ due to Remark 4.4.

Part 2): In view of part (1), it is sufficient to show that

$$
\mathbf{P}(\bar{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\ll j})^{-\gamma},\, \bar{X}'_n \in (\underline{\mathbb{D}}^{\mu}_{\ll j})_{\rho/3}) = o(n^{-j(\alpha-1)}),
$$

for some $\rho > 0$. Noting $\hat{X}_n(t) = (1/n)\sum_{i=0}^{(\lfloor nt \rfloor \wedge r_{N(n)})-1} X_i$ for $t \in [0,1]$, we have that

$$
\{\bar{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\ll j})^{-\gamma},\, \bar{X}'_n \in (\underline{\mathbb{D}}^{\mu}_{\ll j})_{\rho/3}\}
$$

$$
\begin{aligned}
&\subseteq \{\bar{X}'_n \in (\mathbb{D}^{\mu}_{\ll j})_{\rho/3},\, \hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\rho}\} \cup \{\bar{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \hat{X}_n \in (\mathbb{D}^{\mu}_{\ll j})_{\rho}\} \\
&\subseteq \{\bar{X}'_n \in (\mathbb{D}^{\mu}_{\ll j})_{\rho/3},\, \hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\rho}\} \\
&\quad \cup \{\bar{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \hat{X}_n \in (\mathbb{D}^{\mu}_{\ll j-1})_{\rho}\} \\
&\quad \cup \{\bar{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \hat{X}_n \in (\mathbb{D}^{\mu}_{\ll j})_{\rho} \cap (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j-1})^{-\rho}\}.
\end{aligned}
$$

Iterating this procedure $j + k$ times, we obtain that

$$
\begin{aligned}
&\{\bar{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \bar{X}'_n \in (\mathbb{D}^{\mu}_{\ll j})_{\rho/3}\} \\
&\subseteq \{\bar{X}'_n \in (\mathbb{D}^{\mu}_{\ll j})_{\rho/3},\, \hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\rho}\} \cup \{\bar{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \hat{X}_n \in (\mathbb{D}^{\mu}_{0})_{\rho}\} \\
&\quad \cup \bigcup_{i=1}^{j+k-1} \{\bar{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \hat{X}_n \in (\mathbb{D}^{\mu}_{\ll j+1-i})_{\rho} \cap (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j-i})^{-\rho}\}. \qquad (4.7.28)
\end{aligned}
$$

Now, note that

$$
\begin{aligned}
&\{\bar{X}'_n \in (\mathbb{D}^{\mu}_{\ll j})_{\rho/3},\, \hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\rho}\} \\
&\subseteq \{\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\rho}, d_{M_1'}(\bar{X}'_n, \hat{X}_n) \geq \rho/3\}. \qquad (4.7.29)
\end{aligned}
$$

Moreover, for $\rho > 0$ sufficiently small, we have that

$$
\{\bar{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \hat{X}_n \in (\mathbb{D}^{\mu}_{0})_{\rho}\} \subseteq \{R_n \in (\mathbb{D} \setminus \mathbb{D}_{\ll j})^{-\rho}\}, \qquad (4.7.30)
$$

and that

$$
\begin{aligned}
&\{\bar{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \hat{X}_n \in (\mathbb{D}^{\mu}_{\ll j+1-i})_{\rho} \cap (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j-i})^{-\rho}\} \\
&\subseteq \{\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j-i})^{-\rho},\, R_n \in (\mathbb{D} \setminus \mathbb{D}_{\ll i})^{-\rho}\}, \qquad (4.7.31)
\end{aligned}
$$

for all $i \in \{1, \ldots, j + k - 1\}$. In view of (4.7.28)–(4.7.31), we have that

$$
\begin{aligned}
&\mathbf{P}(\bar{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \bar{X}'_n \in (\mathbb{D}^{\mu}_{\ll j})_{\rho/3}) \\
&\leq \mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\rho}, d_{M_1'}(\bar{X}'_n, \hat{X}_n) \geq \rho/3) + \mathbf{P}(R_n \in (\mathbb{D} \setminus \mathbb{D}_{\ll j})^{-\rho}) \\
&\quad + \sum_{i=1}^{j+k-1} \mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j-i})^{-\rho},\, R_n \in (\mathbb{D} \setminus \mathbb{D}_{\ll i})^{-\rho}), \qquad (4.7.32)
\end{aligned}
$$

where the first term in the previous inequality is of order $o(n^{-j(\alpha-1)})$ due to Lemma 4.7.7 (2) above. Turning to estimating the summation in (4.7.32), we define $R_{p,n} = \{R_{p,n}(t), t \in [0,1]\}$ by

$$R_{p,n}(t) = \frac{1}{n} \sum_{i=r_p}^{\lfloor r_{p+1}t \rfloor - 1} X_i.$$

Using the facts that $R_{N(n),n}(t) = R_n(r_{N(n)+1}t/n)$ and $r_{N(n)+1}/n > 1$ a.s., we have that

$$R_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}_{\ll i})^{-\rho} \;\Rightarrow\; R_{N(n),n} \in (\mathbb{D} \setminus \underline{\mathbb{D}}_{\ll i})^{-\rho/2}. \tag{4.7.33}$$

Define $\bar{X}_{\leqslant p,n} = \{\bar{X}_{\leqslant p,n}(t), t \in [0,1]\}$ by $\bar{X}_{\leqslant p,n}(t) = (1/n) \sum_{i=0}^{(\lfloor nt \rfloor \wedge r_{N(n) \wedge p}) - 1} X_i$. In view of (4.7.33), we have that

$$\begin{aligned}
&\mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\ll j-i})^{-\rho},\, R_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}_{\ll i})^{-\rho}) \\
&\le \mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\ll j-i})^{-\rho},\, R_{N(n),n} \in (\mathbb{D} \setminus \underline{\mathbb{D}}_{\ll i})^{-\rho/2}) \\
&\le \mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\ll j-i})^{-\rho},\, R_{N(n),n} \in (\mathbb{D} \setminus \underline{\mathbb{D}}_{\ll i})^{-\rho/2},\, \mathfrak{E}_3^{\epsilon}(n)) + \mathbf{P}(\mathfrak{E}_3^{\epsilon}(n)^c) \\
&= \sum_{p=N_\epsilon^-(n)}^{N_\epsilon^+(n)} \mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\ll j-i})^{-\rho},\, R_{N(n),n} \in (\mathbb{D} \setminus \underline{\mathbb{D}}_{\ll i})^{-\rho/2},\, N(n) = p) \\
&\quad + o(n^{-j(\alpha-1)}) \\
&= \sum_{p=N_\epsilon^-(n)}^{N_\epsilon^+(n)} \mathbf{P}(\bar{X}_{\leqslant p,n} \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\ll j-i})^{-\rho},\, R_{p,n} \in (\mathbb{D} \setminus \underline{\mathbb{D}}_{\ll i})^{-\rho/2},\, N(n) = p) \\
&\quad + o(n^{-j(\alpha-1)}) \\
&\le \sum_{p=N_\epsilon^-(n)}^{N_\epsilon^+(n)} \mathbf{P}(\bar{X}_{\leqslant p,n} \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\ll j-i})^{-\rho}) \mathbf{P}(R_{p,n} \in (\mathbb{D} \setminus \underline{\mathbb{D}}_{\ll i})^{-\rho/2}) + o(n^{-j(\alpha-1)}) \\
&\le \mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\ll j-i})^{-\rho/2}) \sum_{p=N_\epsilon^-(n)}^{N_\epsilon^+(n)} \mathbf{P}(R_{p,n} \in (\mathbb{D} \setminus \underline{\mathbb{D}}_{\ll i})^{-\rho/2}) + o(n^{-j(\alpha-1)}) \\
&= \mathcal{O}(n^{-(j-i)(\alpha-1)}) 2\epsilon \mathcal{O}(n^{-i(\alpha-1)}) + o(n^{-j(\alpha-1)}),
\end{aligned}$$

where in the final step we use Lemma 4.7.7 (2)–(3). Letting $\epsilon \to 0$, we prove that the summation in (4.7.32) is of order $o(n^{-j(\alpha-1)})$. Similarly, it can be shown that $\mathbf{P}(R_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}_{\ll j})^{-\rho})$, and hence, $\mathbf{P}(\bar{X}_n \in (\mathbb{D} \setminus \underline{\mathbb{D}}^{\mu}_{\ll j})^{-\gamma},\, \bar{X}'_n \in (\underline{\mathbb{D}}^{\mu}_{\ll j})_{\rho/3})$ are of order $o(n^{-j(\alpha-1)})$. $\square$

Proof of Lemma 4.7.7. Let $\mathbb{D}^s$ denote the set of all step functions in $\mathbb{D}$. Let $\mathbb{D}^{s,\uparrow}$ denote the set of all non-decreasing step functions in $\mathbb{D}$. Define the mapping $\Psi^{\uparrow}\colon \mathbb{D}^s \to \mathbb{D}^{s,\uparrow}$ by $\zeta = \Psi^{\uparrow}(\xi)$ and

$$\zeta(t) = \inf\{\zeta'(t) \in \mathbb{R}\colon \zeta' \in \mathbb{D}^{s,\uparrow},\, \zeta' \geq \xi\}, \qquad \text{for all } t \in [0,1]. \tag{4.7.34}$$

Basically, $\Psi^{\uparrow}(\xi)$ is the least possible nondecreasing step function such that $\Psi^{\uparrow}(\xi) \geq \xi$.

Part 1): First we show that $\mathbf{P}(d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta) \leq \mathbf{P}(T_2(n^{\beta}) < r_1) + o(n^{-(2-\epsilon)\alpha})$, for any $\beta \in (0,1)$. To begin with, setting $\beta^0 = (1-\beta)/2$ we have that

$$\begin{aligned}\mathbf{P}(d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta) &\leq \mathbf{P}(d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta,\, r_i - r_{i-1} \leq n^{\beta_0}) \\ &\quad + \mathbf{P}(r_i - r_{i-1} > n^{\beta_0}) \\ &= \mathbf{P}(d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta,\, r_i - r_{i-1} \leq n^{\beta_0}) \\ &\quad + o(n^{-(2-\epsilon)\alpha}).\end{aligned}$$

Hence, it is sufficient to show that

$$\mathbf{P}(d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta,\, r_i - r_{i-1} \leq n^{\beta_0}) \leq \mathbf{P}(T_2(n^{\beta}) < r_1). \tag{4.7.35}$$

Note that $d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta$ implies $\|\bar{X}_{i,n} - \bar{X}'_{i,n}\|_{\infty} \geq \delta$, and hence,

$$\delta \leq \sup_{k \leq r_i \wedge n} \left| \frac{1}{n} \sum_{j=r_{i-1}}^{k-1} X_j \right| \leq \sup_{k \leq r_i} \left| \frac{1}{n} \sum_{j=r_{i-1}}^{k-1} X_j \right|.$$

It is sufficient to show that

$$\left\{ \sup_{k \leq r_i} \left| \frac{1}{n} \sum_{j=r_{i-1}}^{k-1} X_j \right| \geq \delta,\, d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta,\, r_i - r_{i-1} \leq n^{\beta_0} \right\}$$

is a subset of $\{T_2(n^{\beta}) < r_1\}$. We distinguish between the cases

1. $\sup_{k \leq r_i} \frac{1}{n} \sum_{j=r_{i-1}}^{k-1} X_j \geq \delta$,
2. $\inf_{k \leq r_i} \frac{1}{n} \sum_{j=r_{i-1}}^{k-1} X_j \leq -\delta$,

and focus on 1), since 2) can be dealt with by replacing X_i by $-X_i$. Note that

$$\sup_{k \le r_i \wedge n} \sum_{j=r_{i-1}}^{k-1} X_j \ge \delta n, \;\; r_i - r_{i-1} \le n^{\beta_0}$$

implies the existence of $k_1 \in \{r_{i-1}, \dots, r_i - 1\}$ such that $X_{k_1} > n^{1-\beta_0} > n^{\beta}$. Now, suppose that $X_k \ge -n^{\beta}$ for all $k \in \{r_{i-1}, \dots, r_i - 1\}$. Then the following statements must hold.

(i) For n sufficiently large, we have

$$\sup_{t\in[0,1]} \Psi^{\uparrow}(\bar{X}_{i,n})(t) - \sup_{t\in[0,1]} \bar{X}'_{i,n}(t) \le n^{-1}(r_i - r_{i-1})n^{\beta} \le n^{\beta+\beta_0-1} \le \delta/3,$$

and hence,

$$\sup_{t\in[0,1]} \bar{X}'_{i,n}(t) \ge \sup_{t\in[0,1]} \Psi^{\uparrow}(\bar{X}_{i,n})(t) - \delta/3 \ge 2/3\delta > 0.$$

Moreover, both $\Psi^{\uparrow}(\bar{X}_{i,n}) \in \mathbb{D}^{s,\uparrow}$ and $\bar{X}'_{i,n} \in \mathbb{D}^{s,\uparrow}$ are nonnegative functions in $\mathbb{D}$. Combining these with $r_i - r_{i-1} \le n^{\beta_0}$, we have that, for sufficiently large n,

$$\begin{aligned} & d_{M_1'}(\Psi^{\uparrow}(\bar{X}_{i,n}), \bar{X}'_{i,n}) \\ & \le \left\{ \sup_{t\in[0,1]} \Psi^{\uparrow}(\bar{X}_{i,n})(t) - \sup_{t\in[0,1]} \bar{X}'_{i,n}(t) \right\} \vee (r_i - r_{i-1})/n \le \delta/3. \end{aligned}$$

(ii) For n sufficiently large,

$$\begin{aligned} d_{M_1'}(\Psi^{\uparrow}(\bar{X}_{i,n}), \bar{X}_{i,n}) &\le \|\Psi^{\uparrow}(\bar{X}_{i,n}) - \bar{X}_{i,n}\|_{\infty} \\ &\le n^{-1}(r_i - r_{i-1})n^{\beta} \le n^{\beta+\beta_0-1} \le \delta/3. \end{aligned}$$

In view of (i) and (ii), we have that

$$d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \le d_{M_1'}(\bar{X}_{i,n}, \Psi^{\uparrow}(\bar{X}_{i,n})) + d_{M_1'}(\Psi^{\uparrow}(\bar{X}_{i,n}), \bar{X}'_{i,n}) \le 2\delta/3,$$

which leads to the contradiction of $d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \ge \delta$. Hence, we prove (4.7.35).

Next we show that $\mathbf{P}(\bar{X}_{i,n} \in (\mathbb{D} \setminus \mathbb{D}_{\ll k})^{-\delta}) = \mathbf{P}(T_k(n^{\beta}) < r_1) + o(n^{-(k-\epsilon)\alpha})$ for any $\beta \in (0,1)$. First we claim that

$d(\xi, \mathbb{D}_{\ll k}) > \delta \;\Rightarrow$

$$\exists (t_0, \dots, t_k) \text{ s.t. } 0 \le t_0 < t_1 < \dots < t_k \le 1,\ |\xi(t_i) - \xi(t_{i-1})| > \delta,\ i = 1, \dots, k. \tag{4.7.36}$$

To see this, assume that the opposite holds. Set $s_0 = 0$ and

$$s_i = \sup\{t \in (s_{i-1}, 1] \colon |\xi(t) - \xi(s_{i-1})| \le \delta\},$$

for $i = 1, \dots, k$. Define $\zeta \in \mathbb{D}$ by $\zeta(t) = \xi(s_i)$ for $s_i \le t < s_{i+1}$. Due to the assumption, we have $\zeta \in \mathbb{D}_{\ll k}$, $d(\xi, \zeta) \le \delta$, and hence, $d(\xi, \mathbb{D}_{\ll k}) \le \delta$. This leads to the contradiction of $d(\xi, \mathbb{D}_{\ll k}) > \delta$. Thus, we proved (4.7.36). Using the fact that $\mathbf{P}(r_1 > n\delta/2)$ decays exponentially, we are able to restrict ourselves to the case where $r_1 \le n\delta/2$. Let $(t_0, \dots, t_k)$ be as in the r.h.s. of (4.7.36). Using the fact that, under the M_1' topology, jumps with the same sign "merge" into one jump in case they are "close", we conclude that $\operatorname{sign}(\xi(t_i))\operatorname{sign}(\xi(t_{i-1})) = -1$ for $i \in \{1, \dots, k\}$. Combining this with the fact that $\mathbf{P}(r_1 > n^{(1-\beta)}) = o(n^{-(k-\epsilon)\alpha})$ we obtain that

$$\begin{aligned}
&\mathbf{P}(\bar{X}_{i,n} \in (\mathbb{D} \setminus \mathbb{D}_{\ll k})^{-\delta}) \\
&= \mathbf{P}(\bar{X}_{i,n} \in (\mathbb{D} \setminus \mathbb{D}_{\ll k})^{-\delta},\ r_1 \le n^{(1-\beta)}) + \mathbf{P}(r_1 > n^{(1-\beta)}) \\
&= \mathbf{P}(T_k(n^\beta) < r_1) + o(n^{-(k-\epsilon)\alpha})
\end{aligned} \tag{4.7.37}$$

for any $\beta \in (0, 1)$.

Now, it remains to show that $\mathbf{P}(T_k(u^\beta) < r_1) = \mathcal{O}(u^{-(k-\epsilon)\alpha})$ as $u \to \infty$. We prove this by induction in k. For the base case we need to show $\mathbf{P}(T_2(n^\beta) < r_1) = \mathcal{O}(n^{-(2-\epsilon)\alpha})$. Recalling $K_\beta^\gamma(u) = \inf\{n > T(u^\beta) \colon |X_n| \le u^\gamma\}$, we have that

$$\begin{aligned}
\mathbf{P}(T_2(u^\beta) < r_1) &= \mathbf{P}(T_2(u^\beta) < K_\beta^\gamma(u)) + \mathbf{P}(T_1(u^\beta) < K_\beta^\gamma(u) < T_2(u^\beta) < r_1) \\
&= \mathbf{P}(T_2(u^\beta) < K_\beta^\gamma(u)) + \mathcal{O}(u^{-(2\beta-\gamma)\alpha}),
\end{aligned} \tag{4.7.38}$$

where $\mathbf{P}(T_1(u^\beta) < K_\beta^\gamma(u) < T_2(u^\beta) < r_1) = \mathcal{O}(u^{-(2\beta-\gamma)\alpha})$ can be deduced by following the arguments as in the proof of Proposition 4.6.1. Applying the dual change of measure $\mathscr{D}$ over the time interval $[0, T_1(u^\beta)]$, we obtain that

$$\begin{aligned}
&u^{(2\beta-\gamma)\alpha}\mathbf{P}(T_2(u^\beta) < K_\beta^\gamma(u)) \\
&= u^{(2\beta-\gamma)\alpha}\mathbf{E}^{\mathscr{D}}\left[e^{-\alpha S_{T(u^\beta)}}\mathbb{1}_{\{T(u^\beta) < r_1\}}\mathbf{P}^{\mathscr{D}}(T_2(u^\beta) < K_\beta^\gamma(u) \,|\, \mathcal{F}_{T(u^\beta)})\right] \\
&= \mathbf{E}^{\mathscr{D}}\left[\mathbb{1}_{\{T(u^\beta) < r_1\}}u^{(\beta-\gamma)\alpha}\mathbf{P}^{\mathscr{D}}(T_2(u^\beta) < K_\beta^\gamma(u) \,|\, \mathcal{F}_{T(u^\beta)})\left|\frac{X_{T(u^\beta)}}{u^\beta Z_{T(u^\beta)}}\right|^{-\alpha}\right].
\end{aligned} \tag{4.7.39}$$

Recalling $\mathfrak{E}_2^\gamma(u) = \{|B_n| \le u^\gamma, \forall 1 \le n < K_\beta^\gamma(u)\}$, we have that, for $|v| \ge 1$

$$\begin{aligned}
&\mathbf{P}^{\mathscr{D}}(T_2(u^\beta) < K_\beta^\gamma(u) \,|\, X_{T(u^\beta)} = vu^\beta) \\
&\le \mathbf{P}^{\mathscr{D}}(|B_n| \le u^\gamma, \forall T(u^\beta) < n < r_1,\, T_2(u^\beta) < K_\beta^\gamma(u) \,|\, X_{T(u^\beta)} = vu^\beta) \\
&\quad + \mathbf{P}^{\mathscr{D}}(\exists T(u^\beta) < n < r_1 \ s.t.\ |B_n| > u^\gamma \,|\, X_{T(u^\beta)} = vu^\beta) \\
&= \mathbf{P}((\mathfrak{E}_2^\beta(u))^c \,|\, X_0 = vu^\beta) = o(u^{-(\beta-\gamma)\alpha})v, \qquad (4.7.40)
\end{aligned}$$

where the tail estimate in (4.7.40) is obtained by following the arguments in the proof of Lemma 4.6.6 and taking advantage of the additional assumption that $\mathbf{E}|B_1|^m < \infty$ for every $m \in \mathbb{Z}_+$. Plugging (4.7.40) into (4.7.39) and using the dominated convergence theorem, we obtain that

$$u^{(2\beta-\gamma)\alpha}\mathbf{P}(T_2(u^\beta) < K_\beta^\gamma(u)) = o(1). \qquad (4.7.41)$$

In view of (4.7.35), (4.7.38), and (4.7.41),

$$\mathbf{P}(T_2(n^\beta) < r_1) = \mathcal{O}(n^{-(2\beta-\gamma)\alpha}) = \mathcal{O}(n^{-(2-\epsilon)\alpha})$$

by choosing $\beta = 1 - \epsilon/3$ and $\gamma = \epsilon/3$. Turning to the inductive step, suppose that $\mathbf{P}(T_k(u^\beta) < r_1) = \mathcal{O}(u^{-(k-\epsilon)\alpha})$. Note that

$$\begin{aligned}
&\mathbf{P}(T_{k+1}(u^\beta) < r_1) \\
&= \mathbf{P}(T_k(u^\beta) < K_\beta^\gamma(u) < T_{k+1}(u^\beta) < r_1) + \mathbf{P}(T_{k+1}(u^\beta) < K_\beta^\gamma(u)),
\end{aligned}$$

where for the first term in the previous sum we have that

$$\begin{aligned}
&\mathbf{P}(T_k(u^\beta) < K_\beta^\gamma(u) < T_{k+1}(u^\beta) < r_1) \\
&\le \mathbf{P}(T_k(u^\beta) < r_1)\mathbf{P}(T(u^\beta) < r_1 | X_0 = u^\gamma) \\
&= \mathcal{O}(u^{-(k-\epsilon')\alpha})\mathcal{O}(u^{-(\beta-\gamma)\alpha}) = \mathcal{O}(u^{-(k+1-\epsilon)\alpha}),
\end{aligned}$$

for suitable choice of β and γ. Hence, it remains to bound $\mathbf{P}(T_{k+1}(u^\beta) < K_\beta^\gamma(u))$. Applying the dual change of measure $\mathscr{D}$ over the time interval $[0, T_1(u^\beta)]$, we obtain that

$$\begin{aligned}
&u^{((k+1)\beta-\gamma)\alpha}\mathbf{P}(T_{k+1}(u^\beta) < K_\beta^\gamma(u)) \\
&= \mathbf{E}^{\mathscr{D}}\left[\mathbb{1}_{\{T(u^\beta)<r_1\}} u^{(k\beta-\gamma)\alpha}\mathbf{P}^{\mathscr{D}}(T_{k+1}(u^\beta) < K_\beta^\gamma(u) \,|\, \mathcal{F}_{T(u^\beta)}) \left|\frac{X_{T(u^\beta)}}{u^\beta Z_{T(u^\beta)}}\right|^{-\alpha}\right]. \qquad (4.7.42)
\end{aligned}$$

Moreover, we have that, for $|v| \geq 1$,

$$
\begin{aligned}
&\mathbf{P}^{\mathscr{D}}(T_{k+1}(u^\beta) < K_\beta^\gamma(u) \,|\, X_{T(u^\beta)} = vu^\beta) \\
&\leq \mathbf{P}^{\mathscr{D}}(\exists T(u^\beta) < n_1 < \cdots < n_k < r_1 \ s.t.\ |B_{n_i}| > u^\gamma,\ \forall i \leq k \,|\, X_{T(u^\beta)} = vu^\beta) \\
&= \mathbf{P}(\exists 0 < n_1 < \cdots < n_k < r_1 \ s.t.\ |B_{n_i}| > u^\gamma,\ \forall i \leq k \,|\, X_0 = vu^\beta) \\
&= \mathbf{P}(\exists 0 < n_1 < \cdots < n_k < r_1 \ s.t.\ |B_{n_i}| > u^\gamma,\ \forall i \leq k) \\
&= o(u^{-(k\beta-\gamma)\alpha}),
\end{aligned} \tag{4.7.43}
$$

where the tail estimate in (4.7.43) is obtained by following the arguments in the proof of Lemma 4.6.6 and taking advantage of the additional assumption that $\mathbf{E}|B_1|^m < \infty$ for every $m \in \mathbb{Z}_+$. Combining (4.7.42) and (4.7.43) with the fact that $|X_{T(u^\beta)}/u^\beta| \leq 1$ we obtain that $\mathbf{P}(T_{k+1}(u^\beta) < K_\beta^\gamma(u))$, and hence, $\mathbf{P}(T_{k+1}(u^\beta) < r_1)$ are of order $\mathcal{O}(u^{-(k+1-\epsilon)\alpha})$.

Part 2): By a similar reasoning as in proving part (1) of Proposition 4.7.2, we have that

$$
\begin{aligned}
&\mathbf{P}(\bar{X}'_n \in (\mathbb{D} \setminus \mathbb{D}^\mu_{\ll j})^{-\gamma},\ d_{M_1'}(\bar{X}'_n, \hat{X}_n) \geq \delta) \\
&\leq \mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\bar{\gamma}}_{\geqslant j},\ \exists i \leq N(n) \ s.t.\ d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta) + o(n^{-j(\alpha-1)}) \\
&= o(n^{-j(\alpha-1)}),
\end{aligned}
$$

where $\mathcal{D}^{\bar{\gamma}}_{\geqslant j}$ is defined as in (4.7.26). It remains to show that, for any $j \geq 1$, $\gamma > 0$, and $\delta > 0$, there exists some $\rho > 0$ so that

$$
\mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^\mu_{\ll j})^{-\gamma},\ \bar{X}'_n \in (\mathbb{D}^\mu_{\ll j})_\rho,\ d_{M_1'}(\bar{X}'_n, \hat{X}_n) \geq \delta) = o(n^{-j(\alpha-1)}),
$$

as $n \to \infty$. Recall, for $\gamma > 0$ and $j \geq 1$, $\mathcal{D}^\gamma_{\geqslant j} = \{\xi \in \mathbb{D} \colon |\mathrm{Disc}_\gamma(\xi)| \geq j\}$, where $\mathrm{Disc}_\gamma(\xi) = \{t \in \mathrm{Disc}(\xi) \colon |\xi(t) - \xi(t^-)| \geq \gamma\}$. Defining $\mathcal{D}^\rho_{=j} = \{\xi \in \mathbb{D} \colon |\mathrm{Disc}_\gamma(\xi)| = j\}$ for $j \in \mathbb{Z}$ and $\rho > 0$, we have that

$$
\begin{aligned}
&\mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^\mu_{\ll j})^{-\gamma},\ \bar{X}'_n \in (\mathbb{D}^\mu_{\ll j})_\rho,\ d_{M_1'}(\bar{X}'_n, \hat{X}_n) \geq \delta) \\
&\leq \mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\rho_0}_{\geqslant j-1},\ d_{M_1'}(\bar{X}'_n, \hat{X}_n) \geq \delta) \\
&\quad + \mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^\mu_{\ll j})^{-\gamma},\ \bar{X}'_n \in (\mathcal{D}^{\rho_0}_{\geqslant j-1})^c) \\
&\leq \mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\rho_0}_{\geqslant j-1},\ d_{M_1'}(\bar{X}'_n, \hat{X}_n) \geq \delta) \\
&\quad + \sum_{i=1}^{j-1} \mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^\mu_{\ll j})^{-\gamma},\ \bar{X}'_n \in \mathcal{D}^{\rho_0}_{=j-i-1})
\end{aligned}
$$

$$= \mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\rho_0}_{\geqslant j-1},\, d_{M_1'}(\bar{X}'_n, \hat{X}_n) \geq \delta) + \sum_{i=1}^{j-1} \mathbf{P}(E_j(i)). \tag{4.7.44}$$

Note that

$$\begin{aligned}
&\mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\rho_0}_{\geqslant j-1},\, d_{M_1'}(\bar{X}'_n, \hat{X}_n) \geq \delta)\\
&\leq \mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\rho_0}_{\geqslant j-1},\, \exists\, i \leq N(n) \ s.t.\ d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta) + o(n^{-j(\alpha-1)})\\
&= \mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\rho_0}_{\geqslant j-1},\, \mathfrak{E}^{\epsilon}_3(n),\, \exists\, i \leq N(n) s.t.\ d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \delta) + o(n^{-j(\alpha-1)})\\
&\leq \mathbf{P}(\exists (i_0, \ldots, i_{j-2}) \in \mathcal{P}(\{1, \ldots, N^+_\epsilon(n)\}, j-1) \text{ s.t.}\\
&\qquad d_{M_1'}(\bar{X}_{i_0,n}, \bar{X}'_{i_0,n}) \geq \delta,\, |X'_{i_p}| \geq n\rho_0, \forall 1 \leq p \leq j-2)\\
&= \mathcal{O}(n^{j-1} n^{-(2-\epsilon)\alpha} n^{-(j-2)\alpha}) + o(n^{-j(\alpha-1)}) = o(n^{-j(\alpha-1)}),
\end{aligned} \tag{4.7.45}$$

where in (4.7.45) we use Lemma 4.7.7 (1) together with the fact that the blocks $\{X_{r_{i-1}}, \ldots, X_{r_i}\}$, $i \geq 1$, are mutually independent, and the final equivalence is obtained by setting $\epsilon < 1/\alpha$. In view of the above computation, it remains to analyze $\mathbf{P}(E_j(k))$, $k \in \{1, \ldots, j-1\}$ as in (4.7.44).

Let $I^* = \{i \leq N(n) \colon d_{M_1'}(\bar{X}_{i,n}, \bar{X}'_{i,n}) \geq \rho_1\}$. Note that

$$\begin{aligned}
&\mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \bar{X}'_n \in \mathcal{D}^{\rho_0}_{=j-k-2})\\
&= \mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \bar{X}'_n \in \mathcal{D}^{\rho_0}_{=j-k-2},\, |I^*| \geq (k+2) \wedge (j-k-2))\\
&\quad + \mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \bar{X}'_n \in \mathcal{D}^{\rho_0}_{=j-k-2},\, 1 \leq |I^*| < (k+2) \wedge (j-k-2))\\
&\quad + \mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \bar{X}'_n \in \mathcal{D}^{\rho_0}_{=j-k-2},\, |I^*| = 0)\\
&= \mathbf{(V.1)} + \mathbf{(V.2)} + \mathbf{(V.3)}.
\end{aligned}$$

Suppose that $k \leq j/2-2$, where the case $k > j/2-2$ can be dealt with similarly. Note that

$$\begin{aligned}
\mathbf{(V.1)} &\leq \mathbf{P}(\bar{X}'_n \in \mathcal{D}^{\rho_0}_{=j-k-2},\, |I^*| \geq k+2,\, \mathfrak{E}^{\epsilon}_3(n)) + o(n^{-j(\alpha-1)})\\
&\leq \mathbf{P}(\exists (i_1, \ldots, i_{j-k-2}) \in \mathcal{P}(\{1, \ldots, N^+_\epsilon(n)\}, j-k-2) \text{ s.t.}\\
&\qquad d_{M_1'}(\bar{X}_{i_p,n}, \bar{X}'_{i_p,n}) \geq \rho, \forall 1 \leq p \leq k+2,\\
&\qquad |X'_{i_q}| \geq n\rho_0, \forall k+3 \leq q \leq j-k-2)\\
&\quad + o(n^{-j(\alpha-1)})\\
&= \mathcal{O}(n^{j-k-2} n^{-(k+2)(2-\epsilon)\alpha} n^{-(j-2k-4)\alpha}) + o(n^{-j(\alpha-1)})\\
&= \mathcal{O}(n^{-j(\alpha-1)} n^{-(k+2)+(k+2)\epsilon\alpha}) + o(n^{-j(\alpha-1)}) = o(n^{-j(\alpha-1)}),
\end{aligned}$$

if $\epsilon < 1/\alpha$. Moreover, we have that $\textbf{(V.3)} = o(n^{-j(\alpha-1)})$ for ρ_0 sufficiently small. Let $I' = \{i \le N(n) \colon \bar{X}'_{i,n} \ge \rho_0\}$. Turning to bounding **(V.2)** we have that

$$\begin{aligned}\textbf{(V.2)} &= \mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \bar{X}'_n \in \mathcal{D}^{\rho_0}_{=j-k-2},\, 1 \le |I^*| \le k+1) \\ &= \sum_{k_1=1}^{k+1} \sum_{k_2=0}^{k_1} \mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \bar{X}'_n \in \mathcal{D}^{\rho_0}_{=j-k-2}, \\ &\qquad\qquad |I^*| = k_1,\, |I' \cap I^*| = k_2,\, \mathfrak{E}^{\epsilon}_3(n)) \\ &\quad + o(n^{-j(\alpha-1)}).\end{aligned}$$

Defining $J = \{(l'_1, \ldots, l'_{k_1}) \colon \mathbf{1}^T(l'_1, \ldots, l'_{k_1}) < k+2+k_2\}$, it is now sufficient to consider

$$\begin{aligned}&\mathbf{P}(\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \bar{X}'_n \in \mathcal{D}^{\rho_0}_{=j-k-2},\, |I^*| = k_1,\, |I' \cap I^*| = k_2,\, \mathfrak{E}^{\epsilon}_3(n)) \\ &\le \mathbf{P}\Bigg(\exists (i_1, \ldots, i_{j-k-2-k_2+k_1}) \in \mathcal{P}(\{1, \ldots, N^+_{\epsilon}(n)\}, j-k-2-k_2+k_1) \text{ s.t.} \\ &\qquad (\bar{X}_{i_1,n}, \ldots, \bar{X}_{i_{k_1},n}) \in \Big(\bigcup_{(l_1,\ldots,l_{k_1}) \in J} \prod_{p=1}^{k_1} \mathbb{D}_{l_{i_p}} \Big)^{-\rho_2}, \\ &\qquad |X'_{i_q}| \ge n\rho_0,\, \forall k_1+1 \le q \le j-k-2-k_2+k_1 \Bigg) \\ &\quad + \mathbf{P}\Bigg(\hat{X}_n \in (\mathbb{D} \setminus \mathbb{D}^{\mu}_{\ll j})^{-\gamma},\, \bar{X}'_n \in \mathcal{D}^{\rho_0}_{=j-k-2},\, |I^*| = k_1,\, |I' \cap I^*| = k_2,\, \mathfrak{E}^{\epsilon}_3(n), \\ &\qquad \exists (i_1, \ldots, i_{j-k-2-k_2+k_1}) \in \mathcal{P}(\{1, \ldots, N^+_{\epsilon}(n)\}, j-k-2-k_2+k_1) \\ &\qquad \text{s.t. } (\bar{X}_{i_1,n}, \ldots, \bar{X}_{i_{k_1},n}) \in \Big(\bigcup_{(l_1,\ldots,l_{k_1}) \in J} \prod_{p=1}^{k_1} \mathbb{D}_{l_{i_p}} \Big)_{\rho_2}, \\ &\qquad |X'_{i_q}| \ge n\rho_0,\, \forall k_1+1 \le q \le j-k-2-k_2+k_1 \Bigg) \\ &= \textbf{(V.2.a)} + \textbf{(V.2.b)}.\end{aligned}$$

Since $0 \le k_2 \le k_1 \le k+1$ we have that

$$\textbf{(V.2.a)} \le \mathcal{O}(n^{j-1})\mathcal{O}(n^{-(k+2+k_2-k_1\delta)\alpha})\mathcal{O}(n^{-(j-k-2-k_2)\alpha})$$

$$= \mathcal{O}(n^{-j(\alpha-1)} n^{k_1\delta\alpha-1}) = o(n^{-j(\alpha-1)}),$$

for $\delta < 1/((k+1)\alpha)$. It remains to show that **(V.2.b)** $= o(n^{-j(\alpha-1)})$. To see this, for $\epsilon > 0$ there exists

$$(\zeta_1, \dots, \zeta_{k_1}) \in \bigcup_{(l_1,\dots,l_{k_1})\in J} \prod_{p=1}^{k_1} \mathbb{D}_{l_{i_p}} \tag{4.7.46}$$

such that $d(\bar{X}_{i_p,n}, \zeta_{i_p}) \leq \rho_2 + \epsilon$, for all $1 \leq p \leq k_1$. Hence, we have that

$$d\left(\hat{X}_n, \bar{X}'_n - \sum_{i\in I'\cap\{i_1,\dots,i_{k_1}\}} \bar{X}'_{i,n} + \sum_{p=1}^{k_1} \zeta_{i_p}\right) \leq \rho_1 \vee (\rho_2 + \epsilon). \tag{4.7.47}$$

For any $c > 0$, define $\Phi_c \colon \mathbb{D} \to \mathbb{D}$ by

$$\Phi_c(\xi)(t) = \sum_{s\in[0,t]\cap \mathrm{Disc}(\xi,c)} (\xi(s) - \xi(s^-)), \quad \text{for } t \in [0,1], \tag{4.7.48}$$

where $\mathrm{Disc}(\xi, c)\{t \in \mathrm{Disc}(\xi) \colon \xi(t) - \xi(t^-) \geq c\}$. Now we claim that

$$\|\bar{X}'_n - \Phi_{\rho_0}(\bar{X}'_n) - \mu \cdot \mathrm{id}\|_\infty > \rho_3. \tag{4.7.49}$$

To see this, suppose $\|\Phi_{\rho_0}(\bar{X}'_n) - \mu \cdot \mathrm{id}\|_\infty \leq \rho_3$. Hence,

$$\begin{aligned} &d\left(\bar{X}'_n - \sum_{i\in I'\cap\{i_1,\dots,i_{k_1}\}} \bar{X}'_{i,n} + \sum_{p=1}^{k_1} \zeta_{i_p}, \mu\cdot\mathrm{id} + \sum_{p=1}^{k_1}\zeta_{i_p} + \sum_{i\in I'\setminus\{i_1,\dots,i_{k_1}\}} \bar{X}'_{i,n}\right) \\ &\leq \left\|\bar{X}'_n - \sum_{i\in I'} \bar{X}'_{i,n} - \mu\cdot\mathrm{id}\right\|_\infty = \|\bar{X}'_n - \Phi_{\rho_0}(\bar{X}'_n) - \mu\cdot\mathrm{id}\|_\infty \leq \rho_3. \end{aligned} \tag{4.7.50}$$

In view of (4.7.47) and (4.7.50) we obtain that

$$d\left(\hat{X}_n, \mu\cdot\mathrm{id} + \sum_{p=1}^{k_1}\zeta_{i_p} + \sum_{i\in I'\setminus\{i_1,\dots,i_{k_1}\}} \bar{X}'_{i,n}\right) \leq \rho_1 \vee (\rho_2+\epsilon) + \rho_3,$$

where

$$\mu\cdot\mathrm{id} + \sum_{p=1}^{k_1}\zeta_{i_p} + \sum_{i\in I'\setminus\{i_1,\dots,i_{k_1}\}} \bar{X}'_{i,n} \in \mathbb{D}^\mu_{\ll j}$$

due to (4.7.46). This leads to the contradiction of $\hat{X}_n \in (\mathbb{D}\setminus\mathbb{D}^{\mu}_{\ll j})^{-\gamma}$ by choosing ρ_1, ρ_2 and ρ_3 small enough. In view of (4.7.49) we have that

$$\begin{aligned}\textbf{(V.2.b)} &\le \mathbf{P}\left(\bar{X}'_n \in \left\{\xi \in \mathbb{D}\colon \xi(t) - \sup_{t\in[0,1]}\left|\Phi_{\rho_0}(\xi)(t) - \mu t\right| > \rho_3\right\}\right)\\ &= o(n^{-j(\alpha-1)}),\end{aligned}$$

by choosing ρ_0 and ρ_3 such that $\rho_3/\rho_0 \notin \mathbb{Z}$ and $\lceil \rho_3/\rho_0\rceil > j$.

Part 3): Since

$$\mathbf{P}(r_{i+1} - r_i > r_i\delta) \le \mathbf{P}(r_{i+1} - r_i > (n-\epsilon')\delta) + \mathbf{P}(r_i \ge n - \epsilon'),$$

$\mathbf{P}(r_{i+1}-r_i > r_i\delta)$ decays exponentially, for $i \in \{N^-_\epsilon(n), \dots, N^+_\epsilon(n)\}$. Combining this with (4.7.36), we are able to utilize the argument as in (4.7.37) and obtain that

$$\mathbf{P}(R_{i,n} \in (\mathbb{D}\setminus\mathbb{D}_{\ll j})^{-\delta}) = \mathbf{P}(T_j(n^\beta) < r_1) + o(n^{-(j-\epsilon)\alpha})$$

for any $\beta \in (0,1)$. Since $\mathbf{P}(T_j(u^\beta) < r_1) = \mathcal{O}(u^{-(j-\epsilon)\alpha})$ for a suitable choice of β, the proof is completed. □

4.8 Proofs of Section 4.3.4

Proof of Theorem 4.3.4. First we generalize the preliminaries in Section 4.2 to the current setting. Note that, thanks to Assumption 4.3.1, Result 4.2.1 and Lemma 4.2.1 can be extended to the current setting. By ($\mathfrak{C}$), it is not difficult to see that Lemma 4.2.2 (1) holds for X_n as in (4.3.11). We want to show that $Z_n \xrightarrow{\text{a.s.}} Z$ as $n\to\infty$. For $n > m$ we have

$$\begin{aligned}Z_n - Z_m &= e^{-S_n} f_n(X_{n-1}) - e^{-S_m}X_m \le e^{-S_{n-1}}X_{n-1} + \overline{B}_n e^{-S_n} - e^{-S_m}X_m\\ &\le e^{-S_m}X_m + \sum_{j=m+1}^{n} \overline{B}_j e^{-S_j} - e^{-S_m}X_m = \sum_{j=m+1}^{n}\overline{B}_j e^{-S_j},\end{aligned}$$

and similarly

$$Z_n - Z_m \ge \sum_{j=m+1}^{n} \underline{B}_j e^{-S_j}.$$

Thus we have that

$$|Z_n - Z_m| \leq \sum_{j=m+1}^{n} (|\underline{B}_j| \vee |\overline{B}_j|) e^{-S_j}.$$

Using the fact that S_n has a positive drift under the α-shifted measure and $\mathbf{E}|\underline{B}|^\alpha + \mathbf{E}|\overline{B}|^\alpha < \infty$, we conclude that $|Z_n - Z_m| \xrightarrow{\text{a.s.}} 0$ as n, $m \to \infty$. This implies that $Z_n \xrightarrow{\text{a.s.}} Z$ for some real-valued random variable Z.

To show Theorem 4.3.4, we need to extend the results in Section 4.6. This can be achieved by using Assumption 4.3.1, especially by utilizing the facts that X_n is monotone in its initial state and is bounded from both below and above by two AR(1) processes. $\square$

Proof of Theorems 4.3.5 and 4.3.6. The proofs are omitted since they are almost identical to the proofs of Theorems 4.3.2 and 4.3.3. $\square$

Chapter 5

Importance sampling of iterated functions

5.1 Introduction

We consider an $\mathbb{R}$-valued Markov chain $\{Z_n\}_{n\in\mathbb{N}}$ defined by

$$Z_{n+1} = \Psi_{n+1}(Z_n), \tag{5.1.1}$$

where $\{\Psi_n\}_{n\in\mathbb{N}}$ is a sequence of independent and identically distributed (i.i.d.) positive random Lipschitz functions (see (5.3.8) below); $Z_0 \in \mathbb{R}$ is independent of the sequence $\{\Psi_n\}_{n\in\mathbb{N}}$. Under mild conditions (see Assumption B1 below), the stationary solution to (5.1.1) has the same distribution as the almost sure limit Z of the sequence $\{\Psi_1 \circ \cdots \circ \Psi_n(Z_0)\}_{n\in\mathbb{N}}$ (for details see [50]). We assume that Ψ_n is such that $\Psi_1 \circ \cdots \circ \Psi_n(Z_0)$ is increasing in n. This chapter develops efficient simulation methods for estimating the tail probability of Z, i.e. we are interested in computing $\mathbf{P}(Z > x) = \mathbf{P}(T(x) < \infty)$ for large x, where $T(x) = \inf\{n \geq 0 : \Psi_1 \circ \cdots \circ \Psi_n(Z_0) > x\}$.

There are two examples that are of particular interest in this setting. The first example is the so-called stochastic perpetuity. More precisely, consider the random difference equation, where $\Psi_n(z) = A_n z + B_n$. The recursion (5.1.1) can be written as

$$Z_{n+1} = A_{n+1} Z_n + B_{n+1}, \tag{5.1.2}$$

where $\{(A_n, B_n)\}_{n\geq 0}$ is a sequence of i.i.d. $\mathbb{R}$-valued random vectors, independent of the initial random variable Z_0. Moreover, noting $T(x) = \inf\{n \geq$

$0\colon \sum_{k=0}^{n} B_{k+1}e^{S_k} > x\}$, the objective here is to estimate $\mathbf{P}(Z > x)$, for which we have the identity $\mathbf{P}(Z > x) = \mathbf{P}(T(x) < \infty)$ in case B_n is positive. Perpetuities occur in the context of ruin problems with investments, in the study of financial time series such as ARCH-type processes (see e.g. [54]), in tail asymptotics for exponential functionals of Lévy processes (see e.g. [88]), etc. Although some particular cases exist that allow for an explicit analysis (see e.g. [116]), it is hard to come up with exact results for the distribution of Z in general. Thus, Monte Carlo simulation arises as a natural approach to deal with the analysis of stochastic perpetuities, including the large deviations regime where x in $\mathbf{P}(Z > x)$ is large, which is the focus of this chapter.

Another example of (5.1.1) is the Lindley recursion that describes the waiting time of a customer in a single-server queue. More precisely, we consider (5.1.1) with $\Psi_n(z) = \max\{0, z + X_n\}$, where X_n, $n \in \mathbb{N}$, is a sequence of i.i.d. $\mathbb{R}$-valued random variables. It is well known (see e.g. [61]) that the stationary solution of the Markov chain $Z_{n+1} = \max\{0, Z_n + X_{n+1}\}$, $n \in \mathbb{N}$, has as representation of the all-time supremum of a random walk, denoted by $\sup_{n\geq 0} S_n$, where $S_n = \sum_{i=1}^{n} X_i$. A similar connection holds of course between e^{Z_n} and $e^{Z_{n+1}}$ in this context. The exponentiated form of the Lindley recursion is actually more suitable for our purposes: by developing a connection between iterated random functions (perpetuities) and the supremum of a random walk, we utilize rare-event simulation techniques for estimating $\mathbf{P}(\sup_{n\geq 0} S_n > x)$ to construct efficient simulation algorithms for estimating the tail probability of the stationary solutions to (5.1.1) and (5.1.2), in a heavy-tailed setup.

Before we give a more precise description of our results, we first mention some related work. In the more general context of iterated random functions, the tail behavior of Z has been studied, for example, in [66], where sufficient conditions of Ψ_n are given for which $\mathbf{P}(Z > x)$ behaves like a power law. A more recent study in this direction is [93]. The main result of [50], which is the most related one to the chapter, states that the tail of Z is slowly varying under the same setting as this chapter (see Assumptions A1–A2 and Assumptions B1–B2 below). The result in [50] is an extension of a classical result on the supremum of random walks by [98] and [115].

Turning to the special case of stochastic perpetuities, sufficient conditions for $\mathbf{P}(Z > x)$ to decay at an exponential rate have been established in [67], where it is assumed that $|A_1|$ is bounded by 1, and the moment generating function of B_1 exists in a neighborhood of the origin. On the other hand, [81] and later on [66] assumed that $\mathbf{E}|A_1|^\alpha = 1$ and $\mathbf{E}|B_1|^\alpha < \infty$ for some $\alpha > 0$, and proved that Z has a power law with exponent α. Moreover, the result of [69], which was generalized by [68], states that the tail of Z is regularly varying of some index, say $-\alpha$, if B_1 is regularly varying of the same index $-\alpha$ and $\mathbf{E}A_1^\alpha < 1$. For a more extensive overview of the literature on this topic see e.g. [26] and references therein.

A study on rare-event simulation that is of primary interest to us is [16], where the authors designed an algorithm for estimating the tail probability of the all-time supremum of heavy-tailed random walks. A major contribution of this chapter is that the algorithm of [16] is extended to the more general setting of [50]. Another study on rare-event simulation for perpetuities and iterated random functions is [7], in which deterministic interest rates are considered. In [23], the authors estimate the tail probability of perpetuities with deterministic premiums (B_n). Later, [20] develops simulation algorithms for perpetuities in the setting where both the discounting factor and premiums are driven by a Markov chain. Furthermore, [34] provides simulation estimators for the tail distribution of Z as in (5.1.1) with $\Psi_n(z) = A_n \max\{z, D_n\} + B_n$.

The contributions of this chapter are the following. For stochastic perpetuities, we propose a strongly efficient simulation algorithm for estimating $\mathbf{P}(Z > x)$. For this, we need to make several assumptions (see Assumptions A1–A5 below). We illustrate the generality of these assumptions by giving examples as well as sufficient conditions in Remarks 5.3–5.7 below. We construct an upper bounding random walk for the stochastic perpetuity, which leads to an asymptotic result for the tail probability of Z under a heavy-tailed assumption on $\log \max\{A_1, B_1\}$. Note that Z is defined over an infinite horizon, and hence, requires an infinite amount of computational effort for generating each sample when using a crude Monte Carlo sampling approach. A natural approach to address such an issue is to work with approximations by finite-time truncation. We study the bias introduced by such approximations and show that our estimator has a vanishing relative bias as $x \to \infty$. By making a slightly stronger, but not restrictive assumption (see Assumption A5 below), we are able to identify the rate at which the bias decays w.r.t. the truncation time. Applying the bias elimination technique studied in [106], we then propose strongly efficient and *unbiased* estimators for $\mathbf{P}(Z > x)$. We finally extend these results to the more general setting as in (5.1.1). In Section 5.3.2, we make a couple of extra assumptions (see Assumptions B0–B1 below) on Ψ_n. Our setting is almost identical to the ones in [66], [37], and [50]. In Remark 5.9 below, we give examples that satisfy our assumptions.

The most important idea behind our results is the following. We connect our class of iterated random function to the supremum of a random walk. To illustrate this idea consider (5.1.2) with $B_n = 1$. We now connect the stochastic perpetuity with the supremum of random walk by observing that, for $\gamma \in (0, -\mathbf{E}S_1)$

$$Z = \sum_{n=0}^{\infty} \exp\{S_n + n\gamma\} \exp(-n\gamma) \leq \exp\left\{\max_{n\geq 0}(S_n + n\gamma)\right\} \frac{1}{1-e^{-\gamma}}, \quad (5.1.3)$$

where $S_n = \sum_{i=1}^{n} \log A_i$. The upper bounding random walk constructed in (5.1.3) allows us to construct a coupling and leverage the importance sampling algorithm designed in [16]. It turns out that we can extend this idea to the general setting of (5.1.1) by constructing a slightly more involved upper bounding random walk. Note that our extension of (5.1.3) leads to a shorter proof of the asymptotic upper bound given in [50] which we believe to be of independent interest.

The rest of the chapter is organized as follows. Section 5.2 deals with basic background information and notation required to state our contributions. Our main results are stated in Section 5.3, where we start with the case of stochastic perpetuity and then extend all results to the iterated function setting. Numerical results are presented in Section 5.4. All proofs can be found in Section 5.5.

5.2 Notations and preliminary results

In this section we will start first with a list of notations that will be employed in this chapter, then we will recall some preliminary results from the literature.

For $(x, y) \in \mathbb{R}^2$, let $x \wedge y \triangleq \min\{x, y\}$ and $x \vee y \triangleq \max\{x, y\}$. For $x \in \mathbb{R}$, let $x^+ = x \vee 0$ denote the positive part of x and let $\log^+ x = 0 \vee \log x = \log(x \vee 1)$. Let $c \in \mathbb{R} \cup \{\pm\infty\}$, let $f(x)$ and $g(x)$ be non-negative real-valued functions. We write $f(x) \sim g(x)$, $f(x) = o(g(x))$, and $f(x) = \mathcal{O}(g(x))$, as $x \to c$, if $\lim_{x\to c} f(x)/g(x) = 1$, $\lim_{x\to c} f(x)/g(x) = 0$, and $\limsup_{x\to c} f(x)/g(x) < \infty$, respectively.

To describe the efficiency of a rare-event simulation algorithm, we adopt a widely applied criterion (for a discussion of efficiency in rare-event simulation, see e.g. [5]). Suppose that we are interested in a sequence of rare events $\mathcal{E}(x)$ that become more rare as $x \to \infty$. Let $L(x)$ be an unbiased estimator of the rare-event probability $\mathbf{P}(\mathcal{E}(x))$. $L(x)$ is said to be strongly efficient if $\mathbf{E}L(x)^2 = \mathcal{O}(\mathbf{P}(\mathcal{E}(x))^2)$ as $x \to \infty$. In particular, strong efficiency implies that the number of simulation runs required to estimate the target probability to a given relative accuracy is bounded w.r.t. x.

As we have mentioned in the introduction, a state-dependent importance sampling scheme will be used in this chapter. We recall the following result that will be very useful in validating our new estimator.

Result 5.2.1. [65, Section 3.9] *Let $Y_n, n \in \mathbb{N}$, be a sequence of random variables on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Let $M_n, n \in \mathbb{N}$, be a non-negative martingale that is adapted to Y_n for which $\mathbf{E}M_0 = 1$. Let Γ be a stopping time adapted to Y_n. Define a sequence of probability measures as $\mathbf{P}_n(A') = \mathbf{E}\mathbb{1}_{A'}M_n$, for $A' \in \mathcal{F}_n \triangleq \sigma(Y_1, \ldots, Y_n)$. Then there exists a probability measure $\tilde{\mathbf{P}}$, such that $\tilde{\mathbf{P}}(A') = \mathbf{P}_n(A')$, for $A' \in \mathcal{F}_n$ and $n \in \mathbb{N}$. Furthermore, we have that $\mathbf{E}\mathbb{1}_{\{\Gamma<\infty\}} = \tilde{\mathbf{E}}\mathbb{1}_{\{\Gamma<\infty\}}M_\Gamma^{-1}$.*

Next, we recall a simulation algorithm proposed in [16], where the authors develop an efficient state-dependent importance sampling strategy for estimating the tail probability of a random walk crossing a certain level. Before we go through the details of the simulation algorithm, we introduce the following definition.

Definition 5.2.1. Let Y be a random variable on $\mathbb{R}$. Let the integrated tail of Y, as a function of x, be defined by

$$x \mapsto 1 \wedge \int_x^\infty \mathbf{P}(Y > t)dt.$$

We say that Y is long tailed, if for every $c \in \mathbb{R}$, we have that

$$\mathbf{P}(Y > t + c) \sim \mathbf{P}(Y > t), \quad \text{as } t \to \infty.$$

We say that Y is subexponential, if

$$\mathbf{P}(Y_{(1)}^+ + Y_{(2)}^+ > t) \sim 2\mathbf{P}(Y^+ > t), \quad \text{as } t \to \infty,$$

where $Y_{(1)}$ and $Y_{(2)}$ are independent copies of Y. Moreover, we say that Y is strongly subexponential, or Y belongs to the class S^*, if

$$2\mathbf{E}Y^+\mathbf{P}(Y > t) \sim \int_0^t \mathbf{P}(Y > t - s)\mathbf{P}(Y > s)ds, \quad \text{as } t \to \infty.$$

Remark 5.1. If Y belongs to S^*, both the distribution of Y and its integrated tail are subexponential (cf. [82, Theorem 3.2]) and, in particular, long-tailed.

Consider a random walk $\{S_n\}_{n\in\mathbb{N}}$ generated by a sequence of i.i.d. random variables $\{X_n\}_{n\in\mathbb{N}}$, i.e., $S_n = \sum_{i=1}^n X_i$. Assume that $\mathbf{E}X_1 < 0$ and X_1 belongs to S^*. Let $P(y, dz)$ denote the transition kernel of the random walk $\{S_n\}_{n\in\mathbb{N}}$. Define a non-negative random variable W that is independent of $\{X_n\}_{n\in\mathbb{N}}$ with tail probability

$$\mathbf{P}(W > y) \triangleq \min\left[1, -\frac{1}{\mathbf{E}X_1}\int_y^\infty \mathbf{P}(X_1 > t)dt\right].$$

Fix $x > 0$. To estimate $\mathbf{P}(\sup_{n\geq 0} S_n > x) = \mathbf{P}(\tau(x) < \infty)$, where $\tau(x) = \inf\{n \geq 0 : S_n > x\}$, Blanchet and Glynn suggest in [16] simulating the random walk via another transition kernel

$$Q_{a_*}(y, dz) \triangleq P(y, dz)\frac{v(z + a_*)}{w(y + a_*)}, \quad \forall y \in (-\infty, x],\ z \in \mathbb{R}, \tag{5.2.1}$$

where

$$v(z) \triangleq \mathbf{P}(W > -(z - x)), \qquad w(y) \triangleq \mathbf{P}(X_1 + W > -(y - x)), \tag{5.2.2}$$

and a_* is such that, for fixed $\delta \in (0, 1)$

$$-\delta \leq \frac{v^2(y) - w^2(y)}{\mathbf{P}(X_1 > -y)w(y)}, \quad \forall y \leq x + a_*. \tag{5.2.3}$$

Let $\mathbf{P}^{Q_{a_*}}$ and $\mathbf{E}^{Q_{a_*}}$ denote respectively the probability measure and the expectation w.r.t. the random process $\{S_n\}_{n\in\mathbb{N}}$ having a one-step transition kernel $Q_{a_*}(y, dz)$ as in (5.2.1). In the following result, we give the simulation estimator proposed in [16], which will prove to be useful in our context.

Result 5.2.2. [16, Theorem 3] *Suppose that $\mathbf{E}X_1 < 0$ and X_1 belongs to S^*. Let v and w be defined as in* (5.2.2). *For fixed $\delta \in (0, 1)$, there exists an $a_* = a_*(\delta) \leq 0$ such that* (5.2.3) *holds. Then*

$$L_\tau(x) = \mathbb{1}_{\{\tau(x)<\infty\}} \prod_{k=1}^{\tau(x)} \frac{w(S_{k-1} + a_*)}{v(S_k + a_*)}$$

is an unbiased estimator of $\mathbf{P}(\sup_{n\geq 0} S_n > x)$ under $\mathbf{P}^{Q_{a_*}}$; moreover, it is strongly efficient, i.e.,

$$\sup_{x>0} \frac{\mathbf{E}^{Q_{a_*}} L_\tau^2(x)}{\mathbf{P}\left(\sup_{n\geq 0} S_n > x\right)^2} < \infty.$$

Remark 5.2. The existence of such an a_* as in Result 5.2.2 is guaranteed by the fact (for details see [16, Proposition 3]) that

$$w(y) - v(y) = o(\,\mathbf{P}(X_1 > -y)\,), \quad \text{as } y \to -\infty. \tag{5.2.4}$$

We will extend this algorithm to the setting of (5.1.1). Unfortunately, it is not straightforward to generate our estimator, say, L such that $\mathbf{E}L = \mathbf{P}(Z > x)$

in a finite computer time. However, there exists a sequence L_n, $n \in \mathbb{N}$, of $\mathcal{L}^2$ approximations (i.e., $\mathbf{E}[(L_n - L)^2] \to 0$, as $n \to \infty$) that can be generated exactly in a finite time. This kind of situation has been considered in [106]. We recall here one of the main results in [106], which will prove to be crucial for our purposes.

Result 5.2.3. [106, Theorem 2] *Let L_n and L be such that $\mathbf{E}[(L_n - L)^2] \to 0$, as $n \to \infty$. Let N be a non-negative integer-valued random variable, independent of L_n, $n \in \mathbb{N}$, such that $\mathbf{P}(N \geq n) > 0$ for all $n \geq 0$. If*

$$\sum_{n=1}^{\infty} \frac{\mathbf{E}[(L_n - L)^2]}{\mathbf{P}(N \geq n)} < \infty,$$

then $\bar{L}$ defined by

$$\bar{L} \triangleq \sum_{n=0}^{N} \frac{L_n - L_{n-1}}{\mathbf{P}(N \geq n)}$$

(with $L_{-1} = 0$) is an unbiased estimator of $\mathbf{E}L$, and

$$\mathbf{E}[\bar{L}^2] = \sum_{n=0}^{\infty} \frac{\mathbf{E}[(L_{n-1} - L)^2] - \mathbf{E}[(L_n - L)^2]}{\mathbf{P}(N \geq n)} < \infty.$$

In order to apply Result 5.2.3 in our context, we conclude this section with the following extension of Result 5.2.2, which basically says the algorithm proposed in [16] gives us an estimator with bounded relative $(2+\epsilon)$-th moment, for some $\epsilon > 0$. The proof of this lemma together with other results presented in this chapter can be found in Section 5.5 below.

Lemma 5.2.1. *Let $S_n = \sum_{i=1}^{n} X_i$ be a random walk. Suppose that $\mathbf{E}X_1 < 0$ and X_1 belongs to S^*. Let v and w be defined as in (5.2.2). For any fixed $\epsilon > 0$ and $\delta \in (0,1)$, there exists an $a_* = a_*(\epsilon, \delta) \leq 0$ such that*

$$-\delta \leq \frac{v^{2+\epsilon}(y) - w^{2+\epsilon}(y)}{\mathbf{P}(X_1 > -y + x) w^{1+\epsilon}(y)}, \quad \forall y \leq x + a_*.$$

Let

$$L_\tau(x) \triangleq \mathbb{1}_{\{\tau(x) < \infty\}} \prod_{k=1}^{\tau(x)} \frac{w(S_{k-1} + a_*)}{v(S_k + a_*)}.$$

Then, $\mathbf{E}^{Q_{a_*}} L_\tau(x) = \mathbf{P}(\sup_{n \geq 0} S_n > x)$ and

$$\sup_{x>0} \frac{\mathbf{E}^{Q_{a_*}} L_\tau^{2+\epsilon}(x)}{\mathbf{P}(\sup_{n \geq 0} S_n > x)^{2+\epsilon}} < \infty.$$

5.3 Main results

In this section we present the main results of this chapter. In Section 5.3.1 we consider the stochastic perpetuity as in (5.1.2). Recall that Z_n, $n \in \mathbb{N}$, was defined by

$$Z_{n+1} = A_{n+1}Z_n + B_{n+1}, \qquad \text{for } n \in \mathbb{N}.$$

Recalling $Z = \sum_{n=0}^{\infty} B_{n+1}e^{S_n}$ and $S_n = \sum_{i=1}^{n} \log A_i$, we are interested in estimating $\mathbf{P}(Z > x)$, where x is large. For this, several assumptions need to be made (see Assumptions A1–A5 below). We will discuss the generality of these assumptions by giving examples as well as sufficient conditions in Remarks 5.3–5.7 below. To construct our simulation estimator, we will find a stochastic upper bound that can be written as a functional of a suitable random walk $S_n(\gamma)$. Then, using this upper bound we define a crossing level $s(x)$ and a stopping time $\tau_\gamma(x) = \inf\{n \geq 0 : S_n(\gamma) > s(x)\}$, such that $\{Z > x\} \subseteq \{\tau_\gamma(x) < \infty\}$. Since the change of measure proposed in [16] is strongly efficient for estimating the tail probability of the supremum of heavy-tailed random walks, a natural strategy is to keep track of the random process $S_n(\gamma), n \in \mathbb{N}$, while simulating the sequence $\sum_{k=0}^{n} B_{k+1}e^{S_k}, n \in \mathbb{N}$, until the stopping time $\tau_\gamma(x)$. By doing this, we can construct a state-dependent change of measure using the path of the random walk until $\tau_\gamma(x)$ according to the method introduced in Section 5.2. Then we continue to simulate the path of the random walk after $\tau_\gamma(x)$ under the original measure. Based on this idea, we will propose a simulation algorithm for estimating $\mathbf{P}(Z > x)$ and discuss its properties such as strong efficiency in the rest of Section 5.3.1. In Section 5.3.2 we extend the results from Section 5.3.1 to the general setting, where Z_n, $n \in \mathbb{N}$, was defined in (5.1.1) as

$$Z_{n+1} = \Psi_{n+1}(Z_n), \qquad \text{for } n \in \mathbb{N},$$

and $\{\Psi_n\}_{n\in\mathbb{N}}$ is a sequence of i.i.d. random functions that is independent of Z_0. Note that all the proofs of the results in this section are given in Section 5.5 below.

5.3.1 Stochastic perpetuity

We consider the Markov chain Z_n, $n \in \mathbb{N}$, given by (5.1.2). Recall that

$$Z_{n+1} = A_{n+1}Z_n + B_{n+1}, \quad n \in \mathbb{N}.$$

To guarantee the positive recurrence of $\{Z_n\}_{n\in\mathbb{N}}$, we assume the following.

Assumption A1. a) $A_1 > 0$ a.s., $\mathbf{E}\log A_1 < 0$ and $\mathbf{E}\log^+|B_1| < \infty$.

b) $\mathbf{E}\log^+(A_1 \vee B_1) < \infty$.

c) $\mathbf{P}(A_1 > x, B_1 \leq -x) = o(\mathbf{P}(A_1 \vee B_1 > x))$.

Recall that, under Assumption A1, the unique stationary distribution of this Markov chain exists, has right-unbounded support and has the same distribution as the random variable $Z \triangleq \sum_{n=0}^{\infty} B_{n+1}e^{S_n}$, where $S_n = \sum_{i=1}^{n}\log A_i$; see, for example, [66] or Chapter 2 of [26] for more detail. As mentioned in the beginning of Section 5.3, we start with developing a connection between perpetuities and the supremum of a random walk. More precisely, we construct an upper bound for Z that can be written as a functional of a suitable random walk $S_n(\gamma)$. We formulate the result in the following lemma.

Lemma 5.3.1. *Let Assumption A1 hold. There exists a constant γ_2 such that*

$$\mathbf{E}\left[(\log^+ B_1^+ - \gamma_2) \vee \log A_1\right] < 0.$$

Moreover, there exists a constant $\gamma_1 \in (0, -\mathbf{E}[\log A_i \vee (\log^+ B_1^+ - \gamma_2)])$ such that

$$Z \leq \exp\left\{\sup_{n\geq 0} S_n(\gamma)\right\}\frac{e^{\gamma_2}}{1-e^{-\gamma_1}} < \infty, \tag{5.3.1}$$

where $S_n(\gamma) = S_n(\gamma_1,\gamma_2) = \sum_{i=1}^{n}[\log A_i \vee (\log^+ B_i^+ - \gamma_2) + \gamma_1]$ and $\mathbf{E}S_1(\gamma) < 0$.

Now from (5.3.1) we can define $s(x) \triangleq \log x - \gamma_2 + \log(1-e^{-\gamma_1})$ and $\tau_\gamma(x) \triangleq \inf\{n \geq 0 : S_n(\gamma) > s(x)\}$, such that the following holds:

$$\{Z > x\} \subseteq \left\{\sup_{n\geq 0} S_n(\gamma) > s(x)\right\}. \tag{5.3.2}$$

As we will see in the proof of Theorem 5.3.2, the asymptotic behavior of $\mathbf{P}(Z > x)$ as $x \to \infty$ will be useful in establishing the strong efficiency of our estimator. Thus, we derive a tail estimate for Z in Theorem 5.3.1 below. To be precise, we are interested in finding a function $f(x)$ such that $\mathbf{P}(Z > x) = \mathcal{O}(f(x))$ as $x \to \infty$. Moreover, we focus on the case, where the following assumption holds.

Assumption A2. The integrated tail (see Definition 5.2.1 above) of $\log(A_1 \vee B_1)$, denoted by $\bar{F}_I$, is subexponential.

Remark 5.3. As mentioned in the introduction, the focus of this chapter is to propose Monte-Carlo estimators for $\mathbf{P}(Z > x)$, which is slowly varying as $x \to \infty$. Indeed, $\mathbf{P}(Z > x)$ is slowly varying under Assumptions A1–A2. A proof can be found e.g. in [50]; in Theorem 5.3.1 (and Theorem 5.3.5) below we give an independent proof for the asymptotic upper bound of $\mathbf{P}(Z > x)$ (under a general setting, see Assumptions B1–B2).

Theorem 5.3.1. *If Assumptions A1 and A2 hold, then*

$$\limsup_{x\to\infty} \frac{\mathbf{P}(Z > x)}{\bar{F}_I(\log(x))} \le -\frac{1}{\mathbf{E}\log A_1}.$$

By constructing the upper bound as in Lemma 5.3.1 above, we have established a connection between perpetuities and the supremum of a random walk. This connection will allow us to utilize rare-event simulation techniques for estimating $\mathbf{P}(\sup_{n\ge 0} S_n > x)$ in designing an efficient simulation estimator for $\mathbf{P}(Z > x)$. To construct our simulation estimator of $\mathbf{P}(Z > x)$, define a non-negative random variable W_γ that is independent of $\{(A_n, B_n)\}_{n\in\mathbb{N}}$ with tail probability

$$\mathbf{P}(W_\gamma > t) \triangleq \min\left[1\,,\ \frac{-1}{\mathbf{E}S_1(\gamma)}\int_t^\infty \mathbf{P}(S_1(\gamma) > s)ds\right],$$

define

$$v_\gamma(z) \triangleq \mathbf{P}(W_\gamma > -(z - s(x))), \quad \text{and} \quad w_\gamma(y) \triangleq \mathbf{P}(S_1(\gamma) + W_\gamma > -(y - s(x))). \tag{5.3.3}$$

Let $P^\gamma(y, dz)$ denote the transition kernel of the random walk $\{S_n(\gamma)\}_{n\in\mathbb{N}}$. For fixed a_*, let $\mathbf{E}^{Q^\gamma_{a_*}}$ denote the expectation w.r.t. the stochastic process $\{S_n(\gamma)\}_{n\in\mathbb{N}}$ having a one-step transition kernel

$$Q^\gamma_{a_*}(y, dz) = \begin{cases} P^\gamma(y, dz) v_\gamma(z + a_*) w_\gamma(y + a_*)^{-1}, & \text{for } n \le \tau_\gamma(x), \\ P^\gamma(y, dz), & \text{for } n > \tau_\gamma(x). \end{cases} \tag{5.3.4}$$

We propose an estimator and show its strong efficiency in Theorem 5.3.2 below. We make a slightly stronger assumption on the tail asymptotics of $\log(A_1 \vee B_1)$:

Assumption A3. The distribution of $\log(A_1 \vee B_1)$ belongs to the class S^*.

Remark 5.4. Assumption A3 is not restrictive in the sense that the class S^* of strongly subexponential random variables includes regularly varying, lognormal and Weibull-type distributions, among many others. For more properties of strongly subexponential distributions we refer to Section 3.4 of [60].

Theorem 5.3.2. *Let Assumptions A1 and A3 hold. Let v_γ and w_γ be defined as in* (5.3.3)*. For fixed $\delta \in (0,1)$, there exists an $a_* = a_*(\delta) \leq 0$ such that*

$$-\delta \leq \frac{v_\gamma^2(y) - w_\gamma^2(y)}{\mathbf{P}(X_1 > -y + s(x))w_\gamma(y)}, \quad \forall y \leq s(x) + a_*. \tag{5.3.5}$$

Let

$$L_T(x) \triangleq \mathbb{1}_{\{T(x)<\infty\}} \prod_{k=1}^{\tau_\gamma(x)} \frac{w_\gamma(S_{k-1}(\gamma) + a_*)}{v_\gamma(S_k(\gamma) + a_*)}. \tag{5.3.6}$$

Then $L_T(x)$ is an unbiased and strongly efficient estimator of $\mathbf{P}(Z > x)$, i.e.,

$$\sup_{x>1} \frac{\mathbf{E}^{Q^\gamma_{a_*}} L_T^2(x)}{\mathbf{P}(Z > x)^2} < \infty.$$

The estimator derived in Theorem 5.3.2 requires the computation of $\mathbb{1}_{\{Z>x\}}$, and hence, is unbiased only if we can generate Z in finite time. Generating a perfect sample from Z in our current setting is not straightforward. To address this issue, we will apply the bias elimination technique introduced in [106]. The plan for the rest of this section can be described as follows. First we propose a family of simulation algorithms by approximating the path $\{Z_n\}_{n>\tau_\gamma(x)}$ with $\{Z_n\}_{\tau_\gamma(x)<n\leq\tau_\gamma(x)+M}$ for a fixed and sufficiently large M; we show that the latter family of simulation algorithms gives biased estimators with vanishing relative bias as $x \to \infty$; consequently, we are able to apply the bias elimination technique and obtain an unbiased estimator that is strongly efficient and runs in finite time. To begin with, note that

$$Z = \sum_{n=0}^{\tau_\gamma(x)} B_{n+1} e^{S_n} + e^{S_{\tau_\gamma(x)}} \underbrace{\sum_{n=\tau_\gamma(x)+1}^{\infty} B_{n+1} e^{S_n - S_{\tau_\gamma(x)}}}_{\triangleq Z'},$$

where Z' is independent of $\sum_{n=0}^{\tau_\gamma(x)} B_{n+1}e^{S_n}$ and $e^{S_{\tau_\gamma(x)}}$, and has the same distribution as Z. A natural choice for approximating the distribution of Z' is a truncated sum. More precisely, letting $M \in \mathbb{N}$ be fixed, our modified estimator takes the form

$$L_T^\triangle(x, M) = \mathbb{1}_{\{\tau_\gamma(x)<\infty,\, \sum_{n=0}^{\tau_\gamma(x)+M} B_{n+1}e^{S_n}>x\}} \prod_{k=1}^{\tau_\gamma(x)} \frac{w_\gamma(S_{k-1}(\gamma) + a_*)}{v_\gamma(S_k(\gamma) + a_*)}. \tag{5.3.7}$$

We give a simulation algorithm for generating one sample of $L_T^\triangle(x, M)$.

Algorithm 5 Generating one sample of $L_T^{\Delta}(x, M)$

```
1: fixed δ ∈ (0,1)
2: a_* ← a_*(δ) ≤ 0 satisfying (5.3.5)
3: m ← 1, n ← 0, Z ← 0, S_n(γ) ← 0, L_T(x) ← 1
4: while n < τ_γ(x) do
5:     S_{n+1}(γ) ← S_n(γ) + X_{n+1}(γ)    ▷ X_{n+1}(γ) ∼ P(· | X_{n+1}(γ) + W > s(x) − S_n(γ) − a_*
6:     Sample (A_{n+1}, B_{n+1}) conditional on the value of X_{n+1}(γ)
7:     L_T(x) ← L_T(x) w_γ(S_n(γ) + a_*) v_γ(S_{n+1}(γ) + a_*)^{−1}
8:     Z ← Z + B_{n+1} ∏_{i=1}^{n} A_i
9:     n ← n + 1
10: while m ≤ M do
11:     Sample (A_{τ_γ(x)+m}, B_{τ_γ(x)+m}) under the original measure
12:     Z ← Z + B_{τ_γ(x)+m} ∏_{i=1}^{τ_γ(x)+m} A_i
13:     m ← m + 1
14: if Z > x then return L_T^Δ(x, M)
15: else return 0
```

Remark 5.5. In Step 6 of Algorithm 5, sampling (A_{n+1}, B_{n+1}) conditionally on $X_{n+1}(\gamma)$ under the change of measure is equivalent to sampling it conditionally on $X_{n+1}(\gamma)$ under the original measure. To see this, note that, for $n+1 \leq \tau_\gamma(x)$ and any measurable set $C \subseteq \mathbb{R}^2$

$$
\begin{aligned}
&\mathbf{E}^{Q^{\gamma}_{a_*}}[\mathbb{1}_{(A_{n+1},\, B_{n+1})\in C} \mid S_i(\gamma), i \leq n, X_{n+1}(\gamma)] \\
&= \mathbf{E}\left[\left(\prod_{i=1}^{n+1} \frac{w_\gamma(S_{i-1}(\gamma)+a_*)}{v_\gamma(S_i(\gamma)+a_*)}\right) \,\middle|\, S_i(\gamma), i \leq n, X_{n+1}(\gamma)\right]^{-1} \\
&\quad \cdot \mathbf{E}\left[\mathbb{1}_{(A_{n+1},\, B_{n+1})\in C}\left(\prod_{i=1}^{n+1} \frac{w_\gamma(S_{i-1}(\gamma)+a_*)}{v_\gamma(S_i(\gamma)+a_*)}\right) \,\middle|\, S_i(\gamma), i \leq n, X_{n+1}(\gamma)\right] \\
&= \left(\prod_{i=1}^{n+1} \frac{w_\gamma(S_{i-1}(\gamma)+a_*)}{v_\gamma(S_i(\gamma)+a_*)}\right)^{-1} \\
&\quad \cdot \left(\prod_{i=1}^{n+1} \frac{w_\gamma(S_{i-1}(\gamma)+a_*)}{v_\gamma(S_i(\gamma)+a_*)}\right)\mathbf{E}\left[\mathbb{1}_{(A_{n+1},\, B_{n+1})\in C} \,\middle|\, S_i(\gamma), i \leq n, X_{n+1}(\gamma)\right] \\
&= \mathbf{E}\left[\mathbb{1}_{(A_{n+1},\, B_{n+1})\in C} \,\middle|\, S_i(\gamma), i \leq n, X_{n+1}(\gamma)\right] \\
&= \mathbf{P}\left((A_{n+1},\, B_{n+1}) \in C \mid X_{n+1}(\gamma)\right).
\end{aligned}
$$

Next we will analyze the performance of our modified estimator. We show in Theorem 5.3.3 below that, under the following assumption,

$$\mathbf{E}^{Q^{\gamma}_{a_*}} L^{\Delta}_{T}(x, M)/\mathbf{P}(T(x) < \infty) \to 1$$

as $x \to \infty$, establishing that the relative bias of L^{Δ}_{T} vanishes.

Assumption A4. $B_1 \geq 0$ a.s.

Remark 5.6. Under Assumption A4, Assumption A1 c) is redundant.

Theorem 5.3.3. *Under Assumptions A1, A3 and A4, $L^{\Delta}_{T}(x, M)$ as in* (5.3.7) *is asymptotically unbiased as $x \to \infty$, i.e.,*

$$\lim_{x\to\infty} \mathbf{E}^{Q^{\gamma}_{a_*}} L^{\Delta}_{T}(x, M)/\mathbf{P}(T(x) < \infty) = 1,$$

uniformly in $M \in \mathbb{N}$.

We are now ready to apply the bias elimination technique in Result 5.2.3 to the estimators proposed in (5.3.7) as mentioned in the paragraph above Algorithm 5. By analyzing the asymptotic behavior of the relative bias as $M \to \infty$ for fixed x (see Lemma 5.5.2 in Section 5.5), we are able to apply the bias elimination technique and obtain an unbiased estimator for $\mathbf{P}(Z > x)$. We introduce the following assumption; the unbiased estimator is then given in Theorem 5.3.4 below.

Assumption A5. a) The Markov chain $\{Z_n\}_{n\in\mathbb{N}}$ given by (5.1.2) is irreducible and aperiodic.

b) There exists $q \geq 2$ such that $\mathbf{E}|\log A_1|^q + \mathbf{E}|\log B_1^+|^q < \infty$.

Remark 5.7. Assumption A5a) is satisfied, for example, if (A_1, B_1) has a Lebesgue density (see [26, Lemma 2.2.2]).

Theorem 5.3.4. *Let Assumptions A1 and A3–A5 hold. Let v_γ and w_γ be defined as in* (5.3.3). *For fixed $\delta \in (0, 1)$ and $\beta \in (0, 1)$, there exists an $a_* = a_*(\delta) \leq 0$ satisfying*

$$-\delta \leq \frac{v_\gamma^{\frac{2-\beta}{1-\beta}}(y) - w_\gamma^{\frac{2-\beta}{1-\beta}}(y)}{\mathbf{P}(X_1 > -y + s(x))w_\gamma^{\frac{1}{1-\beta}}(y)}, \quad \forall y \leq s(x) + a_*.$$

Moreover, it is possible to construct a random variable N independent of x, such that
$$\sum_{n=0}^{\infty} \frac{\mathbf{E}^{Q^{\gamma}_{a_*}}(L_T^{\Delta}(x, 2^{n-1}) - L_T(x))^2}{\mathbf{P}(Z > x)^2 \mathbf{P}(N \geq n)} < \infty,$$
and hence, the estimator $L_T^{RG}(x)$ defined by
$$L_T^{RG}(x) \triangleq \sum_{n=0}^{N} \frac{L_T^{\Delta}(x, 2^n) - L_T^{\Delta}(x, 2^{n-1})}{\mathbf{P}(N \geq n)}$$
with L_T^{Δ} as in (5.3.7) *is unbiased and strongly efficient.*

Remark 5.8. As we will see in the proof of Theorem 5.3.4, one possible choice is to sample N with $\mathbf{P}(N \leq n) = 1 - (1-p)^n$ for $n \geq 1$, where $p < 1 - 2^{-(q-1)}$ and q is as in Assumption A5 b). In general, the bias elimination scheme of [106] is not guaranteed to produce non-negative estimators, which might not be ideal in the context of estimating (rare-event) probabilities. However, in our case, $L_T^{\Delta}(x, M)$ increases w.r.t. M by Assumption A4, and hence, the resulting unbiased estimator $L_T^{RG}(x)$ is always non-negative.

5.3.2 Iterated random functions

We consider the Markov chain $\{Z_n\}_{n \geq 0}$ with $Z_{n+1} = \Psi_n(Z_n)$, where Ψ_n satisfies the following assumption. For similar settings of analyzing Markov chains that are generated by iterated random functions, see e.g. [66], [37], and [50].

Assumption B0. $\{\Psi_n\}_{n \in \mathbb{N}}$ is a sequence of i.i.d. random Lipschitz functions with
$$\mathrm{Lip}(\Psi_n) \triangleq \sup_{z_1 \neq z_2} \left| \frac{\Psi_n(z_1) - \Psi_n(z_2)}{z_1 - z_2} \right|. \tag{5.3.8}$$
Moreover, there exists a sequence of i.i.d. random vectors $\{(A_n, B_n, D_n)\}_{n \in \mathbb{N}}$ such that
$$A_n z + B_n - D_n \leq \Psi_n(z) \leq A_n z^+ + B_n^+ + D_n, \quad \text{for all } z \in \mathbb{R}. \tag{5.3.9}$$
In addition, we can sample Ψ_n from the conditional distribution given $(\log^+(B_n^+ + D_n) - \gamma_2) \vee \log A_n$, for γ_2 as in Lemma 5.3.2 below.

The goal of this section is to extend the results from Section 5.3.1 to the setting as described above. To achieve this, we will introduce a list of additional assumptions that are extensions of Assumptions A1–A5. To begin with, we consider an extension of Assumption A1.

Assumption B1. Assume that (5.3.9) holds and (A_1, B_1, D_1) satisfies the following conditions:

a) $A_1, D_1 > 0$ a.s., $\mathbf{E} \log A_1 > -\infty$, and $\mathbf{E} \log \mathrm{Lip}(\Psi_1) < 0$. Moreover, $\mathbf{E} \log^+ |B_1 + D_1| < \infty$ and $\mathbf{E} \log^+ |B_1 - D_1| < \infty$.

b) $\mathbf{E} \log^+(A_1 \vee B_1) < \infty$.

c) Let the following tail behaviors hold

$$\mathbf{P}(\max(A_1, B_1 + D_1) > x) \sim \mathbf{P}(\max(A_1, B_1) > x),$$

$$\mathbf{P}(\max(A_1, B_1 - D_1) > x) \sim \mathbf{P}(\max(A_1, B_1) > x),$$

and

$$\mathbf{P}(A_1 > x, B_1 - D_1 \leq -x) = o(\mathbf{P}(\max(A_1, B_1) > x)).$$

Define

$$\Psi_{1:n}(z) \triangleq \Psi_1 \circ \Psi_2 \circ \cdots \circ \Psi_n(z)$$

for each $z \in \mathbb{R}$, and $Z \triangleq \lim_{n\to\infty} \Psi_{1:n}(Z_0)$. Recall that (cf. [50, Theorem 3.1]), under Assumption B1, the unique stationary solution to (5.1.1) exists, is finite, has the same distribution as Z and has right-unbounded support. Moreover, the distribution of Z does not depend on the initial condition Z_0. Thus, w.l.o.g. we set $Z_0 = 0$. Note that Z can be bounded from above by a stochastic perpetuity $\bar{Z}$ that is given by

$$\bar{Z} \triangleq \sum_{n=0}^{\infty} \bar{B}_{n+1} e^{S_n},$$

where $\bar{B}_n \triangleq \max(B_n^+ + D_n, 1)$ and $S_n = \sum_{i=1}^n \log A_i$. Analogously to the previous section, we construct an upper bound for $\bar{Z}$ (and thus for Z) that can be written as a functional of the supremum of a suitable random walk $S_n(\gamma)$. We claim the following lemma.

Lemma 5.3.2. *Under Assumption B1, there exists a constant γ_2 such that*

$$\mathbf{E}\left[\max\left(\log^+\left(B_1^+ + D_1\right) - \gamma_2, \log A_1\right)\right] < 0.$$

Moreover, there exists a constant $\gamma_1 \in (0, -\mathbf{E}[\log A_i \vee (\log^+ B_1 - \gamma_2)])$ such that

$$Z \leq \exp\left\{\max_{n\geq 0} S_n(\gamma)\right\} \frac{e^{\gamma_2}}{1 - e^{-\gamma_1}} < \infty, \tag{5.3.10}$$

where $S_n(\gamma) = S_n(\gamma_1, \gamma_2) = \sum_{i=1}^n [\log A_i \vee (\log^+(B_i^+ + D_i) - \gamma_2) + \gamma_1]$.

Let $S_n(\gamma)$ be as in Lemma 5.3.2. Now from (5.3.10) we can define $s(x) \triangleq \log x - \gamma_2 + \log(1-e^{-\gamma_1})$ and $\tau_\gamma(x) \triangleq \inf\{n \geq 0 : S_n(\gamma) > s(x)\}$, such that (5.3.2) holds. Thanks to [50], under subexponentiality assumptions on the random variable $\log(A_1 \vee B_1)$, the tail asymptotics can be described using the integrated tail function of $\log(A_1 \vee B_1)$. However, the upper bound we derived in Lemma 5.3.2 yields a shorter proof for the asymptotic upper bound in [50, Theorem 3.1].

Assumption B2. The integrated tail of $\log(A_1 \vee B_1)$, denoted by $\bar{F}_I$, is subexponential.

Theorem 5.3.5. *If Assumptions B1 and B2 hold, then*

$$\limsup_{x\to\infty} \frac{\mathbf{P}(Z > x)}{\bar{F}_I(\log(x))} \leq -\frac{1}{\mathbf{E}\log A_1}.$$

For fixed $a_* \leq 0$, v_γ and w_γ as in (5.3.3), recall that $\mathbf{P}^{Q^\gamma_{a_*}}$ and $\mathbf{E}^{Q^\gamma_{a_*}}$ denote respectively the probability measure and the expectation w.r.t. the stochastic process $\{S_n(\gamma)\}_{n\in\mathbb{N}}$ having a one-step transition kernel $Q^\gamma_{a_*}$ as in (5.3.4). Given the asymptotic behavior of $\mathbf{P}(Z > x)$, we are able to show the strong efficiency (under $\mathbf{P}^{Q^\gamma_{a_*}}$) of our estimator in Theorem 5.3.6 below.

Assumption B3. The distribution of $\log(A_1 \vee B_1)$ belongs to the class S^*.

Theorem 5.3.6. *Let Assumptions B1 and B3 hold. Let v_γ and w_γ be as in* (5.3.3)*. For fixed $\delta \in (0,1)$, one can choose $a_* = a_*(\delta) \leq 0$ such that* (5.3.5) *holds. Let*

$$L_T(x) \triangleq \mathbb{1}_{\{T(x)<\infty\}} \prod_{k=1}^{\tau_\gamma(x)} \frac{w_\gamma(S_{k-1}(\gamma)+a_*)}{v_\gamma(S_k(\gamma)+a_*)}.$$

Then $L_T(x)$ is a strongly efficient estimator of $\mathbf{P}(Z > x)$ under $\mathbf{P}^{Q^\gamma_{a_*}}$.

Recall that $\Psi_{1:n}(z) \triangleq \Psi_1 \circ \Psi_2 \circ \cdots \circ \Psi_n(z)$ for each $z \in \mathbb{R}$. Define

$$L^\Delta_T(x, M) \triangleq \mathbb{1}_{\{\tau_\gamma(x)<\infty,\, \Psi_{1:\tau_\gamma(x)+M}(Z_0)>x\}} \prod_{k=1}^{\tau_\gamma(x)} \frac{w_\gamma(S_{k-1}(\gamma)+a_*)}{v_\gamma(S_k(\gamma)+a_*)}. \tag{5.3.11}$$

As in Section 5.3.1, we approximate $L_T(x)$ by $L^\Delta_T(x, M)$. Analogously to Theorem 5.3.3, we show that the estimator as in (5.3.11) is asymptotically unbiased in Theorem 5.3.7 below.

Assumption B4. For each z, $\Psi_{1:n}(z)$ is increasing in n.

Theorem 5.3.7. *Under Assumptions B1, B3, and B4,* $L_T^{\Delta}(x, M)$ *in* (5.3.11) *is asymptotically unbiased as* $x \to \infty$, *i.e.,*

$$\lim_{x\to\infty} \frac{\mathbf{E}^{Q_{a_*}} L_T^{\Delta}(x, M)}{\mathbf{P}(T(x) < \infty)} = 1,$$

uniformly in $M \in \mathbb{N}$.

Applying again Result 5.2.3, we construct an unbiased estimator for estimating $\mathbf{P}(Z > x)$ in Theorem 5.3.8 below. To do this, we need the following assumptions.

Assumption B5. a) The Markov chain $\{Z_n\}_{n\in\mathbb{N}}$ given by (5.1.1) is irreducible and aperiodic.

b) There exists $q \geq 2$ such that $\mathbf{E}|\log A_1|^q + \mathbf{E}|\log B_1^+|^q + \mathbf{E}|\log D_1|^q < \infty$.

Assumption B6. There exists $\underline{z}$ such that $\Psi_n([\underline{z}, \infty)) \subseteq [\underline{z}, \infty)$ and Ψ_n is bijective on $[\underline{z}, \infty)$ almost surely.

Remark 5.9. Assumptions B4 and B6 are satisfied if, for instance, the stochastic equation is given by

$$Z_{n+1} = \sqrt{A_{n+1} Z_n^2 + B_{n+1} Z_n + C_{n+1}}.$$

This corresponds to a second-order random polynomial equation, which is studied in [66]. Other examples are, for instance, $\Psi_n(z) = \max\{A_n z, B_n\}$ and $\Psi_n(z) = A_n \max\{z, B_n\} + C_n$.

Theorem 5.3.8. *Let Assumptions B1 and B3–B6 hold. Let* v_γ *and* w_γ *be as in* (5.3.3)*. For fixed* $\delta \in (0, 1)$ *and* $\beta \in (0, 1)$*, there exists an* $a_* = a_*(\delta) \leq 0$ *satisfying*

$$-\delta \leq \frac{v_\gamma^{\frac{2-\beta}{1-\beta}}(y) - w_\gamma^{\frac{2-\beta}{1-\beta}}(y)}{\mathbf{P}(X_1 > -y + s(x)) w_\gamma^{\frac{1}{1-\beta}}(y)}, \quad \forall y \leq s(x) + a_*.$$

Then it is possible to construct a random variable N *independent of* x*, such that*

$$\sum_{n=0}^{\infty} \frac{\mathbf{E}^{Q_{a_*}^{\gamma}} (L_T^{\Delta}(x, 2^{n-1}) - L_T(x))^2}{\mathbf{P}(Z > x)^2 \mathbf{P}(N \geq n)} < \infty,$$

and hence, the estimator $L_T^{RG}(x)$ *defined by*

$$L_T^{RG}(x) \triangleq \sum_{n=0}^{N} \frac{L_T^{\Delta}(x, 2^n) - L_T^{\Delta}(x, 2^{n-1})}{\mathbf{P}(N \geq n)}$$

with $L_T^{\Delta}(x, M)$ *as in* (5.3.11) *is unbiased and strongly efficient.*

Remark 5.10. As in Remark 5.8, N can be chosen such that $\mathbf{P}(N \leq n) = 1 - (1-p)^n$ for $n \geq 1$, where $p < 1 - 2^{-(q-1)}$ and q is as in Assumption B5 b).

5.4 Numerical results

Here we investigate our algorithm numerically based on a stochastic perpetuity with $B_n = 1$. We consider the increment $\log A_n \stackrel{\mathrm{D}}{=} \mathcal{W} - 3/2$ where $\mathcal{W}$ is a random variable with Weibull distribution:

$$\mathbf{P}(\mathcal{W} > t) = \exp\left(-2t^{1/2}\right).$$

For the algorithmic parameters, we choose $a_* = -10$, $\gamma = 0.5$. Moreover, we use a geometrically distributed random truncation index with parameter 0.5. Figure 5.1 shows the change of the estimated probability with respect to the various choices of M for 4 different values of $x = 10^8$, $x = 10^{16}$, $x = 10^{32}$, and $x = 10^{64}$ in each of the four plots. One can see that the estimated probability stabilizes as M grows, which confirms that our estimator is consistent as $M \to \infty$. Comparing the four plots, one can also tell that the initial bias for small M decreases as x increases, which is consistent with the conclusion of Theorem 5.3.3 (vanishing relative bias). Table 5.1 reports the estimated probabilities, their 95%-confidence intervals and the estimated coefficients of variation, that is, the estimated standard deviation divided by the sample mean (based on 200000 samples), for different values of x and M. In the last column, we present the results produced with the unbiased algorithm as introduced in Theorem 5.3.4. We can see that, on the one hand the ratio between the estimated probability and the standard deviation stays roughly constant over a range of x values and M values; on the other hand, the estimated probability using the fixed truncation method tends to converge to the estimated probability produced with the unbiased algorithm as M grows. These observations illustrate the strong efficiency (Theorems 5.3.2 and 5.3.4) of our estimators.

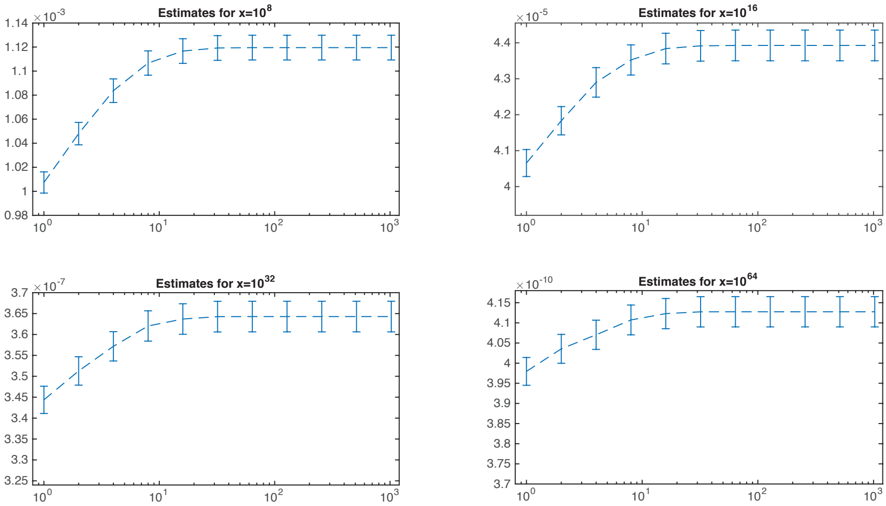

Figure 5.1: Estimated probabilities for different values of M. The y-axis values indicate the estimated rare-event probabilities and the vertical bars indicate the 95% confidence intervals. The x-axis values indicate the truncation index M.

Est CI CV	$M = 2^2$	$M = 2^4$	$M = 2^6$	$M = 2^8$	RG
	1.083×10^{-3}	1.117×10^{-3}	1.120×10^{-3}	1.120×10^{-3}	1.119×10^{-3}
$x = 10^8$	$\pm 0.009 \times 10^{-3}$	$\pm 0.010 \times 10^{-3}$	$\pm 0.010 \times 10^{-3}$	$\pm 0.010 \times 10^{-3}$	$\pm 0.013 \times 10^{-3}$
	2.06	2.10	2.10	2.10	2.70
	4.271×10^{-5}	4.373×10^{-5}	4.383×10^{-5}	4.383×10^{-5}	4.375×10^{-5}
$x = 10^{16}$	$\pm 0.041 \times 10^{-5}$	$\pm 0.042 \times 10^{-5}$	$\pm 0.043 \times 10^{-5}$	$\pm 0.043 \times 10^{-5}$	$\pm 0.053 \times 10^{-5}$
	2.17	2.22	2.22	2.22	2.76
	3.583×10^{-7}	3.646×10^{-7}	3.650×10^{-7}	3.650×10^{-7}	3.663×10^{-7}
$x = 10^{32}$	$\pm 0.035 \times 10^{-7}$	$\pm 0.037 \times 10^{-7}$	$\pm 0.037 \times 10^{-7}$	$\pm 0.037 \times 10^{-7}$	$\pm 0.045 \times 10^{-7}$
	2.25	2.28	2.29	2.29	2.81
	4.079×10^{-10}	4.120×10^{-10}	4.123×10^{-10}	4.123×10^{-10}	4.115×10^{-10}
$x = 10^{64}$	$\pm 0.037 \times 10^{-10}$	$\pm 0.037 \times 10^{-10}$	$\pm 0.038 \times 10^{-10}$	$\pm 0.038 \times 10^{-10}$	$\pm 0.041 \times 10^{-10}$
	2.05	2.06	2.06	2.06	2.27

Table 5.1: Estimated rare-event probability and 95% confidence intervals.

5.5 Proofs

In this section we provide proofs of the results presented in this chapter. Let $\tilde{v}(z) \triangleq \mathbf{P}(W > -z)$, $\tilde{w}(z) \triangleq \mathbf{P}(X_1 + W > -z)$, and

$$\tilde{Q}(y, dz) \triangleq P(y, dz)\tilde{v}(z)/\tilde{w}(y).$$

For $y \leq 0$, let $\mathbf{E}_y^{\tilde{Q}}$ denote the expectation operator associated with $\tilde{S}_n \triangleq y + S_n$ having the transition kernel $\tilde{Q}$, conditionally on $\tilde{S}_0 = y$. Let $\Gamma = \inf\{n \geq 0 \colon \tilde{S}_n > 0\}$.

Lemma 5.5.1. *Let $\epsilon > 0$ be given. Suppose that there exist constants $\delta_1, \delta_2 > 0$ and a finite-valued function $h : \mathbb{R} \longrightarrow [\delta_1, \infty)$ such that*

$$\tilde{w}^{1+\epsilon}(y) \int \tilde{v}(z)h(z)P(y, dz) \leq h(y)\tilde{v}^{2+\epsilon}(y), \tag{5.5.1}$$

for $y \leq 0$. If $h(z) \geq 1$ for $z > 0$ and $\tilde{v}(z) \geq \delta_2 > 0$ for $z > 0$, then we have that

$$\mathbf{E}_y^{Q'} \mathbb{1}_{\{\Gamma < \infty\}} \prod_{k=1}^{\Gamma} \frac{\tilde{w}^{2+\epsilon}(\tilde{S}_{k-1})}{\tilde{v}^{2+\epsilon}(\tilde{S}_k)} \leq \delta_1^{-1}\delta_2^{-(2+\epsilon)}\tilde{v}^{2+\epsilon}(y)h(y), \quad \textit{for } y \leq 0.$$

Proof of Lemma 5.5.1. Let $\mathbf{E}_y$ denote the expectation operator associated with $\{\tilde{S}_n\}_{n\geq 0}$ having the transition kernel P, conditionally on $\tilde{S}_0 = y$. Recall Theorem 2 (iii) of [16], where it is proved that if there exists a finite-valued non-negative function $\tilde{h}$ such that

$$(K\tilde{h})(y) \leq \tilde{h}(y) - \eta(y), \quad \text{for } y \leq 0,$$

where

$$(K\tilde{h})(y) = \int_{(-\infty,0]} \tilde{h}(z)\frac{\tilde{w}^{1+\epsilon}(y)}{\tilde{v}^{1+\epsilon}(z)}P(y, dz) \text{ and } \eta(y) = \int_{(0,\infty)} \frac{\tilde{w}^{1+\epsilon}(y)}{\tilde{v}^{1+\epsilon}(z)}P(y, dz),$$

then

$$\mathbf{E}_y \mathbb{1}_{\{\Gamma < \infty\}} \prod_{k=1}^{\Gamma} \frac{\tilde{w}^{1+\epsilon}(\tilde{S}_{k-1})}{\tilde{v}^{1+\epsilon}(\tilde{S}_k)} \leq \tilde{h}(y), \quad \text{for } y \leq 0. \tag{5.5.2}$$

Define $\tilde{h}(\cdot) = \delta_1^{-1}\delta_2^{-(2+\epsilon)}h(\cdot)v^{2+\epsilon}(\cdot)$. Note that

$$\delta_1^{-1}\delta_2^{-(2+\epsilon)}\tilde{w}^{1+\epsilon}(y) \int \tilde{v}(z)h(z)P'(y, dz)$$

$$= \delta_1^{-1}\delta_2^{-(2+\epsilon)}\tilde{w}^{1+\epsilon}(y)\left(\int_{(-\infty,0]} + \int_{(0,\infty)}\right)\tilde{v}(z)h(z)P(y,dz)$$
$$= (K\tilde{h})(y) + \delta_1^{-1}\delta_2^{-(2+\epsilon)}\tilde{w}^{1+\epsilon}(y)\int_{(0,\infty)}\tilde{v}(z)h(z)P(y,dz).$$

Thus, (5.5.1) is equivalent to

$$(K\tilde{h})(y) \le \tilde{h}(y) - \delta_1^{-1}\delta_2^{-(2+\epsilon)}\tilde{w}^{1+\epsilon}(y)\int_{(0,\infty)}\tilde{v}(z)h(z)P(y,dz). \tag{5.5.3}$$

On the other hand, we have that

$$\eta(y) = \int_{(0,\infty)}\frac{\tilde{w}^{1+\epsilon}(y)}{\tilde{v}^{1+\epsilon}(z)}P(y,dz)$$
$$\le \delta_2^{-(2+\epsilon)}\tilde{w}^{1+\epsilon}(y)\int_{(0,\infty)}\tilde{v}(z)P(y,dz)$$
$$\le \delta_1^{-1}\delta_2^{-(2+\epsilon)}\tilde{w}^{1+\epsilon}(y)\int_{(0,\infty)}h(z)\tilde{v}(z)P(y,dz). \tag{5.5.4}$$

Using (5.5.3) and (5.5.4), (5.5.1) implies that

$$(K\tilde{h})(y) \le \tilde{h}(y) - \eta(y).$$

From Theorem 2 (iii) of [16], we conclude that (5.5.2) holds, which is the desired result. □

Proof of Lemma 5.2.1. We first find h that satisfies (5.5.1) in Lemma 5.5.1. Define

$$h(y) = 1 - \delta\mathbb{1}(y > -a_*).$$

We will find a suitable $a_* \le 0$ later. Note that (5.5.1) in this case can be written as

$$\tilde{w}(y+a_*)^{-1}\mathbf{E}\tilde{v}(X_1+y+a_*)h(X+y) \le h(y)\left(\frac{\tilde{v}(y+a_*)}{\tilde{w}(y+a_*)}\right)^{2+\epsilon}$$
$$\le \left(\frac{\tilde{v}(y+a_*)}{\tilde{w}(y+a_*)}\right)^{2+\epsilon}, \tag{5.5.5}$$

for $y \le 0$. Here we use the fact that $h(y) = 1$ for $y \le 0$. By the definition of $\tilde{v}$,

$$\mathbf{E}\tilde{v}(X_1+y+a_*)h(X_1+y)$$

$$
\begin{aligned}
&= \mathbf{E}\big[\mathbf{E}[\mathbb{1}(W > -X_1 - y - a_*)|X_1]\big] \\
&\quad - \delta\mathbf{E}\big[\mathbf{E}[\mathbb{1}(W > -X_1 - y - a_*)|X_1]\mathbb{1}(X_1 + y > -a_*)\big] \\
&= \mathbf{P}(W + X_1 > -y - a_*) - \delta\mathbf{P}(W + X_1 > -y - a_*,\ X_1 + y > -a_*).
\end{aligned}
$$

Since $\tilde{w}(y + a_*) = \mathbf{P}(X_1 + W > -y - a_*)$, (5.5.5) is equivalent to

$$
1 - \delta\mathbf{P}(X_1 + y > -a_*|W + X_1 > -y - a_*) \le \frac{\tilde{v}^{2+\epsilon}(y + a_*)}{\tilde{w}^{2+\epsilon}(y + a_*)},
$$

which is equivalent to

$$
\begin{aligned}
&- \delta \le \frac{\tilde{v}^{2+\epsilon}(y + a_*) - \tilde{w}^{2+\epsilon}(y + a_*)}{\mathbf{P}(X_1 > -y - a_*)\tilde{w}^{1+\epsilon}(y + a_*)}, && \text{for all } y \le 0, \\
\Leftrightarrow &- \delta \le \frac{v^{2+\epsilon}(x + y + a_*) - w^{2+\epsilon}(x + y + a_*)}{\mathbf{P}(X_1 > -y - a_*)w^{1+\epsilon}(x + y + a_*)}, && \text{for all } y \le 0, \\
\Leftrightarrow &- \delta \le \frac{v^{2+\epsilon}(y) - w^{2+\epsilon}(y)}{\mathbf{P}(X_1 > -y + x)w^{1+\epsilon}(y)}, && \text{for all } y \le x + a_*. \quad (5.5.6)
\end{aligned}
$$

Using the definition of w and the non-negativity of W, (5.2.4) implies that $w(y) - v(y) = o(w(y))$, and hence $v(y) \sim w(y)$, as $y \to -\infty$. Therefore, there exists an a_* satisfying (5.5.6), and hence, (5.5.1). Since $\inf_{z \ge 0} \tilde{v}(z + a_*) = \mathbf{P}(W > -a_*)$, Lemma 5.5.1 applies to give

$$
\mathbf{E}\mathbb{1}_{\{\tilde{\tau}(0)<\infty\}} \prod_{k=1}^{\tilde{\tau}(0)} \frac{\tilde{w}^{1+\epsilon}(\tilde{S}_{k-1} + a_*)}{\tilde{v}^{1+\epsilon}(\tilde{S}_k + a_*)} \le \delta^{-1}\mathbf{P}(W > -a_*)^{-(2+\epsilon)}\tilde{v}^{2+\epsilon}(y), \text{ for } y \le 0.
$$

Recall the Pakes-Veraverbeke's theorem ([115], [118]):

$$
\mathbf{P}\left(\sup_{n \ge 0} \tilde{S}_n > 0\right) \sim -\frac{1}{\mathbb{E}X_1}\int_{-y}^{\infty} \mathbb{P}(X_1 > t)dt, \tag{5.5.7}
$$

as $y \to -\infty$. This implies that, for any fixed y

$$
\mathbf{P}\left(\sup_{n \ge 0} \tilde{S}_n > 0\right) \sim \tilde{v}(y), \quad \text{as } y \to -\infty.
$$

Combining this with the fact that $\mathbf{P}\left(\sup_{n\ge 0} \tilde{S}_n > 0\right)/\tilde{v}(y)$ is bounded as a function of y on compact sets, we obtain that

$$
\sup_{y<0} \mathbf{P}\left(\sup_{n \ge 0} \tilde{S}_n > 0\right)^{-2} \mathbf{E}\mathbb{1}_{\{\tilde{\tau}(0)<\infty\}} \prod_{k=1}^{\tilde{\tau}(0)} \frac{\tilde{w}^{1+\epsilon}(\tilde{S}_{k-1} + a_*)}{\tilde{v}^{1+\epsilon}(\tilde{S}_k + a_*)} < \infty,
$$

which is equivalent to

$$\sup_{x>0} \frac{\mathbf{E}^{Q_{a*}} L_\tau^{2+\epsilon}(x)}{\mathbf{P}(\sup_{n\geq 0} S_n > x)^{2+\epsilon}} < \infty.$$

□

Proof of Lemma 5.3.1. Note that $\max\{((\log^+ B_1 - \gamma_2') \vee \log A_1), 0\} \leq |\log^+ B_1 \vee \log A_1|$, and $\min\{((\log^+ B_1 - \gamma_2') \vee \log A_1), 0\}$ is bounded from above and non-increasing w.r.t. γ_2'. Since $(\log^+ B_1 - \gamma_2') \vee \log A_1 = \max\{((\log^+ B_1 - \gamma_2') \vee \log A_1), 0\} + \min\{((\log^+ B_1 - \gamma_2') \vee \log A_1), 0\}$, we can apply bounded convergence for the maximum and monotone convergence for the minimum to get

$$\lim_{\gamma_2' \to \infty} \mathbf{E}[(\log^+ B_1 - \gamma_2') \vee \log A_1] = \mathbf{E} \lim_{\gamma_2' \to \infty} (\log^+ B_1 - \gamma_2') \vee \log A_1 = \mathbf{E} \log A_1 < 0.$$

Therefore, there exists a γ_2 such that $\mathbf{E}[(\log^+ B_1 - \gamma_2') \vee \log A_1] < 0$.

Now we have that

$$Z \leq \sum_{n=0}^{\infty} \max(B_{n+1}, 1)\, e^{S_n} = e^{\gamma_2} \sum_{n=0}^{\infty} e^{(\log^+ B_{n+1} - \gamma_2) + S_n} \leq e^{\gamma_2} \sum_{n=0}^{\infty} e^{S_n'}, \quad (5.5.8)$$

where $S_n' = S_{n-1}' + (\log^+ B_n - \gamma_2) \vee \log A_n$. Note that the last inequality can be checked by comparing S_{n+1}' with $(\log^+ B_{n+1} - \gamma_2) + S_n$ for each n—i.e.,

$$\begin{aligned}
&\big(\log^+ B_{n+1} - \gamma_2\big) + S_n \\
&= \big(\log^+ B_{n+1} - \gamma_2\big) + \sum_{k=1}^{n} \log A_k \\
&\leq (\log^+ B_{n+1} - \gamma_2) \vee \log A_{n+1} + \sum_{k=1}^{n} (\log^+ B_k - \gamma_2) \vee \log A_k \\
&= S_{n+1}'.
\end{aligned}$$

Now let $\gamma_1 \in (0, -\mathbf{E}\left[(\log^+ B_1 - \gamma_2) \vee \log A_1\right])$ be fixed. From (5.5.8), we observe that

$$\begin{aligned}
Z &\leq e^{\gamma_2} \sum_{n=0}^{\infty} e^{S_n'} = e^{\gamma_2} \sum_{n=0}^{\infty} \exp\{S_n' + n\gamma_1\} \exp(-n\gamma_1) \\
&\leq \exp\left\{\max_{n\geq 0} S_n(\gamma)\right\} \frac{e^{\gamma_2}}{1 - e^{-\gamma_1}},
\end{aligned}$$

where $\gamma = (\gamma_1, \gamma_2)$, $S'_n = S'_{n-1} + (\log^+ B_n - \gamma_2) \vee \log A_n$ and $S_n(\gamma) = S'_n + n\gamma_1$. Note that $\mathbf{E}S_1(\gamma) < 0$ by the choice of γ_1. Hence $\sup_{n\geq 0} S_n(\gamma)$ is finite a.s. $\square$

Proof of Theorem 5.3.1. From the upper bound we constructed in Lemma 5.3.1, we know that

$$\mathbf{P}(Z > x) \leq \mathbf{P}\left(\sup_{n\geq 0} S_n(\gamma) > s(x)\right). \tag{5.5.9}$$

Due to Assumption A2, we know that the integrated tail of $\log^+(A_1 \vee B_1^+)$ is also subexponential. Moreover, it is straightforward to check that

$$\log^+(A_1 \vee B_1^+) - \gamma_2 \leq \log\left(\max\left\{A_1, e^{-\gamma_2} B_1^+, e^{-\gamma_2}\right\}\right) \leq \log^+(A_1 \vee B_1^+).$$

Therefore, the increments of the random walk $S_n(\gamma)$ have a subexponential integrated tail. Using the Pakes-Veraverbeke theorem we get the following relationship for the r.h.s. of (5.5.9), namely

$$\mathbf{P}\left(\sup_{n\geq 0} S_n(\gamma) > s(x)\right) \sim -\frac{1}{\mathbf{E}[(\log^+ B_1^+ - \gamma_2) \vee \log A_1] + \gamma_1} \bar{F}_I(\log(x)). \tag{5.5.10}$$

Thus, we have that

$$\limsup_{x\to\infty} \frac{\mathbf{P}(Z > x)}{\bar{F}_I(\log x)} \leq -\frac{1}{\mathbf{E}[(\log^+ B_1^+ - \gamma_2) \vee \log A_1] + \gamma_1}.$$

Now, letting $\gamma_2 \to \infty$ and $\gamma_1 \to 0$ we obtain the result. $\square$

Proof of Theorem 5.3.2. Let

$$M_n^{-1} = \prod_{k=1}^{n} \frac{w_\gamma(S_{k-1}(\gamma) + a_*)}{v_\gamma(S_k(\gamma) + a_*)}.$$

Obviously, $\{M_n\}_{n\in\mathbb{N}}$ is a martingale, and therefore, $\{M_{n\wedge\tau_\gamma(x)}\}_{n\in\mathbb{N}}$ is also a martingale. Since $\tau_\gamma(x) \leq T(x)$ we can apply Lemma 5.2.1 to get

$$\mathbf{E}^{Q^\gamma_{a_*}} L_T(x) = \mathbf{P}(T(x) < \infty) = \mathbf{P}(Z > x).$$

For the strong efficiency we have that

$$\frac{\mathbf{E}^{Q^\gamma_{a_*}} L_T^2(x)}{\mathbf{P}(Z > x)^2} = \frac{\mathbf{E}^{Q^\gamma_{a_*}} \mathbb{1}_{\{Z>x\}} M_{\tau_\gamma}^{-2}(x)}{\mathbf{P}(Z > x)^2}$$

$$\leq \frac{\mathbf{E}^{Q^{\gamma}_{a*}} \mathbb{1}_{\{\sup_{n\geq 0} S_n(\gamma) > s(x)\}} M_{\tau_\gamma}^{-2}(x)}{\mathbf{P}(Z > x)^2}$$

$$= \frac{\mathbf{E}^{Q^{\gamma}_{a*}} \mathbb{1}_{\{\sup_{n\geq 0} S_n(\gamma) > s(x)\}} M_{\tau_\gamma}^{-2}(x)}{\mathbf{P}\left(\sup_{n\geq 0} S_n(\gamma) > s(x)\right)^2} \left(\frac{\mathbf{P}\left(\sup_{n\geq 0} S_n(\gamma) > s(x)\right)}{\mathbf{P}\left(Z > x\right)} \right)^2, \quad (5.5.11)$$

where the first term in the last equation is guaranteed to be bounded over $x \in (1, \infty)$ due to Result 5.2.2. Hence, only the latter term remains to be analyzed. From [50, Theorem 3.1] we have that

$$\liminf_{x\to\infty} \frac{\mathbf{P}(Z > x)}{\bar{F}_I(\log(x))} \geq -\frac{1}{\mathbf{E}\log A_1}. \quad (5.5.12)$$

Since by assumption the integrated tail $\bar{F}_I$ is subexponential, it is in particular long-tailed. Combining (5.5.10) and (5.5.12) we obtain that

$$\limsup_{x\to\infty} \frac{\mathbf{P}\left(\sup_{n\geq 0} S_n(\gamma) > s(x)\right)}{\mathbf{P}(Z > x)} \leq \frac{\mathbf{E}\log A_1}{\mathbf{E}[\max(\bar{B}_1 - \gamma_2, \log A_1)] + \gamma_1}. \quad (5.5.13)$$

Using the fact that the l.h.s. of (5.5.13) is bounded over a compact interval, we obtain the result. $\square$

Proof of Theorem 5.3.3. Note that the Markov chain that has been considered in Section 5.3.1 is a special case of (5.1.1). Thus, for details we refer to the proof of Theorem 5.3.7 below. $\square$

Proof of Theorem 5.3.4. We wish to apply Result 5.2.3 to $L_T^{\Delta}(x, 2^n)$, $n \in \mathbb{N}$. Therefore, we should check the existence of a random variable N such that

$$\sum_{n=0}^{\infty} \frac{\mathbf{E}^{Q^{\gamma}_{a*}}[(L_T^{\Delta}(x, 2^{n-1}) - L_T(x))^2] - \mathbf{E}^{Q^{\gamma}_{a*}}[(L_T^{\Delta}(x, 2^n) - L_T(x))^2]}{\mathbf{P}(N \geq n)\mathbf{P}(Z > x)^2} < \infty. \quad (5.5.14)$$

We will bound

$$\frac{\mathbf{E}^{Q^{\gamma}_{a*}}(L_T^{\Delta}(x, 2^n) - L_T(x))^2}{\mathbf{P}(Z > x)^2} \quad (5.5.15)$$

by a geometrically decreasing function of n that does not depend on x. For $\beta \in (0, 1)$, using Hölder's inequality we get that

$$\frac{\mathbf{E}^{Q^{\gamma}_{a*}}(L_T^{\Delta}(x, 2^n) - L_T(x))^2}{P(Z > x)^2}$$

$$
\begin{aligned}
&= \frac{\mathbf{E}^{Q^{\gamma}_{a*}} \mathbb{1}_{\{\tau_\gamma(x)<\infty, \sum_{k=0}^{\tau_\gamma(x)+2^n} B_{k+1}e^{S_k}\le x, Z>x\}} (M^{-1}_{\tau_\gamma}(x))^2}{\mathbf{P}(Z>x)^2} \\
&= \mathbf{E}^{Q^{\gamma}_{a*}} \left(\mathbb{1}_{\{\tau_\gamma(x)<\infty, \sum_{k=0}^{\tau_\gamma(x)+2^n} B_{k+1}e^{S_k}\le x, Z>x\}} M^{-1}_{\tau_\gamma}(x)\right)^{\beta} \\
&\quad \times \left(\mathbb{1}_{\{T(x)<\infty\}} M^{-1}_{\tau_\gamma}(x)\right)^{2-\beta} \Big/ \mathbf{P}(Z>x)^2 \\
&\le \frac{\left(\mathbf{E}^{Q^{\gamma}_{a*}} \mathbb{1}_{\{\tau_\gamma(x)<\infty, \sum_{k=0}^{\tau_\gamma(x)+2^n} B_{k+1}e^{S_k}\le x, Z>x\}} M^{-1}_{\tau_\gamma}(x)\right)^{\beta}}{\mathbf{P}(Z>x)^{\beta}} \\
&\quad \times \frac{\left(\mathbf{E}^{Q^{\gamma}_{a*}} \mathbb{1}_{\{T(x)<\infty\}} M^{-1}_{\tau_\gamma}(x)^{\frac{2-\beta}{1-\beta}}\right)^{1-\beta}}{\mathbf{P}(Z>x)^{2-\beta}} \\
&= \underbrace{\left(\frac{\mathbf{E}^{Q^{\gamma}_{a*}} \mathbb{1}_{\{\tau_\gamma(x)<\infty, \sum_{k=0}^{\tau_\gamma(x)+2^n} B_{k+1}e^{S_k}\le x, Z>x\}} M^{-1}_{\tau_\gamma}(x)}{\mathbf{P}(Z>x)}\right)^{\beta}}_{\triangleq(\mathbf{II})} \\
&\quad \times \underbrace{\left(\frac{\mathbf{E}^{Q^{\gamma}_{a*}} L_T^{\frac{2-\beta}{1-\beta}}(x)}{\mathbf{P}(Z>x)^{\frac{2-\beta}{1-\beta}}}\right)^{1-\beta}}_{\triangleq(\mathbf{I})}.
\end{aligned}
$$

Now we analyze terms ($\mathbf{I}$) and ($\mathbf{II}$). Using the same argument around (5.5.11) in the proof of Theorem 5.3.2, to bound ($\mathbf{I}$), it is sufficient to analyze

$$
\mathbf{P}\left(\tau_\gamma(x)<\infty\right)^{-(2+\epsilon)} \mathbf{E}^{Q^{\gamma}_{a*}} L^{2+\epsilon}_{\tau_\gamma}(x).
$$

From Lemma 5.2.1, we see that $\mathbf{P}(\tau_\gamma(x)<\infty)^{-(2+\epsilon)}\mathbf{E}^{Q^{\gamma}_{a*}} L^{2+\epsilon}_{\tau_\gamma}(x)$ is bounded w.r.t. x; therefore, ($\mathbf{I}$) is also bounded w.r.t. x. Turning to ($\mathbf{II}$), we claim that it can be bounded by $\kappa 2^{-n(q-1)}$, for some constant $\kappa>0$. To see this, note that

$$
\begin{aligned}
(\mathbf{II}) &= \frac{\mathbf{P}\left(\tau_\gamma(x)<\infty, \sum_{k=0}^{\tau_\gamma(x)+2^n} B_{k+1}e^{S_k}\le x, Z>x\right)}{\mathbf{P}(Z>x)} \\
&= \underbrace{\mathbf{P}^{(x)}\left(\sum_{k=0}^{\tau_\gamma(x)+2^n} B_{k+1}e^{S_k}\le x, Z>x\right)}_{\triangleq(\mathbf{III})} \frac{\mathbf{P}(\tau_\gamma(x)<\infty)}{\mathbf{P}(Z>x)},
\end{aligned}
$$

where $\mathbf{P}^{(x)}(\cdot)$ denotes the conditional distribution $\mathbf{P}(\,\cdot\,|\,\tau_\gamma(x)<\infty)$. Hence, it is sufficient to analyze the behavior of **(III)** w.r.t. M. Note that

$$Z = \underbrace{\sum_{k=0}^{\tau_\gamma(x)} B_{k+1}e^{S_k}}_{\triangleq B'_x} + \underbrace{e^{S_{\tau_\gamma(x)}}}_{\triangleq A'_x} \underbrace{\sum_{k=1}^{\infty} B_{\tau_\gamma(x)+k} e^{S_{\tau_\gamma(x)+k}-S_{\tau_\gamma(x)}}}_{\triangleq Z'}, \tag{5.5.16}$$

and

$$\sum_{k=0}^{\tau_\gamma(x)+M} B_{k+1}e^{S_k} = B'_x + A'_x \underbrace{\sum_{k=1}^{M} B_{\tau_\gamma(x)+k} e^{S_{\tau_\gamma(x)+k}-S_{\tau_\gamma(x)}}}_{\triangleq Z'^{(M)}}. \tag{5.5.17}$$

Combining (5.5.16), (5.5.17) with Assumption A4, we obtain that

$$\begin{aligned}
\textbf{(III)} &= \int \mathbb{1}_{\left\{Z'^{(M)}\le \frac{x-B'_x}{A'_x},\, Z'>\frac{x-B'_x}{A'_x}\right\}} d\mathbf{P}^{(x)} \\
&= \int \mathbf{P}^{(x)}(Z'^{(M)} \le y, Z' > y) d\mathbf{P}^{(x)}((x-B'_x)/A'_x \le y) \\
&= \int \left\{\mathbf{P}^{(x)}(Z'>y) - \mathbf{P}^{(x)}(Z'^{(M)}>y)\right\} d\mathbf{P}^{(x)}((x-B'_x)/A'_x \le y).
\end{aligned}$$

Using the strong Markov property we have that $Z'^{(M)} \overset{\mathrm{D}}{=} \sum_{k=0}^{M} B_{k+1}e^{S_k}$ and $Z' \overset{\mathrm{D}}{=} Z$ under $\mathbf{P}^{(x)}$. Hence, we have that

$$\textbf{(III)} = \int \left\{\mathbf{P}^{(x)}(Z>y) - \mathbf{P}^{(x)}\left(\sum_{k=0}^{M} B_{k+1}e^{S_k} > y\right)\right\} d\mathbf{P}^{(x)}\left(\frac{x-B'_x}{A'_x} \le y\right).$$

Combining this with the fact that the backward iteration $\sum_{k=0}^{M} B_{k+1}e^{S_k}$ has the same distribution as Z_M defined in (5.1.2), we obtain that

$$\textbf{(III)} = \int \left\{\mathbf{P}\left(Z>y\right) - \mathbf{P}\left(Z_M>y\right)\right\} d\mathbf{P}^{(x)}((x-B'_x)/A'_x \le y) \le d_{TV}(Z_M, Z), \tag{5.5.18}$$

where d_{TV} denotes the total variation distance. To get a handle on this quantity, we apply the Lyapunov criterion in [77, Theorem 3.6], which implies a polynomial convergence rate of the M-step transition kernel to the invariant distribution in the total variation norm. In view of Lemma 5.5.2 below, there exists a constant κ such that **(III)** $\leq d_{TV}(Z_M, Z) \leq \kappa M^{-(q-1)}$, for all $M \in \mathbb{N}$. It should be noted that an exact expression for the constant κ can be obtained in a few special cases—see e.g. [44], [79] and the references therein. By choosing N such that $\mathbf{P}(N \leq n) = 1 - (1-p)^n$ for $n \geq 1$ with $p < 1 - 2^{-(q-1)}$, we conclude that the left-hand-side of (5.5.14) is bounded; and hence, by applying Result 5.2.3, we can get rid of this constant and obtain an unbiased, strongly efficient estimator. $\square$

Lemma 5.5.2. *Let Z_n be a Markov chain as in* (5.1.2) *such that Assumption A5 holds. Then there exists a constant κ such that $d_{TV}(Z_n, Z) \leq \kappa n^{-(q-1)}$.*

Proof of Lemma 5.5.2. W.l.o.g. we can assume $q \geq 2$ is an integer; otherwise one can take the greatest integer q' such that $q' < q$ and Assumption A5 holds for q'. We wish to apply Theorem 3.6 in [77]. In order to establish the Lyapunov condition as in (5.5.20) below, let $V(x) = 1 \vee (\log x)^q$ and $PV(x) \triangleq \mathbf{E}V(A_1x+B_1)$. Note that $V(x) = (\log x)^q \mathbb{1}_{\{x>e\}} + \mathbb{1}_{\{x\leq e\}}$ and hence the binomial expansion gives

$$
\begin{aligned}
&PV(x)\\
&= \mathbf{E}\left(\log\left(A_1x+B_1\right)\right)^q \mathbb{1}_{\{A_1x+B_1>e\}} + \mathbf{P}(A_1x+B_1 \leq e)\\
&\leq \mathbf{E}\left(\log\left(A_1x+B_1^+\right)\right)^q \mathbb{1}_{\{A_1x+B_1>e\}} + \mathbf{P}(A_1x+B_1 \leq e)\\
&= \mathbf{E}\left(\log\frac{A_1x+B_1^+}{x} + \log x\right)^q \mathbb{1}_{\{A_1x+B_1>e\}} + \mathbf{P}(A_1x+B_1 \leq e)\\
&= \mathbf{E}(\log x)^q \mathbb{1}_{\{A_1x+B_1>e\}}\\
&\quad + \sum_{i=1}^{q} \binom{q}{i} (\log x)^{q-i} \mathbf{E}\left(\log\frac{A_1x+B_1^+}{x}\right)^i \mathbb{1}_{\{A_1x+B_1>e\}}\\
&= \mathbf{E}[(\log x)^q](\mathbb{1}_{\{A_1x+B_1>e\}} + \mathbb{1}_{\{x>e\}} - \mathbb{1}_{\{x>e\}}) + \mathbb{1}_{\{x\leq e\}} - \mathbb{1}_{\{x\leq e\}}\\
&\quad + \sum_{i=1}^{q} \binom{q}{i} (\log x)^{q-i} \mathbf{E}\left(\log\frac{A_1x+B_1^+}{x}\right)^i \mathbb{1}_{\{A_1x+B_1>e\}}\\
&\quad + \mathbf{P}(A_1x+B_1 \leq e)\\
&= V(x) + \mathbf{E}(\log x)^q(\mathbb{1}_{\{A_1x+B_1>e\}} - \mathbb{1}_{\{x>e\}}) + \mathbf{P}(A_1x+B_1 \leq e) - \mathbb{1}_{\{x\leq e\}}\\
&\qquad + q(\log x)^{q-1}\mathbf{E}\left(\log\frac{A_1x+B_1^+}{x}\right)\mathbb{1}_{\{A_1x+B_1>e\}}\\
&\qquad + \sum_{i=2}^{q} \binom{q}{i} (\log x)^{q-i} \mathbf{E}\left[\left(\log\frac{A_1x+B_1^+}{x}\right)^i \mathbb{1}_{\{A_1x+B_1>e\}}\right].
\end{aligned}
$$

For $x > e$,

$$\begin{aligned}
PV(x) &\leq V(x) + \mathbf{P}(A_1 x + B_1 \leq e) \\
&\quad + q(\log x)^{q-1} \mathbf{E}\left(\log \frac{A_1 x + B_1^+}{x} \mathbb{1}_{\{A_1 x + B_1 > e\}}\right) \\
&\quad + \sum_{i=2}^{q} \binom{q}{i} (\log x)^{q-i}\, \mathbf{E}\left[\left(\log \frac{A_1 x + B_1^+}{x}\right)^i \mathbb{1}_{\{A_1 x + B_1 > e\}}\right].
\end{aligned}$$

Note that

$$\begin{aligned}
\log A_1 \leq \log \frac{A_1 x + B_1^+}{x} &\leq \log(A_1 + B_1^+) \\
&\leq \log(2(A_1 \vee B_1^+)) = \log(A_1 \vee B_1^+) + \log 2 \\
&= (\log A_1) \vee (\log B_1^+) + \log 2 \\
&\leq |\log A_1| + |\log B_1^+| + \log 2,
\end{aligned}$$

and hence

$$\left|\log \frac{A_1 x + B_1^+}{x}\right| \leq |\log A_1| + |\log B_1^+| + \log 2. \tag{5.5.19}$$

Moreover, the right-hand-side of (5.5.19) does not depend on x and has finite q-th moment. Thus, there have to be constants c_i, $i \geq 1$, such that

$$\begin{aligned}
&\sum_{i=2}^{q} \binom{q}{i} (\log x)^{q-i}\, \mathbf{E}\left[\left(\log \frac{A_1 x + B_1^+}{x}\right)^i \mathbb{1}_{\{A_1 x + B_1 > e\}}\right] \\
&\leq \sum_{i=0}^{q-2} c_i (\log x)^i \leq \epsilon (\log x)^{q-1}
\end{aligned}$$

for sufficiently large x. On the other hand, note that $\log \frac{A_1 x + B_1^+}{x} \mathbb{1}_{\{A_1 x + B_1 > e\}}$ converges to $\log A_1$ almost surely as $x \to \infty$, and hence, by dominated convergence

$$\mathbf{E}\left(\log \frac{A_1 x + B_1^+}{x}\right) \mathbb{1}_{\{A_1 x + B_1 > e\}} \to \mathbf{E} \log A_1 < 0.$$

Therefore, for any fixed $\epsilon > 0$,

$$q(\log x)^{q-1} \mathbf{E}\left(\log \frac{A_1 x + B_1^+}{x} \mathbb{1}_{\{A_1 x + B_1 > e\}}\right) \leq (q\mathbf{E} \log A_1 + \epsilon)(\log x)^{q-1}$$

for sufficiently large x. Choosing ϵ so that $q\mathbf{E}\log A_1 + 3\epsilon < 0$ and noting that $\mathbf{P}(A_1x + B_1 \leq e) \to 0$ as $x \to \infty$, as well as $(\log x)^{q-1} = \big((\log x)^q \mathbb{1}_{\{x>e\}} + \mathbb{1}_{\{x\leq e\}}\big)^{(q-1)/q}$ for $x > e$, we conclude that there exists K such that

$$\begin{aligned} PV(x) &\leq V(x) + \epsilon(\log x)^{q-1} + (q\mathbf{E}\log A_1 + \epsilon)(\log x)^{q-1} + \epsilon(\log x)^{q-1} \\ &\leq V(x) - cV^{(q-1)/q}(x) \end{aligned}$$

for $x > K$, where $c = -(q\mathbf{E}\log A_1 + 3\epsilon) > 0$. Finally, since $PV(x)$, $V(x)$ and $V^{(q-1)/q}(x)$ are bounded on $[0, K]$, there exists a constant b such that

$$PV(x) \leq V(x) - cV^{(q-1)/q}(x) + b\mathbb{1}_{[0,K]}, \tag{5.5.20}$$

which is the sufficient condition in [77, Theorem 3.6] for polynomial ergodicity. Thus we obtain the result. □

Proof of Lemma 5.3.2. Recall that Z can be bounded by a stochastic perpetuity $\bar{Z}$ that is given by

$$\bar{Z} \triangleq \sum_{n=0}^{\infty} \bar{B}_{n+1} e^{S_n},$$

where $\bar{B}_n \triangleq \max\left(B_n^+ + D_n, 1\right)$ and $S_n = \sum_{i=1}^{n} \log A_i$. Applying the same technique as in the proof of Lemma 5.3.1 to $\bar{Z}$, we obtain the result. □

Proof of Theorems 5.3.5 and 5.3.6. We omit the details here, since they can be proved analogously as Theorems 5.3.1 and 5.3.2. □

The following lemma is useful in proving Theorem 5.3.7.

Lemma 5.5.3. *Consider the sets $E_n^{(1)}$, $E_n^{(2)}$, and $E_n^{(3)}$ as in the proof of Theorem 5.3.7. Then, for $\nu, \epsilon > 0$, there exists $K > 0$, such that*

$$\mathbf{P}\left(\bigcap_{n\geq 1}\left(E_n^{(1)} \cap E_n^{(2)} \cap E_n^{(3)}\right)\right) \geq 1 - \epsilon.$$

Proof of Lemma 5.5.3. In the proof of Theorem 1 in [99], the authors state that for any $\nu > 0$ and any i.i.d. sequence $\{Y_n\}_{n\geq 0}$ with $\mathbf{E}\left[\log^+ |Y_1|\right] < \infty$, it holds that

$$\mathbf{P}\left(|Y_j| \leq e^{\nu j + K},\, j \leq n\right) \to 1,$$

as $K \to \infty$ uniformly w.r.t. n. Using this argument we conclude that $\mathbf{P}\left(E_n^{(3)}\right) \to 1$ as $K \to \infty$ uniformly w.r.t. n. Further, combining this fact with the SLLN for $\{S_n\}_{n\geq 0}$ and $\{S_n(\gamma)\}_{n\geq 0}$ (for details see e.g. [9, Lemma 3.1]), we can always take K large enough such that

$$\mathbf{P}\left(E_n^{(1)} \cap E_n^{(2)} \cap E_n^{(3)}\right) \geq 1 - \epsilon,$$

for all $n \in \mathbb{N}$. Finally, since the sequence of sets $E_n^{(1)} \cap E_n^{(2)} \cap E_n^{(3)}, n \geq 0$ is decreasing in the sense of inclusion, we obtain the result. □

Proof of Theorem 5.3.7. Recall that $\Psi_{1:n}(Z_0) \triangleq \Psi_1 \circ \Psi_2 \circ \cdots \circ \Psi_n(Z_0)$. Due to the fact that $\{\tau_\gamma(x) < \infty, \Psi_{1:\tau_\gamma(x)+M}(Z_0) > x\} \subseteq \{T(x) < \infty\}$, in order to prove the vanishing relative bias result, it is sufficient to show that

$$\liminf_{x\to\infty} \frac{\mathbf{P}\left(\tau_\gamma(x) < \infty, \Psi_{1:\tau_\gamma(x)+M}(Z_0) > x\right)}{\mathbf{P}\left(T(x) < \infty\right)} \geq 1. \qquad (5.5.21)$$

Recall that $S_n = \sum_{i=1}^n \log A_i$ and $S_n(\gamma) = n\gamma_1 + \sum_{i=1}^n [(\log^+(B_i^+ + D_i) - \gamma_2) \vee \log A_i]$. Let $\mu \triangleq -\mathbf{E}S_1$ and $\mu_\gamma \triangleq -\mathbf{E}S_1(\gamma)$. For $\nu, K > 0$ consider the sets

$$E_n^{(1)} = E_n^{(1)}(K,\nu) = \{S_j \in (-j(\mu+\nu) - K, -j(\mu-\nu) + K), \, j \leq n\},$$

$$E_n^{(2)} = E_n^{(2)}(K,\nu) = \{S_j(\gamma) \in (-j(\mu_\gamma+\nu) - K, -j(\mu_\gamma-\nu) + K), \, j \leq n\},$$

and

$$E_n^{(3)} = E_n^{(3)}(K,\nu) = \left\{|\underline{B}_j| \leq e^{\nu j + K}, \, j \leq n\right\},$$

where $\underline{B}_j = B_j - D_j$. Define

$$\begin{aligned} E_n = {} & E_n^{(1)} \cap E_n^{(2)} \cap E_n^{(3)} \cap \{\Psi_{n+2}(Z_0) > \nu\} \\ & \cap \left\{\max\left(A_{n+1}, \underline{B}_{n+1}\right) > xe^{n(\mu+\nu)+L+K}, \underline{B}_{n+1} \geq -xe^{n(\mu-\nu)-K}\right\}, \end{aligned}$$

where $L > 0$ is chosen to be large enough, such that the sets $\{E_n\}_{n\geq 1}$ are disjoint. The existence of such an L is guaranteed by the fact that $E_n \subseteq \{\tau_\gamma(x) = n+1\}$ (see below). Now, we show that $E_n \subseteq \{\tau_\gamma(x) = n+1, \Psi_{1:\tau_\gamma(x)+1}(Z_0) > x\}$. To see $E_n \subseteq \{\tau_\gamma(x) = n+1\}$, note that $\{S_j(\gamma)\}_{j\leq n}$ is bounded by K, $\mu > \mu_\gamma$—due to the fact that $S_1 \leq S_1(\gamma)$—and

$$S_{n+1}(\gamma) = S_n(\gamma) + \log\left(\max\left(\bar{B}_{n+1}e^{-\gamma_2}, A_{n+1}\right)\right) + \gamma_1$$

$$
\begin{aligned}
&> -n\left(\mu_\gamma + \nu\right) - K + \log\left(\max\left(\underline{B}_{n+1}, A_{n+1}\right)\right) - \gamma_2 + \gamma_1 \\
&> \log x + n\left(\mu - \mu_\gamma\right) + L - \gamma_2 + \gamma_1 \\
&> \log x + L - \gamma_2 + \gamma_1 > s(x),
\end{aligned}
$$

for sufficiently large L that does not depend on x. Thus, we conclude that $\tau_\gamma(x) = n+1 < \infty$ by taking x sufficiently large. To see $E_n \subseteq \{\tau_\gamma(x) = n+1, \Psi_{1:\tau_\gamma(x)+1}(Z_0) > x\}$, note that $\Psi_{1:n}(z) \geq \sum_{k=0}^{n-1} \underline{\mathrm{B}}_{k+1} \prod_{j=1}^{k} A_j + z \prod_{j=1}^{n} A_j$ from (5.3.9) and Assumption B4. Moreover, $|\underline{B}_{k+1}| \prod_{j=1}^{k} A_j = |\underline{B}_{k+1}| e^{S_k} \leq e^{\nu(k+1)+K} e^{-k(\mu-\nu)+K} = e^{-k(\mu-2\nu)+2K+\nu}$ on E_n, and hence,

$$
\begin{aligned}
&\Psi_{1:\tau_\gamma(x)+1}(Z_0) = \Psi_{1:n+2}(Z_0) = \Psi_{1:n}(\Psi_{n+1:n+2}(Z_0)) \\
&\geq \sum_{k=0}^{n-1} \underline{B}_{k+1} \prod_{j=1}^{k} A_j + \left(\underline{B}_{n+1} + \Psi_{n+2}(Z_0) A_{n+1}\right) \prod_{j=1}^{n} A_j \\
&\geq -\sum_{k=0}^{n-1} |\underline{B}_{k+1}| \prod_{j=1}^{k} A_j + \left(\underline{B}_{n+1} + x e^{n(\mu-\nu)-K} + \Psi_{n+2}(Z_0) A_{n+1}\right) \prod_{j=1}^{n} A_j \\
&\quad - x e^{n(\mu-\nu)-K} \prod_{j=1}^{n} A_j \\
&\geq -\frac{e^{2K+\nu}}{1-e^{-\mu+2\nu}} + \left(\underline{B}_{n+1} + x e^{n(\mu-\nu)-K} + \nu A_{n+1}\right) \prod_{j=1}^{n} A_j - x \\
&\geq -\frac{e^{2K+\nu}}{1-e^{-\mu+2\nu}} + \min(\nu,1) \max\left(A_{n+1}, \underline{B}_{n+1} + x e^{n(\mu-\nu)-K}\right) e^{-n(\mu+\nu)-K} - x \\
&\geq -\frac{e^{2K+\nu}}{1-e^{-\mu+2\nu}} + \min(\nu,1) \max\left(A_{n+1}, \underline{B}_{n+1}\right) e^{-n(\mu+\nu)-K} - x \\
&\geq -\frac{e^{2K+\nu}}{1-e^{-\mu+2\nu}} + \min(\nu,1) x e^{L} - x > x,
\end{aligned}
$$

for sufficiently large L that does not depend on x. Note that from Lemma 5.5.3 above,

$$
\begin{aligned}
\mathbf{P}(E_n) &= \mathbf{P}(E_n^{(1)} \cap E_n^{(2)} \cap E_n^{(3)}) \cdot \mathbf{P}(\Psi_{n+2}(Z_0) > \nu) \\
&\qquad \cdot \mathbf{P}\Big(\max\left(A_{n+1}, \underline{B}_{n+1}\right) > x e^{n(\mu+\nu)+L+K}, \underline{B}_{n+1} \geq -x e^{n(\mu-\nu)-K}\Big) \\
&= \mathbf{P}(E_n^{(1)} \cap E_n^{(2)} \cap E_n^{(3)}) \cdot \mathbf{P}(\Psi_1(Z_0) > \nu) \\
&\qquad \cdot \mathbf{P}\Big(\max\left(A_1, \underline{B}_1\right) > x e^{n(\mu+\nu)+L+K}, \underline{B}_1 \geq -x e^{n(\mu-\nu)-K}\Big)
\end{aligned}
$$

$$
\begin{aligned}
&\geq (1-\epsilon)\,\mathbf{P}(\Psi_1(Z_0) > \nu) \\
&\qquad \cdot \Big\{ \mathbf{P}\big(\max(A_1, \underline{B}_1) > x e^{n(\mu+\nu)+L+K} \big) \\
&\qquad\qquad - \mathbf{P}\big(\max(A_1, \underline{B}_1) > x e^{n(\mu+\nu)+L+K}, \underline{B}_1 < -x e^{n(\mu-\nu)-K} \big) \Big\} \\
&\geq (1-\epsilon)\,\mathbf{P}(\Psi_1(Z_0) > \nu) \\
&\qquad \cdot \Big\{ \mathbf{P}\big(\max(A_1, \underline{B}_1) > x e^{n(\mu+\nu)+L+K} \big) \\
&\qquad\qquad - \mathbf{P}\big(A_1 > x e^{n(\mu+\nu)+L+K}, \underline{B}_1 < -x e^{n(\mu-\nu)-K} \big) \Big\}.
\end{aligned}
$$

Moreover, since $E_n \subseteq \{\tau_\gamma(x) < \infty, \Psi_{1:\tau_\gamma(x)+M}(Z_0) > x\}$, and E_n, $n \geq 1$ are disjoint, we have that

$$
\begin{aligned}
&\mathbf{P}\left(\tau_\gamma(x) < \infty, \Psi_{1:\tau_\gamma(x)+M}(Z_0) > x\right) \geq \sum_{n\geq 0} \mathbf{P}(E_n) \\
&\geq (1-\epsilon)\mathbf{P}(\Psi_1(Z_0) > \nu) \sum_{n\geq 0} \Big\{ \mathbf{P}(\max(A_1, \underline{B}_1) > x e^{n(\mu+\nu)+L+K}) \\
&\qquad\qquad - \mathbf{P}(A_1 > x e^{n(\mu+\nu)+L+K}, \underline{B}_1 < -x e^{n(\mu-\nu)-K}) \Big\}.
\end{aligned}
\tag{5.5.22}
$$

From Assumption B1c) we conclude that, for any $\epsilon' > 0$, by taking sufficiently large x,

$$
\begin{aligned}
&\mathbf{P}\left(A_1 > x e^{n(\mu+\nu)+L+K}, \underline{B}_1 < -x e^{n(\mu-\nu)-K} \right) \\
&\leq \mathbf{P}\left(A_1 > x e^{n(\mu-\nu)-K}, \underline{B}_1 < -x e^{n(\mu-\nu)-K} \right) \\
&\leq \epsilon' \mathbf{P}\left(\max(A_1, \underline{B}_1) > x e^{n(\mu-\nu)-K} \right).
\end{aligned}
$$

Combining this with (5.5.22), we obtain that

$$
\begin{aligned}
&\mathbf{P}\left(\tau_\gamma(x) < \infty, \Psi_{1:\tau_\gamma(x)+M}(Z_0) > x\right) \\
&\geq (1-\epsilon)\mathbf{P}(\Psi_1(Z_0) > \nu) \\
&\quad \sum_{n\geq 0} \Big\{ \mathbf{P}\left(\max(A_1, \underline{B}_1) > x e^{n(\mu+\nu)+L+K} \right) \\
&\qquad\qquad - \epsilon' \mathbf{P}\left(\max(A_1, \underline{B}_1) > x e^{n(\mu-\nu)-K} \right) \Big\}.
\end{aligned}
\tag{5.5.23}
$$

For a given $\epsilon'' > 0$, let y be sufficiently large so that

$$\left|1 - \frac{\mathbf{P}(\log\max(A_1, \underline{B}_1) > y)}{\mathbf{P}(\log\max(A_1, B_1) > y)}\right| \leq \epsilon''.$$

Since $\mathbf{P}\left(\max\left(A_1, \underline{B}_1\right) > y\right)$ is decreasing in y,

$$\begin{aligned}
&\sum_{n\geq 0} \mathbf{P}\left(\max\left(A_1, \underline{B}_1\right) > xe^{n(\mu+\nu)+L+K}\right) \\
&\geq \sum_{n\geq 0} \frac{1}{\mu+\nu} \int_{\log x+L+K+n(\mu+\nu)}^{\log x+L+K+(n+1)(\mu+\nu)} \mathbf{P}\left(\log\max\left(A_1, \underline{B}_1\right) > y\right) dy \\
&\geq \sum_{n\geq 0} \frac{1-\epsilon''}{\mu+\nu} \int_{\log x+L+K+n(\mu+\nu)}^{\log x+L+K+(n+1)(\mu+\nu)} \mathbf{P}\left(\log\max\left(A_1, B_1\right) > y\right) dy \\
&= \frac{1-\epsilon''}{\mu+\nu} \bar{F}_I(\log x + L + K),
\end{aligned}$$

and that

$$\begin{aligned}
&\sum_{n\geq 0} \mathbf{P}\left(\max\left(A_1, \underline{B}_1\right) > xe^{n(\mu-\nu)-K}\right) \\
&\leq \sum_{n\geq 0} \frac{1+\epsilon''}{\mu-\nu} \int_{\log x-K+(n-1)(\mu-\nu)}^{\log x-K+n(\mu-\nu)} \mathbf{P}\left(\log\max\left(A_1, \underline{B}_1\right) > y\right) dy \\
&\leq \sum_{n\geq 0} \frac{1+\epsilon''}{\mu-\nu} \int_{\log x-K+(n-1)(\mu-\nu)}^{\log x-K+n(\mu-\nu)} \mathbf{P}\left(\log\max\left(A_1, \underline{B}_1\right) > y\right) dy \\
&= \frac{1+\epsilon''}{\mu-\nu} \bar{F}_I(\log x - K - \mu + \nu).
\end{aligned}$$

Moreover, using the fact that $\bar{F}_I$ is long-tailed, we obtain from (5.5.23) that

$$\begin{aligned}
&\mathbf{P}\left(\tau_\gamma(x) < \infty, \Psi_{1:\tau_\gamma(x)}(\Psi_{\tau_\gamma(x)+1:\tau_\gamma(x)+M}(Z_0)) > x\right) \\
&\geq (1-\epsilon)\mathbf{P}(\Psi_1(Z_0) > \nu)\Bigg(\frac{1-\epsilon''}{\mu+\nu}\bar{F}_I(\log x + L + K) \\
&\qquad\qquad - \frac{\epsilon'(1+\epsilon'')}{\mu-\nu}\bar{F}_I(\log x + L + K - \mu + \nu)\Bigg) \\
&\sim (1-\epsilon)\mathbf{P}(\Psi_1(Z_0) > \nu)\left(\frac{1-\epsilon''}{\mu+\nu} - \frac{\epsilon'(1+\epsilon'')}{\mu-\nu}\right)\bar{F}_I(\log x)
\end{aligned}$$

$$\sim \mu(1-\epsilon)\mathbf{P}(\Psi_1(Z_0) > \nu)\left(\frac{1-\epsilon''}{\mu+\nu} - \frac{\epsilon'(1+\epsilon'')}{\mu-\nu}\right)\mathbf{P}(T(x) < \infty), \tag{5.5.24}$$

where in (5.5.24) we use [50, Theorem 3.1]. Recall that the distribution of the stationary solution to (5.1.1) does not depend on the initial condition Z_0, and hence, w.l.o.g. we can set $Z_0 = 0$. Noting that $\Psi_1(0) \geq 0$ and hence $\mathbf{P}(\Psi_1(Z_0) > \nu) \to 1$ as $\nu \to 0$, we let $\epsilon, \epsilon', \epsilon'', \nu \to 0$ to obtain (5.5.21). This implies that the relative bias converges to 0, since the numerator in (5.5.21) is always smaller than the denominator. $\square$

The following lemma is useful in proving Theorem 5.3.8.

Lemma 5.5.4. *Let Z_n be a Markov chain as in* (5.1.1) *such that Assumption B5 hold. Then there exists a constant κ such that $d_{TV}(Z_n, Z) \leq \kappa n^{-(q-1)}$.*

Proof of Lemma 5.5.4. Let $V(x) = 1 \vee (\log x)^q$ and $PV(x) \triangleq \mathbf{E}V(\Psi_1(x))$. By noting that

$$\begin{aligned} PV(x) &\triangleq \mathbf{E}(\log(\Psi_1(x)))^q \mathbb{1}_{\{\Psi_1(x)>e\}} + \mathbf{P}(\Psi_1(x) \leq e) \\ &\leq \mathbf{E}(\log(A_1 x + B_1^+ + D_1))^q \mathbb{1}_{\{\Psi_1(x)>e\}} + \mathbf{P}(\Psi_1(x) \leq e), \end{aligned}$$

the rest follows immediately from similar arguments as in the proof of Lemma 5.5.2. $\square$

Proof of Theorem 5.3.8. Note that this result can be proved by following similar arguments as in the proof of Theorem 5.3.4, hence we only give a sketch of the proof. Recall that, in the context of iterated random functions, the estimator $L_T^{\Delta}(x, M)$ is given by

$$L_T^{\Delta}(x, M) \triangleq \mathbb{1}_{\{\tau_\gamma(x)<\infty,\, \Psi_{1:\tau_\gamma(x)+M}(Z_0)>x\}} \prod_{k=1}^{\tau_\gamma(x)} \frac{w_\gamma(S_{k-1}(\gamma) + a_*)}{v_\gamma(S_k(\gamma) + a_*)}.$$

Analogously to the proof of Theorem 5.3.4, we wish to bound

$$\frac{\mathbf{E}^{Q_{a_*}^\gamma}(L_T^{\Delta}(x, 2^n) - L_T(x))^2}{\mathbf{P}(Z > x)^2} \tag{5.5.25}$$

by a decreasing function of n independent of x. Again, by using Hölder's inequality it is sufficient to bound

$$\textbf{(I')} \triangleq \frac{\mathbf{E}^{Q_{a_*}^\gamma} \mathbb{1}_{\{\tau_\gamma(x)<\infty, \Psi_{1:\tau_\gamma(x)+M}(Z_0)\leq x, Z>x\}} M_{\tau_\gamma}^{-1}(x)}{\mathbf{P}(Z > x)}$$

$$= \underbrace{\mathbf{P}^{(x)}\left(\Psi_{1:\tau_\gamma(x)}(\Psi_{\tau_\gamma(x)+1:\tau_\gamma(x)+M}(Z_0)) \leq x, \Psi_{1:\tau_\gamma(x)}(Z') > x\right)}_{\triangleq\textbf{(II')}} \frac{\mathbf{P}(\tau_\gamma(x) < \infty)}{\mathbf{P}(Z > x)},$$

where $Z' \triangleq \lim_{M\to\infty} \Psi_{\tau_\gamma(x)+1:\tau_\gamma(x)+M}(Z_0) \stackrel{\mathrm{D}}{=} Z$ and $\mathbf{P}^{(x)}(\cdot)$ denotes the conditional distribution $\mathbf{P}(\,\cdot\,|\,\tau_\gamma(x) < \infty)$. Since Ψ_n is Lipschitz and bijective, $\Psi^{-1}_{1:\tau_\gamma(x)}$ is either strictly increasing or strictly decreasing. W.l.o.g. we assume that $\Psi^{-1}_{1:\tau_\gamma(x)}$ is strictly increasing, since the case of $\Psi^{-1}_{1:\tau_\gamma(x)}$ being strictly decreasing can be dealt with similarly. Using the strong Markov property we obtain that

$$\textbf{(II')} = \int \{\mathbf{P}(Z > y) - \mathbf{P}(Z_M > y)\}\, d\mathbf{P}^{(x)}(\Psi^{-1}_{1:\tau_\gamma(x)}(x) \leq y) \leq d_{TV}(Z_M, Z).$$

By Lemma 5.5.4 above we obtain the result. □

Bibliography

[1] H. Albrecher, J. Beirlant, and J. Teugels. *Reinsurance: Actuarial and statistical aspects*. Wiley Series in Probability and Statistics. John Wiley & Sons, 2017.

[2] H. Albrecher, C. Robert, and J. Teugels. Joint asymptotic distributions of smallest and largest insurance claims. *Risks*, 2(3):289–314, 2014.

[3] H. Ammeter. The rating of "Largest Claim" reinsurance covers. *Quarterly letter from the Algemeene Reinsurance Companies Jubilee*, (2):5–17, 1964.

[4] S. Asmussen and H. Albrecher. *Ruin probabilities*. Advanced Series on Statistical Science & Applied Probability, 14. World Scientific, Second edition, 2010.

[5] S. Asmussen and P. Glynn. *Stochastic simulation: Algorithms and analysis*, volume 57 of *Stochastic Modelling and Applied Probability*. Springer-Verlag New York, 1st edition, 2007.

[6] S. Asmussen and C. Klüppelberg. Large deviations results for subexponential tails, with applications to insurance risk. *Stochastic Processes and their Applications*, 64(1):103–125, 1996.

[7] S. Asmussen and H. M. Nielsen. Ruin probabilities via local adjustment coefficients. *Journal of Applied Probability*, 32(3):736–755, 1995.

[8] S. Asmussen and M. Pihlsgård. Performance analysis with truncated heavy-tailed distributions. *Methodology and Computing in Applied Probability*, 7(4):439–457, Dec 2005.

[9] S. Asmussen, H. Schmidli, and V. Schmidt. Tail probabilities for non-standard risk and queueing processes with subexponential jumps. *Advances in Applied Probability*, 31(2):422–447, 1999.

[10] S. Aspandiiarov and R. Iasnogorodski. General criteria of integrability of functions of passage-times for nonnegative stochastic processes and their applications. *Theory of Probability & Its Applications*, 43(3):343–369, 1999.

[11] K. Athreya and P. Ney. *Branching processes.* Grundlehren der mathematischen Wissenschaften. Springer-Verlag Berlin Heidelberg, 1972.

[12] K. B. Athreya and P. Ney. A new approach to the limit theory of recurrent Markov chains. *Transactions of the American Mathematical Society*, 245:493–501, 1978.

[13] B. Berliner. Correlations between excess of loss reinsurance covers and reinsurance of the n largest claims. *ASTIN Bulletin: The Journal of the IAA*, 6(3):260–275, 1972.

[14] P. Billingsley. *Convergence of probability measures.* John Wiley & Sons, 2013.

[15] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1987.

[16] J. Blanchet and P. Glynn. Efficient rare-event simulation for the maximum of heavy-tailed random walks. *The Annals of Applied Probability*, 18(4):1351–1378, August 2008.

[17] J. Blanchet, P. Glynn, and J. Liu. Efficient rare event simulation for heavy-tailed multiserver queues. Technical report, 2007.

[18] J. Blanchet, P. Glynn, and S. Meyn. Large deviations for the empirical mean of an M/M/1 queue. *Queueing Systems*, 73(4):425–446, Apr 2013.

[19] J. Blanchet, H. Lam, Q. Tang, and Z. Yuan. Applied robust performance analysis for actuarial applications. Technical report, 01 2017.

[20] J. Blanchet, H. Lam, and B. Zwart. Efficient rare-event simulation for perpetuities. *Stochastic Processes and their Applications*, 122:3361 – 3392, 2012.

[21] J. Blanchet and J. Liu. State-dependent importance sampling for regularly varying random walks. *Advances in Applied Probability*, 40(4):1104–1128, Dezember 2008.

[22] J. Blanchet and J. Liu. Efficient importance sampling in ruin problems for multidimensional regularly varying random walks. *Journal of Applied Probability*, 47(2):301–322, June 2010.

[23] J. Blanchet and B. Zwart. Importance sampling of compounding processes. In *2007 Winter Simulation Conference*, pages 372–379, December 2007.

[24] A. A. Borovkov and K. A. Borovkov. *Asymptotic analysis of random walks: Heavy-tailed distributions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2008.

[25] O. J. Boxma, E. J. Cahen, D. Koops, and M. Mandjes. Linear stochastic fluid networks: Rare-event simulation and Markov modulation. *Methodology and Computing in Applied Probability*, 21(1):125–153, Mar 2019.

[26] D. Buraczewski, E. Damek, and T. Mikosch. *Stochastic models with power-law tails: The equation X=AX+B*. Springer Series in Operations Research and Financial Engineering. Springer International Publishing, 1st edition, 2016.

[27] D. Buraczewski, E. Damek, T. Mikosch, and J. Zienkiewicz. Large deviations for solutions to stochastic recurrence equations under Kesten's condition. *The Annals of Probability*, 41(4):2755–2790, 2013.

[28] A. Castaño-Martìnez, G. Pigueiras, and M. Sordo. On a family of risk measures based on largest claims. *Insurance: Mathematics and Economics*, 2019.

[29] B. Chen, J. Blanchet, C.-H. Rhee, and B. Zwart. Efficient rare-event simulation for multiple jump events in regularly varying random walks and compound Poisson processes. *Mathematics of Operations Research*, 2019.

[30] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *The Annals of Mathematical Statistics*, 23(4):493–507, Dec 1952.

[31] V. Chistyakov. A theorem on sums of independent positive random variables and its applications to branching random processes. *Theory of Probability & Its Applications*, 9:640–648, Jan 1964.

[32] J. Chover, P. Ney, and S. Wainger. Functions of probability measures. *Journal d'Analyse Mathématique*, 26:255–302, 1973.

[33] A. Clauset, C. Shalizi, and M. Newman. Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703, 2009.

[34] J. F. Collamore, G. Diao, and A. N. Vidyashankar. Rare event simulation for processes generated via stochastic fixed point equations. *The Annals of Applied Probability*, 24(5):2143–2175, October 2014.

[35] J. F. Collamore and A. Höing. Small-time ruin for a financial process modulated by a Harris recurrent Markov chain. *Finance and Stochastics*, 11(3):299–322, Jul 2007.

[36] J. F. Collamore and S. Mentemeier. Large excursions and conditioned laws for recursive sequences generated by random matrices. *The Annals of Probability*, forthcoming.

[37] J. F. Collamore and A. N. Vidyashankar. Tail estimates for stochastic fixed point equations via nonlinear renewal theory. *Stochastic Processes and their Applications*, 123(9):3378–3429, 2013.

[38] R. Cont and E. Voltchkova. Integro-differential equations for option prices in exponential lévy models. *Finance and Stochastics*, 9(3):299–325, Jul 2005.

[39] H. Cramér. Sur un nouveau théorème-limite de la théorie des probabilités. *Actual. Sci. Ind.*, 736:5–23, 1938.

[40] L. de Haan. *On regular variation and its application to the weak convergence of sample extremes*. PhD thesis, 1970.

[41] K. Debicki and M. Mandjes. *Queues and Lévy fluctuation theory*. Universitext. Springer International Publishing, 1st edition, 2015.

[42] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2009.

[43] D. Denisov, A. B. Dieker, and V. Shneer. Large deviations for random walks under subexponentiality: The big-jump domain. *The Annals of Probability*, 36(5):1946–1991, 2008.

[44] R. Douc, E. Moulines, and P. Soulier. Computable convergence rates for sub-geometric ergodic Markov chains. *Bernoulli*, 13(3):831–848, August 2007.

[45] P. Dupuis, K. Leder, and H. Wang. Importance sampling for sums of random variables with regularly varying tails. *ACM Transitions on Modeling and Computer Simulation*, 17(3), July 2007.

[46] P. Dupuis, A. D. Sezer, and H. Wang. Dynamic importance sampling for queueing networks. *The Annals of Applied Probability*, 17(4):1306–1346, August 2007.

[47] P. Dupuis and H. Wang. Importance sampling, large deviations, and differential games. *Stochastics and Stochastic Reports*, 76(6):481–508, 2004.

[48] P. Dupuis and H. Wang. On the convergence from discrete to continuous time in an optimal stopping problem. *The Annals of Applied Probability*, 15(2):1339–1366, 2005.

[49] P. Dupuis and H. Wang. Importance sampling for Jackson networks. *Queueing Systems*, 62(1):113–157, 2009.

[50] P. Dyszewski. Iterated random functions and slowly varying tails. *Stochastic Processes and their Applications*, 126(2):392–413, 2016.

[51] R. S. Ellis. Large deviations for a general class of random vectors. *The Annals of Probability*, 12(1):1–12, Feb 1984.

[52] P. Embrechts and C. M. Goldie. On convolution tails. *Stochastic Processes and their Applications*, 13(3):263 – 278, 1982.

[53] P. Embrechts, C. M. Goldie, and N. Veraverbeke. Subexponentiality and infinite divisibility. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 49(3):335–347, 1979.

[54] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events for insurance and finance*, volume 33 of *Stochastic Modelling and Applied Probability.* Springer-Verlag Berlin Heidelberg, 1st edition, 1997.

[55] P. Embrechts and J. Villasenor. Ruin estimates for large claims. *Insurance: Mathematics and Economics*, 7(4):269 – 274, 1988.

[56] W. Feller. *An introduction to probability theory and its applications*, volume 2 of *Wiley Series in Probability and Statistics*. Wiley, 1971.

[57] S. Foss, T. Konstantopoulos, and S. Zachary. Discrete and continuous time modulated random walks with heavy-tailed increments. *Journal of Theoretical Probability*, 20(3):581–612, Sep 2007.

[58] S. Foss and D. Korshunov. Heavy tails in multi-server queue. *Queueing Systems: Theory and Applications*, 52(1):31–48, January 2006.

[59] S. Foss and D. Korshunov. On large delays in multi-server queues with heavy tails. *Mathematics of Operations Research*, 37(2):201–218, 2012.

[60] S. Foss, D. Korshunov, and S. Zachary. *An introduction to heavy-tailed and subexponential distributions*, volume 38 of *Springer Series in Operations Research and Financial Engineering*. Springer-Verlag New York, 2nd edition, 2013.

[61] A. Ganesh, N. O'Connell, and D. Wischik. *Big queues*, volume 1838 of *Lecture Notes in Mathematics*. Springer-Verlag Berlin Heidelberg, 1st edition, 2004.

[62] J. Gärtner. On large deviations from the invariant measure. *Theory of Probability & Its Applications*, 22:24–39, 1977.

[63] P. Glasserman and S.-G. Kou. Analysis of an importance sampling estimator for tandem queues. *ACM Transactions on Modeling and Computer Simulation*, 5(1):22–42, January 1995.

[64] P. Glasserman and Y. Wang. Counterexamples in importance sampling for large deviations probabilities. *The Annals of Applied Probability*, 7(3):731–746, August 1997.

[65] P. Glynn. Lecture notes for MS&E Stochastic Calculus and Control. *Stanford University*, 2012.

[66] C. M. Goldie. Implicit renewal theory and tails of solutions of random equations. *The Annals of Applied Probability*, 1(1):126–166, 1991.

[67] C. M. Goldie and R. Grübel. Perpetuities with thin tails. *Advances in Applied Probability*, 28(2):463–480, 1996.

[68] D. R. Grey. Regular variation in the tail behaviour of solutions of random difference equations. *The Annals of Applied Probability*, 4(1):169–183, February 1994.

[69] H. Grincevičius. Limit theorems for products of random linear transformations on the line. *Lithuanian Mathematical Journal*, 15:568–579, 1975.

[70] T. Gudmundsson and H. Hult. Markov chain Monte Carlo for computing rare-event probabilities for a heavy-tailed random walk. *Journal of Applied Probability*, 51(2):359–376, June 2014.

[71] J. M. Harrison and M. I. Reiman. Reflected Brownian motion on an orthant. *The Annals of Probability*, 9(2):302–308, April 1981.

[72] E. Hashorva and J. Li. ECOMOR and LCR reinsurance with gamma-like claims. *Insurance: Mathematics and Economics*, 53(1):206–215, 2013.

[73] X. Hu and H. Cui. Generating multi-dimensional discrete distribution random number. In *2010 Sixth International Conference on Natural Computation*, volume 3, pages 1102–1104, August 2010.

[74] H. Hult and F. Lindskog. Extremal behavior of stochastic integrals driven by regularly varying Lévy processes. *The Annals of Probability*, 35(1):309–339, 2007.

[75] H. Hult, F. Lindskog, T. Mikosch, and G. Samorodnitsky. Functional large deviations for multivariate regularly varying random walks. *The Annals of Applied Probability*, 15(4):2651–2680, 2005.

[76] Z. I. Botev, A. Ridder, and L. Rojas-Nandayapa. Semiparametric cross entropy for rare-event simulation. *Journal of Applied Probability*, 53:633–649, 09 2016.

[77] S. F. Jarner and G. O. Roberts. Polynomial convergence rates of Markov chains. *The Annals of Applied Probability*, 12(1):224–247, 02 2002.

[78] J. Jiang and Q. Tang. Reinsurance under the LCR and ECOMOR treaties with emphasis on light-tailed claims. *Insurance: Mathematics and Economics*, 43(3):431–436, 2008.

[79] V. Kalashnikov and G. Tsitsiashvili. Tails of waiting times and their bounds. *Queueing Systems*, 32(1):257–283, 1999.

[80] J. Karamata. Sur un mode de croissance régulière des fonctions. *Mathematica (Cluj)*, 4:38–53, 1930.

[81] H. Kesten. Random difference equations and renewal theory for products of random matrices. *Acta Mathematica*, 131:207–248, 1973.

[82] C. Klüppelberg. Subexponential distributions and integrated tails. *Journal of Applied Probability*, 25(1):132–141, March 1988.

[83] D. G. Konstantinides and T. Mikosch. Large deviations and ruin probabilities for solutions to stochastic recurrence equations with heavy-tailed innovations. *The Annals of Probability*, 33(5):1992–2035, 2005.

[84] A. E. Kyprianou. *Introductory lectures on fluctuations of Lévy processes with applications*. Universitext. Springer-Verlag, Berlin, 2006.

[85] S. A. Ladoucette and J. L. Teugels. Reinsurance of large claims. *Journal of Computational and Applied Mathematics*, 186(1):163–190, 2006.

[86] J. Li. Asymptotics for large claims reinsurance in a time-dependent renewal risk model. *Scandinavian Actuarial Journal*, 2015(2):172–183, 2015.

[87] F. Lindskog, S. Resnick, and J. Roy. Regularly varying measures on metric spaces: Hidden regular variation and hidden jumps. *Probability Surveys*, 11:270–314, 2014.

[88] K. Maulik and B. Zwart. Tail asymptotics for exponential functionals of Lévy processes. *Stochastic Processes and their Applications*, 116:156–177, 2006.

[89] S. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, New York, NY, USA, 2nd edition, 2009.

[90] T. Mikosch and G. Samorodnitsky. Ruin probability with claims modeled by a stationary ergodic stable process. *The Annals of Probability*, 28(4):1814–1851, Oct 2000.

[91] T. Mikosch and O. Wintenberger. Precise large deviations for dependent regularly varying sequences. *Probability Theory and Related Fields*, 156(3):851–887, Aug 2013.

[92] T. Mikosch and O. Wintenberger. A large deviations approach to limit theory for heavy-tailed time series. *Probability Theory and Related Fields*, 166(1):233–269, Oct 2016.

[93] M. Mirek. Heavy tail phenomenon and convergence to stable laws for iterated Lipschitz maps. *Probability Theory and Related Fields*, 151(3-4):705–734, 2011.

[94] A. A. Mogulskii. Large deviations for trajectories of multi-dimensional random walks. *Theory of Probability & Its Applications*, 21(2):300–315, 1976.

[95] A. V. Nagaev. Integral limit theorems taking large deviations into account when Cramér's condition does not hold. I. *Theory of Probability & Its Applications*, 14(1):51–64, 1969.

[96] A. V. Nagaev. On a property of sums of independent random variables. *Theory of Probability & Its Applications*, 22(2):326–338, 1978.

[97] L. Nguyen-Ngoc and M. Yor. Lookback and barrier options under general Lévy processes. *Handbook of financial econometrics*, 2003.

[98] A. G. Pakes. On the tails of waiting-time distributions. *Journal of Applied Probability*, 12(3):555–564, 1975.

[99] Z. Palmowski and B. Zwart. Tail asymptotics of the supremum of a regenerative process. *Journal of Applied Probability*, 44(2):349–365, 2007.

[100] L. Peng. Joint tail of ECOMOR and LCR reinsurance treaties. *Insurance: Mathematics and Economics*, 58:116–120, 2014.

[101] S. Ramasubramanian. A subsidy-surplus model and the Skorokhod problem in an orthant. *Mathematics of Operations Research*, 25(3):509–538, 2000.

[102] S. I. Resnick. Point processes, regular variation and weak convergence. *Advances in Applied Probability*, 18(1):66–138, 1986.

[103] S. I. Resnick. *Extreme values, regular variation and point processes.* Springer Series in Operations Research and Financial Engineering. Springer-Verlag New York, 1st edition, 1987.

[104] S. I. Resnick. *Heavy-tail phenomena.* Springer Series in Operations Research and Financial Engineering. Springer-Verlag New York, 1st edition, 2007.

[105] C.-H. Rhee, J. Blanchet, and B. Zwart. Sample path large deviations for heavy-tailed Lévy processes and random walks. *eprint arXiv:1606.02795*, 2016.

[106] C.-H. Rhee and P. Glynn. Unbiased estimation with square root convergence for SDE models. *Articles in Advance*, pages 1–18, 2015.

[107] G. Rubino and B. Tuffin. *Rare event simulation using Monte Carlo methods.* John Wiley & Sons, 2009.

[108] W. Schoutens. *Lévy processes in finance: Pricing financial derivatives.* Wiley Series in Probability and Statistics. John Wiley & Sons, 2003.

[109] A. V. Skorokhod. Stochastic equations for diffusion processes in a bounded region. *Theory of Probability & Its Applications*, 6(3):264–274, 1961.

[110] A. V. Skorokhod. Stochastic equations for diffusion processes in a bounded region. II. *Theory of Probability & Its Applications*, 7(1):3–23, 1962.

[111] P. Tankov and R. Cont. *Financial modelling with jump processes.* Chapman and Hall/CRC Financial Mathematics Series. Taylor & Francis, 2nd edition, 2015.

[112] J. L. Teugels. The class of subexponential distributions. *The Annals of Probability*, 3(6):1000–1011, Dez 1975.

[113] A. Thépaut. Une nouvelle forme de réassurance: Le traité d'excédent du coût moyen relatif (ecomor). *Bulletin Trimestriel de l'Institut des Actuaires Français*, (49):273–343, 1950.

[114] G. Torrisi. Simulating the ruin probability of risk processes with delay in claim settlement. *Stochastic Processes and their Applications*, 112(2):225–244, 2004.

[115] N. Veraverbeke. Asymptotic behaviour of Wiener-Hopf factors of a random walk. *Stochastic Processes and their Applications*, 5(1):27–37, 1977.

[116] W. Vervaat. On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Advances in Applied Probability*, 11(4):750–783, 1979.

[117] W. Whitt. *Stochastic-process limits.* Springer Series in Operations Research and Financial Engineering. Springer-Verlag New York, 1st edition, 2002.

[118] S. Zachary. A note on Veraverbeke's theorem. *Queueing Systems*, 46(1-2):9–14, January 2004.

[119] B. Zwart. *Queueing systems with heavy tails.* PhD thesis, Department of Mathematics and Computer Science, 2001.

[120] B. Zwart, S. Borst, and M. Mandjes. Exact asymptotics for fluid queues fed by multiple heavy-tailed on–off flows. *The Annals of Applied Probability*, 14(2):903–957, May 2004.

Summary

Heavy tails: asymptotics, algorithms, applications

My research has been mainly focusing on large deviations theory and rare-event simulation in heavy-tailed settings. The starting point of the dissertation is a heavy-tailed sample path large deviations (SPLD) result recently established by Rhee, Blanchet, and Zwart [105]. By utilizing this result, we develop a simulation algorithm for a general class of rare events. We test the developed simulation algorithm on several applications. Moreover, we extend the result in [105] to the setting where the increments are no longer independent. Finally, we consider a rare event simulation problem, where the underlying processes possess super heavy-tailed features.

In the first part of the dissertation (Chapters 2 and 3), we propose a class of strongly efficient rare event simulation estimators for random walks and compound Poisson processes with regularly varying increments in a general large deviations regime. Our estimator is based on an importance sampling strategy that hinges on the heavy-tailed SPLD result in [105]. The estimators are straightforward to implement and can be used to systematically evaluate the probability of a wide range of rare events with bounded relative error. They are "universal" in the sense that a single importance sampling scheme applies to a very general class of rare events that arise in heavy-tailed systems. In particular, our estimators can deal with rare events that are caused by multiple big jumps as well as multidimensional processes such as the buffer content process of a queueing network.

At the end of Chapter 2, as well as, in Chapter 3, we illustrate the versatility of our approach with several applications that arise in the context of mathematical finance, actuarial science, and queueing theory.

In Chapter 4, we extend the result in [105] to stochastic processes with

increments that are driven by an autoregressive process. To relate our problem to the one in [105], we first identify a sequence of regeneration times and split the Markov chain into i.i.d. blocks. By aggregating the trajectory over regeneration cycles, we obtain a regenerative process with i.i.d. jump distributions and the regeneration times as renewals. Under a set of classical assumptions, we adapt a large deviation change of measure argument and further establish that the area under a typical regeneration cycle has an asymptotic power law. This approach brings us close to the framework of [105] and allows us to derive an SPLD result for the aggregated process. Finally, by showing that the residual process is negligible, we state an SPLD result for the original process.

In Chapter 5, we consider the stationary solution of an autoregressive process, which is also called stochastic perpetuity. We propose a strongly efficient simulation algorithm for estimating the tail probability of the perpetuities. Our most important idea behind these results is to connect the perpetuities to the maximum of a random walk. This allows us to construct a coupling and leverage an existing importance sampling algorithm. Since the stationary distribution is defined over an infinite horizon, it requires an infinite amount of computational effort for generating each sample when using a crude Monte Carlo sampling approach. To address this issue, we work with approximations by finite-time truncation. We study the bias introduced by such approximations and show that our estimator has a vanishing relative bias. Moreover, by making a slightly stronger assumption and applying a bias elimination technique, we propose strongly efficient and unbiased estimators.

We generalize the results in Chapters 4 and 5 to the case where the associated Markov chain is defined via a sequence of i.i.d. iterated random functions.

Curriculum Vitae

Bohan Chen was born in Nanjing (China) on October 12, 1988. He went to Germany as an exchange student in October 2005 and graduated from Johann-Vanotti-Gymnasium Ehingen in June 2008. From 2008 to 2011, Bohan proceeded to study Mathematics at the Technical University Munich. Subsequently, he completed his master's programme Mathematical Finance and Actuarial Science at the same university in July 2014 and obtained a master's degree of "passed with distinction". In September 2015, he started a Ph.D project at Centrum Wiskunde & Informatica in the Stochastics group under the supervision of Bert Zwart and Chang-Han Rhee. The results of this project are presented in this dissertation. Since September 2019 he is employed at Munich Re Group.