NON-GAUSSIAN SEMI-STABLE DISTRIBUTIONS AND THEIR STATISTICAL APPLICATIONS

Ritwik Chaudhuri

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Statistics and Operations Research in the University of North Carolina at Chapel Hill

Chapel Hill
2014

Approved by:

Shankar Bhamidi

Amarjit Budhiraja

Chuanshu Ji

Andrew Nobel

Vladas Pipiras

ABSTRACT

RITWIK CHAUDHURI: NON-GAUSSIAN SEMI-STABLE DISTRIBUTIONS AND THEIR STATISTICAL APPLICATIONS
(Under the direction of Vladas Pipiras)

The dissertation is motivated by problems arising in modern communication networks such as the Internet. Over these networks, information is sent in the form of data packets which are further grouped into flows. For example, a flow can be associated with a certain (document, music, movie or other) file. Knowing the structure of flows is of great interest to network operators and networking researchers. One quantity of particular interest is the distribution of flow sizes (the number of packets in a flow).

Each packet carries information on the flow it belongs to. Hence, examining all packets allows reconstructing and studying the associated flows. Examining all packets, however, is becoming cumbersome due to the ever increasing amount of data and processing costs. To overcome these issues, packet sampling has become prevalent. One common sampling scheme is probabilistic sampling wherein each packet is sampled independently and with the same probability. The basic problem then becomes inference of the characteristics of original flows (e.g. the flow size distribution) from sampled packets (forming sampled flows).

This problem, known as an inversion problem, has attracted much attention in the networking community. In particular, a well-known nonparametric estimator of the flow size distribution is available under probabilistic sampling, based on sampled packets and sampled flows. From the application perspective, the focus of the dissertation is on some statistical properties of this nonparametric estimator. Under suitable and restrictive assumptions, the estimator has been known to be asymptotically normal. Going beyond these assumptions, it is shown in the dissertation that the estimator can be asymptotically semi-stable.

To achieve this goal, the domains of attraction of semi-stable distributions are reexamined here. As a main theoretical contribution, general, sufficient and practical conditions are provided for a distribution to be in the domain of attraction of a semi-stable distribution. They lead to practical conditions for the aforementioned nonparametric estimator to be attracted to a semi-stable law. Examples of probability distributions and illustrations of the main results are provided throughout the dissertation. One practical consequence of the results is a confidence interval for the distribution of flow sizes, based on critical values of semi-stable distributions.

Semi-stable distributions do not have closed forms in general. In order to compute their critical values, numerical calculation of semi-stable densities is also considered in the dissertation. This is carried out by using a celebrated method of Joseph Abate and Ward Whitt, allowing numerical calculation of a density given its characteristic function (Laplace transform). The code implementing the method for semi-stable densities is included.

The numerically calculated densities are used to assess the goodness of approximations involving semi-stable distributions and, as indicated above, the computation of confidence intervals. These points are explored in numerical simulations throughout the dissertation. Finally, some multivariate extensions of the results and further directions are also discussed.

ACKNOWLEDGEMENTS

I wish to express my sincerest gratitude to my advisor Prof. Vladas Pipiras. He introduced me to the topic of heavy-tailed distributions and without his constant support, encouragement and invaluable insights this work would not have been possible. I thank him for his constant support and patience with me and teaching me the importance of hard work in every walk of life. I thank him for helping me whenever I needed it. I feel grateful and very fortunate to have him as my mentor and the lessons that I learned through this journey will stay with me for the rest of my life. I will take this opportunity to specially thank Prof. Amarjit Budhiraja for teaching me Measure Theory in my first semester at UNC Chapel Hill. I learnt a lot about probability through this class. My special thanks goes to Prof. Shankar Bhamidi for giving me honest suggestions whenever I needed it. He motivated me a lot through every discussion we had. I also thank Prof. Amarjit Budhiraja, Prof. Shankar Bhamidi, Prof. Chuanshu Ji and Prof. Andrew Nobel for being members of my dissertation committee and providing many useful comments.

I would like to thank many of my friends at Chapel Hill and the friends from Indian Statistical Institute dispersed all over India and the US. They have definitely made my life more colorful. I specially thank Sayan Dasgupta and Abhishek Pal Majumder for being my ever dependable roommates. Their helpfulness can not be described in words and any amount of acknowledgments will fall short. I am grateful to a number of friends for their constant support and for their trust. Special thanks are due for Anjishnu Bannerjee, Siddhartha Mandal, Pourab Roy, Pratyaydipta Rudra, Sujatro Chakladar, Poulomi Maitra, Suman Chakrobarty, Parthasarathi Mukherjee, Ritendranath Mitra, Suprateek Kundu and Swarnava Mukhopadhyaya. Without their constant support, this journey would not have been possible. My parents, my elder brother, my cousin sisters, my grandma are always in my mind. I would take this opportunity to thank my parents as the tiniest bit of token of appreciation for believing in me and for all their sacrifices throughout their life to make my life a little better than the day before. Their motivation helped me to achieve whatever I wanted to achieve in life. I thank my elder brother for teaching me Mathematics when I was a little boy. He will remain as ever supporting, most trusted and the best friend forever to me. I would not be here without their sacrifices, continuous support and unconditional love. Their pride in anything that I achieve is the warmest motivation for me. Finally I will take this opportunity to thank everyone whom I came across in my life. Their unconditional friendship shaped my life and helped me to become a better person.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

$\mathbb{N}$	Set of natural numbers
$\mathbb{N}_0$	Set of non-negative integers
$\mathbb{R}$	Set of real numbers
$\mathbb{R}_+$	Set of positive real numbers
$\mathbb{Z}$	Set of integers
$[x]$	Largest integer smaller than or equal to x
$\lceil x \rceil$	Smallest integer larger than or equal x
$\lceil x \rceil_+$	Smallest integer larger than x

CHAPTER 1: INTRODUCTION

We describe here the motivation behind this dissertation (Section 1.1), and outline its structure with main result (Section 1.2).

1.1 Motivation

Let W, W_i, $i = 1, 2, \ldots, N$, be i.i.d. integer-valued random variables with the probability mass function (p.m.f.) $f_W(w)$, $w \geq 1$. Let also $\mathrm{Bin}(n,\, q)$ denote a binomial distribution with parameters $n \geq 1$, $q \in (0, 1)$. Consider random variables $W_q, W_{q,i}$, $i = 1, 2, \ldots, N$, obtained from W, W_i, $i = 1, 2, \ldots, N$, through the relationships $W_q = \mathrm{Bin}(W, q)$ and $W_{q,i} = \mathrm{Bin}(W_i, q)$, $i = 1, 2, \ldots, N$ (independently across i). Note that W_q takes values in $0, 1, 2, \ldots, W$. Let the probability mass function of W_q be $f_{W_q}(s)$, $s \geq 0$. The basic interpretation of W_q is as follows. If an object consists of W points (a finite point process) and each point is sampled with a probability q, then the number of sampled points is $W_q = \mathrm{Bin}(W, q)$.

One application of the above setting arises in modern communication networks. A finite point process (an object) is associated with the so-called packet flow (and a point is associated with a single packet). Sampling is used in order to reduce the amount of data being collected and processed. One basic problem that has attracted much attention recently is the inference of f_W from the observed sampled data $W_{q,i}$, $i = 1, 2, \ldots, N$ (in principle, $W_{q,i} = 0$ is not observed directly, but the inference about the number of times $W_{q,i} = 0$ is made through other means). See, for example, Duffield, Lund and Thorup [14], Hohn and Veitch [16], Yang and Michailidis [28].

We are interested in this dissertation in some statistical properties of a nonparametric estimator of $f_W(w)$, introduced in Hohn and Veitch [16] and also considered in Antunes and Pipiras [2]. We first briefly outline how the estimator is derived. Estimation of $f_W(w)$ is based on a theoretical inversion of the relation

$$f_{W_q}(s) = \sum_{w=s}^{\infty} P(W_q = s | W = w) P(W = w) = \sum_{w=s}^{\infty} \binom{w}{s} q^s (1-q)^{w-s} f_W(w), \qquad s \geq 0. \tag{1.1.1}$$

In terms of the moment generating functions $G_{W_q}(z) = \sum_{s=0}^{\infty} z^s f_{W_q}(s)$ and $G_W(z) = \sum_{w=1}^{\infty} z^w f_W(w)$, the relation (1.1.1) can be written as $G_{W_q}(z) = G_W(zq + 1 - q)$. By changing the variables $zq + 1 - q = x$, one has $G_W(x) = G_{W_q}(q^{-1}x - q^{-1}(1-q))$ which has the earlier form but with q replaced by q^{-1}(and z replaced by x). This suggests that (1.1.1) can be inverted as

$$\begin{aligned} f_W(w) &= \sum_{s=w}^{\infty} \binom{s}{w} (q^{-1})^w (1-q^{-1})^{s-w} f_{W_q}(s) \\ &= \sum_{s=w}^{\infty} \binom{s}{w} \frac{(-1)^{s-w}}{q^s} (1-q)^{s-w} f_{W_q}(s), \quad w \geq 1. \end{aligned} \tag{1.1.2}$$

Antunes and Pipiras [2], Proposition 4.1, showed that the inversion relation (1.1.2) holds when

$$\sum_{s=n}^{\infty} \binom{s}{n} \frac{(1-q)^{s-n}}{q^s} f_{W_q}(s) = \sum_{w=n}^{\infty} \binom{w}{n} 2^{w-n} (1-q)^{w-n} f_W(w) < \infty, \quad n \geq 1. \tag{1.1.3}$$

Observe that (1.1.3) always holds when $q \in (0.5, 1)$. But when $q \in (0, 0.5]$, the finiteness of the above expression depends on the behavior of $f_W(w)$ as $w \to \infty$. We shall make the assumption (1.1.3) throughout this dissertation.

In view of (1.1.2), a natural nonparametric estimator of f_W is

$$\widehat{f}_W(w) = \sum_{s=w}^{\infty} \binom{s}{w} \frac{(-1)^{s-w}}{q^s} (1-q)^{s-w} \widehat{f}_{W_q}(s), \quad w \geq 1, \tag{1.1.4}$$

where

$$\widehat{f}_{W_q}(s) = \frac{1}{N} \sum_{i=1}^{N} 1_{\{W_{q,i}=s\}}, \quad s \geq 0, \tag{1.1.5}$$

is the empirical p.m.f. of f_{W_q}, and 1_A denotes the indicator function of an event A. Note that, by using (1.1.4) and (1.1.2),

$$\sqrt{N}(\widehat{f}_W(w) - f_W(w)) = \sum_{s=w}^{\infty} \binom{s}{w} \frac{(-1)^{s-w}}{q^s} (1-q)^{s-w} \sqrt{N}(\widehat{f}_{W_q}(s) - f_{W_q}(s)). \tag{1.1.6}$$

Since

$$\{\sqrt{N}(\widehat{f}_{W_q}(s) - f_{W_q}(s))\}_{s=0}^{\infty} \xrightarrow{d} \{\xi(s)\}_{s=0}^{\infty}, \tag{1.1.7}$$

where $\{\xi(s)\}_{s=0}^{\infty}$ is a Gaussian process with zero mean and covariance structure

$$E(\xi(s_1)\xi(s_2)) = f_{W_q}(s_1) 1_{\{s_1=s_2\}} - f_{W_q}(s_1) f_{W_q}(s_2),$$

one may naturally expect that under suitable assumptions, (1.1.6) is asymptotically normal in the sense that

$$\{\sqrt{N}(\widehat{f}_W(w) - f_W(w))\}_{w=1}^{\infty} \xrightarrow{d} \{S(\xi)_w\}_{w=1}^{\infty}, \tag{1.1.8}$$

where $\{S(\xi)_w\}_{w=1}^{\infty}$ is a Gaussian process. Antunes and Pipiras [2], Theorem 4.1, showed that (1.1.8) holds indeed if $R_{q,w} < \infty,\ w \geq 1$, where

$$\begin{aligned} R_{q,w} &= \sum_{s=w}^{\infty} \binom{s}{w}^2 \frac{(1-q)^{2(s-w)}}{q^{2s}} f_{W_q}(s) \\ &= \sum_{i=w}^{\infty} f_W(i)(1-q)^{i-2w} \binom{i}{w} \sum_{s=w}^{i} \binom{s}{w} \binom{i-w}{s-w} (q^{-1}-1)^s. \end{aligned} \tag{1.1.9}$$

We are interested in $\widehat{f}_W(w)$ when the condition $R_{q,w} < \infty,\ w \geq 1$, is not satisfied. In fact, such a situation is expected with many distributions. For example, we show in Section 3.2 (Chapter 3) below that if $f_W(w) = (1-c)c^{w-1}$, $w \geq 1$, is a geometric distribution with parameter $c \in (0,1)$, then the distribution of $f_{W_q}(s)$ is given by

$$f_{W_q}(s) = \begin{cases} \frac{(1-q)(1-c)}{1-c(1-q)}, & \text{if } s = 0, \\ \frac{1}{c} c_q^s (1-c_q), & \text{if } s \geq 1, \end{cases} \tag{1.1.10}$$

where $c_q = \frac{cq}{1-c(1-q)}$. Moreover, the condition $R_{q,w} < \infty$ holds if and only if $c < \frac{q}{1-q}$ (see Section 3.2). Thus, for example, we are interested what happens with $\widehat{f}_W(w)$ when W_q has p.m.f. given by (1.1.10) with $c \geq \frac{q}{1-q}$.

To understand what happens when $R_{q,w} = \infty$, observe from (1.1.4) and (1.1.5) that $\widehat{f}_W(w)$ can also be written as

$$\widehat{f}_W(w) = \frac{1}{N} \sum_{i=1}^{N} X_i, \tag{1.1.11}$$

where X_i, $i = 1, 2, \ldots, N$, are i.i.d. random variables defined as

$$X_i = \binom{W_{q,i}}{w} \frac{(-1)^{W_{q,i}-w}}{q^{W_{q,i}}} (1-q)^{W_{q,i}-w} 1_{\{W_{q,i} \geq w\}}. \tag{1.1.12}$$

Focus on the key term $\frac{(1-q)^{W_{q,i}}}{q^{W_{q,i}}} = (q^{-1}-1)^{W_{q,i}}$ entering (1.1.12). For example, when W is geometric with parameter c, $W_{q,i}$ has p.m.f. in (1.1.10). One then expects that

$$\begin{aligned} P((q^{-1}-1)^{W_{q,i}} > x) &= P\left(W_{q,i} > \frac{\log x}{\log(q^{-1}-1)}\right) \\ &\approx \frac{1}{c} c_q^{\frac{\log x}{\log(q^{-1}-1)}} = \frac{1}{c} x^{-\alpha}, \end{aligned} \tag{1.1.13}$$

where $\alpha = \frac{\log c_q^{-1}}{\log(q^{-1}-1)}$. This suggests that the distribution of X_i, $i = 1, 2, \dots, N$, has heavy tail and that the estimator $\widehat{f}_W(w)$ is asymptotically non-Gaussian stable when $\alpha < 2$. In fact, the story turns out to be more complex. Because of the discrete nature of $W_{q,i}$, the relation (1.1.13) does not hold in the asymptotic sense as $x \to \infty$. An appropriate setting in this case involves the so-called semi-stable laws. In the semi-stable context, moreover, the convergence of (1.1.11) is expected only along subsequences of N.

Semi-stable laws have been studied quite extensively (see Section 2.1 for references). In particular, necessary and sufficient conditions are known for a distribution to be attracted to a semi-stable law (see Theorem 2.1.1 below). Verifying whether the distribution of (1.1.12) satisfies these conditions for a large class of p.m.f.'s f_{W_q}, however, turns out to be highly nontrivial. Much of this dissertation, in fact, concerns this problem. We identify a large class of p.m.f.'s, including (1.1.10), for which the distribution X_i satisfies the sufficient (and necessary) conditions and hence is attracted to a semi-stable law. We also study what this result means for the nonparametric estimator $\widehat{f}_W(w)$.

1.2 Structure of dissertation and main results

More specifically, the structure of the dissertation is as follows. Chapter 2 concerns semi-stable distributions. In Section 2.1, we recall the definition and basic properties of semi-stable distributions. We also provide known necessary and sufficient conditions for a distribution to be in the domain of attraction of a semi-stable distribution (Theorem 2.1.1).

For later applications, we will need to have critical values for semi-stable distributions. The distribution function and the density of a semi-stable distribution are not available in closed form (except special cases) but its Laplace transformation is. In Section 2.2, we recall a popular method of Abate and Whitt [1] to calculate numerically the density of a distribution given its Laplace transform. Numerical illustrations of the method applied to semi-stable laws can be found in Section 2.3.

In Section 2.3, we also evaluate numerically how well a semi-stable law approximates partial sums in its domain of attraction. Our findings here are that the goodness of the approximation very much depends on a parameter involved in the convergence result (more specifically the constant c in (2.1.5) below) with the performance degrading as the parameter increases.

Chapter 3 concerns the domains of attraction of semi-stable distributions. Though necessary and sufficient conditions for a distribution to be in the domain of attraction are available (and recalled in Section 2.1 as indicated above), verifying those conditions is quite nontrivial in general. In Section 3.1, we provide a large family of distributions which satisfy these conditions and hence are attracted to semi-stable law (Theorem 3.1.1 and Corollary 3.1.2). The proof of the main result (Theorem 3.1.1) is quite lengthy. The form of the family is motivated by applications to sampling.

In Section 3.1, we not only show the convergence to a semi-stable law but also address the following related issues: the forms of a subsequence (along which there is a convergence) and normalizing constants (Theorem 3.1.1); centering constants (Proposition 3.1.1); the behavior of the partial sums across the sequence of all indices

(Proposition 3.1.2). These results are important for practical applications, as in the case of sampling of finite point processes.

In Section 3.2, we consider a concrete example of a distribution from the derived family attracted to a semi-stable law. We also study numerically the goodness of approximation (of a partial sum by the limit semi-stable law) in finite samples.

In Chapter 4 and Section 4.1, in particular, we return to the sampling framework described in Section 1.1 and are interested in the asymptotics of the nonparametric estimator $f_W(w)$ in (1.1.4) of the p.m.f. $f_W(w)$ of the number of points in a finite point process. Under suitable assumptions and by using the results of Chapter 3 we conclude that $\widehat{f}_W(w)$ can be asymptotically semi-stable (Theorem 4.1.1). This result can be used to construct confidence intervals for $f_W(w)$ (Proposition 4.1.1).

In Section 4.2, we illustrate the results of Section 4.1 with several concrete examples of p.m.f. $f_W(w)$. In the first example, W follows a geometric distribution, and in the second example, W follows a negative binomial distribution.

In Chapter 5, we discuss several open problems and future directions.

In Appendix A1, we give all the auxiliary results that have been used in proving the main result in Section 3.1. In Appendix A2, we also provide the code for calculating numerically some semi-stable densities.

CHAPTER 2: SEMI-STABLE DISTRIBUTIONS

This chapter concerns semi-stable distributions. Basic facts about semi-stable distributions and known results on their domains of attraction are recalled in Section 2.1. The Abate and Whitt [1] method for calculating numerically a density function is described in Section 2.2, and it is applied to semi-stable distributions in Section 2.3.

2.1 Definitions and domains of attractions

One way to characterize a semi-stable distribution is through its characteristic function (Maejima [20]).

Definition 2.1.1. A probability distribution μ on $\mathbb{R}$ (or a random variable with distribution μ) is called *semi-stable* if there exist $r, b \in (0,1)$ and $c \in \mathbb{R}$ such that

$$\widehat{\mu}(\theta)^r = \widehat{\mu}(b\theta)e^{ic\theta}, \text{ for all } \theta \in \mathbb{R}, \tag{2.1.1}$$

and $\widehat{\mu}(\theta) \neq 0$, for all $\theta \in \mathbb{R}$, where

$$\widehat{\mu}(\theta) = \int_{\mathbb{R}} e^{i\theta x}\mu(dx)$$

denotes the characteristic function of μ.

A semi-stable distribution is known to be infinitely divisible (Maejima [20]) with a location parameter $\eta \in \mathbb{R}$, a Gaussian part with variance $\sigma^2 \geq 0$ and a non-Gaussian part with Lévy measure characterized by (distribution) functions

$$L(x) = \frac{M_L(x)}{|x|^\alpha}, \quad x < 0, \qquad R(x) = -\frac{M_R(x)}{x^\alpha}, \quad x > 0, \tag{2.1.2}$$

where $\alpha \in (0,2)$, $M_L(c^{\frac{1}{\alpha}}x) = M_L(x)$ when $x < 0$, and $M_R(c^{\frac{1}{\alpha}}x) = M_R(x)$ when $x > 0$, for some $c > 0$. The functions M_L and M_R are thus periodic with multiplicative period $c^{\frac{1}{\alpha}}$. The functions $L(x)$ and $R(x)$ are left-continuous and non-decreasing on $(-\infty, 0)$ and right-continuous and non-decreasing on $(0,\infty)$, respectively. The characteristic function of a semi-stable distribution with a location parameter η and without a Gaussian part is given by

$$\log\widehat{\mu}(t) = i\eta t + \int_{-\infty}^{0}(e^{itx} - 1 - \frac{itx}{1+x^2})dL(x) + \int_0^\infty (e^{itx} - 1 - \frac{itx}{1+x^2})dR(x). \tag{2.1.3}$$

Stable distributions are special cases of semi-stable distributions corresponding to $M_L(x) \equiv c_1$ and $M_R(x) \equiv c_2$, where $c_1 \geq 0$ and $c_2 \geq 0$ are two constants.

Semi-stable distributions arise as limits of partial sums of i.i.d. random variables. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with a common distribution function F. Consider the sequence of partial sums

$$S_n^* = \frac{1}{A_{k_n}}\Big\{\sum_{j=1}^{k_n} X_j - B_{k_n}\Big\}, \tag{2.1.4}$$

where $\{A_{k_n}\}$ and $\{B_{k_n}\}$ are normalizing and centering sequences. Semi-stable laws arise as limits of partial sums S_n^*, supposing that $\{k_n\}$ satisfies

$$k_n \to \infty,\ k_n \leq k_{n+1},\ \lim_{n\to\infty}\frac{k_{n+1}}{k_n} = c \in [1,\infty). \tag{2.1.5}$$

Moreover, if S_n^* converges to a nontrivial limit (semi-stable distribution), the distribution F of X_j is said to be in the *domain of attraction* of the limiting semi-stable law. In this case and supposing the limiting law is non-Gaussian semi-stable, it is known that the normalizing sequence $\{A_{k_n}\}$ necessarily satisfies

$$A_{k_n} \to \infty,\ A_{k_n} \le A_{k_{n+1}},\ \lim_{n\to\infty} \frac{A_{k_{n+1}}}{A_{k_n}} = c^{\frac{1}{\alpha}}, \tag{2.1.6}$$

where $\alpha \in (0,2)$.

Megyesi [23], Grinevich and Khokhlov [15] gave necessary and sufficient conditions for a distribution to be in the domain of attraction of a semi-stable distribution. Recall that a function L is *slowly varying at infinity* if it is positive and, for every $a > 0$,

$$\frac{L(ax)}{L(x)} \to 1, \qquad \text{as } x \to \infty.$$

Examples of slowly varying functions at infinity are:

- $L(x) = c, x > 0$; more generally, $L(x) \sim c > 0$, as $x \to \infty$.
- $L(x) = \log x, x > 1$.
- $L(x) = (\log x)^\beta, x > 1, \beta \in \mathbb{R}$.
- $L(x) = \log\log x, x > e$.

Theorem 2.1.1. (Megyesi [23], Corollary 3) *Distribution F is in the domain of attraction of a non-Gaussian semi-stable distribution with the characteristic function (2.1.3) along the subsequence k_n with normalizing constants A_{k_n} satisfying (2.1.5) and (2.1.6) if and only if for all $x > 0$ large enough,*

$$F_{\text{-}}(-x) = x^{-\alpha} l^*(x)(M_L(-\delta(x)) + h_L(x)), \tag{2.1.7}$$

$$1 - F(x) = x^{-\alpha} l^*(x)(M_R(\delta(x)) + h_R(x)), \tag{2.1.8}$$

where l^* is a right-continuous function, slowly varying at ∞, $\alpha \in (0, 2)$, F_- is the left-continuous version of F and the error functions h_R and h_L are such that

$$h_K(A_{k_n} x_0) \to 0, \ \text{as } n \to \infty, \tag{2.1.9}$$

for every continuity point x_0 of M_R, if $K = R$, and $-x_0$ of M_L, if $K = L$. M_K, $K \in \{L, R\}$, are two periodic functions with common multiplicative period $c^{\frac{1}{\alpha}}$ and for all large enough x, $\delta(x)$ is defined as

$$\delta(x) = \frac{x}{a(x)} \in [1, c^{\frac{1}{\alpha}} + \epsilon], \tag{2.1.10}$$

where $\epsilon > 0$ is any fixed number, with

$$a(x) = A_{k_n} \ \text{if } A_{k_n} \leq x < A_{k_{n+1}}. \tag{2.1.11}$$

Grinevich and Khokhlov [15] also showed that, in the sufficiency part of the theorem above, k_n can be chosen as follows. First, choose a sequence $\{\tilde{A}_n\}$ such that

$$\lim_{n\to\infty} n\tilde{A}_n^{-\alpha} l^*(\tilde{A}_n) = 1 \tag{2.1.12}$$

and

$$\tilde{A}_n \to \infty, \quad \tilde{A}_n \leq \tilde{A}_{n+1} \quad \text{and} \quad \lim_{n\to\infty} \frac{\tilde{A}_{n+1}}{\tilde{A}_n} = 1. \tag{2.1.13}$$

Define a new sequence $\{a_n\}$ by setting $a_n = A_{k_n}$ for every n, where A_{k_n} appears in

(2.1.11). Then, the natural numbers k_n can be chosen as

$$\tilde{A}_{k_n} \leq a_n < \tilde{A}_{k_{n+1}}. \tag{2.1.14}$$

The centering constants B_{k_n} in (2.1.4) can be chosen as (Csörgö and Megyesi [11])

$$B_{k_n} = k_n \int_{\frac{1}{k_n}}^{1-\frac{1}{k_n}} Q(s)ds, \tag{2.1.15}$$

where, for $0 < s \leq 1$,

$$Q(s) = \inf_y \{F(y) \geq s\} \tag{2.1.16}$$

is the quantile functions. The location parameter η of the limiting semi-stable law in (2.1.3) is then given by

$$\eta = \Theta(\psi_1) - \Theta(\psi_2), \tag{2.1.17}$$

where

$$\Theta(\psi_i) = \int_0^1 \frac{\psi_i(s)}{1+\psi_i^2(s)} ds - \int_1^\infty \frac{\psi_i^3(s)}{1+\psi_i^2(s)} ds, \quad i = 1, 2, \tag{2.1.18}$$

and

$$\psi_1(s) = \inf_{x<0} \{L(x) > s\}, \qquad \psi_2(s) = \inf_{x<0} \{-R(-x) > s\}. \tag{2.1.19}$$

It is also worth mentioning that the slowly varying function $l^*(x)$ entering in (2.1.7) and (2.1.8) can be replaced by two different, asymptotically equivalent slowly varying functions $l_1^*(x)$ and $l_2^*(x)$. The proof of this result is given in Lemma A.1.5 in Appendix A.1.

For later reference we also provide a result on the domains of attraction of semi-stable laws exspressed using the quantile function in (2.1.16). A function l is *slowly* *varying at zero* if the function $L(x) = l(\frac{1}{x})$ is slowly varying at infinity.

Theorem 2.1.2. (Megyesi [23], Corollary 3) *Distribution F is in the domain of attraction of a non-Gaussian semi-stable distribution with the characteristic function (2.1.3) along the subsequence k_n satisfying (2.1.5) and centering constants B_{k_n} as given in (2.1.15) if and only if for all $s \in (0,1)$ small enough the quantile function Q pertaining to F is of the form*

$$Q_+(s) = -s^{-\frac{1}{\alpha}} l(s)(M_1(\gamma(s)) + h_1(s)), \tag{2.1.20}$$

$$Q(1-s) = s^{-\frac{1}{\alpha}} l(s)(M_2(\gamma(s)) + h_2(s)), \tag{2.1.21}$$

where $0 < \alpha < 2$, l is right-continuous and slowly varying at 0, and h_1 and h_2 are some right-continuous functions such that

$$h_j(\frac{s}{k_n}) \to 0, \qquad \text{as } n \to \infty, \tag{2.1.22}$$

for every continuity point s of M_j, $j = 1, 2$. The function $\gamma(s)$ appearing above is defined as $\gamma(s) := sk_{n*(s)}$ where $k_{n*(s)}$ is uniquely determined by the relation $\frac{1}{k_{n*(s)}} \leq s < \frac{1}{k_{n*(s)-1}}$ for all small enough values of s. M_1 and M_2 are non-negative, right-continuous functions, at least one of which is not identically equal to zero. Moreover, if $M_j \neq 0$, then it is bounded away from zero and infinity and has a multiplicative period c, that is, $M_j(cs) = M_j(s), s > 0$.

A number of papers can be found in the literature on semi-stable distributions. A detailed characterization of semi-stable distributions and their domains of attraction was first presented by Shimizu [26], Pillai [24] and Kruglov [19]. As indicated above, Theorem 2.1.1 is stated following Megyesi [23], Grinevich and Khokhlov [15]. A probabilistic approach to the domains of partial attraction was initiated by Csörgö, Haeusler and Mason [9]. A few extensions of their work can be found in Csörgö [5], [6]. Properties of semi-stable distributions are considered in Watanabe and Yamamuro [27], Bouzar [4], Meerschaert and Scheffler [21].

St. Petersburg paradox is one of the most famous paradoxes and it is related to semi-stable distributions. A number of research papers can be found in the literature which discuss various extensions of St. Petersburg paradox. See the papers by Csörgö [7], Csörgö and Dodunekova [8], Csörgö and Simons [13]. A brief description of St. Petersburg problem is as follows.

John offers David to toss a coin until a head is obtained and pays him $r^{\frac{k}{\alpha}}$ dollars if this happens in the kth toss where $k \in \mathbb{N}$. Moreover $r = \frac{1}{1-p}$ and p is the probability of obtaining a head in each toss. The parameter $\alpha > 0$ is known as the pay off parameter. This particular example is a generalized version of St. Petersburg problem. If X is the amount won by David, then $P(X = r^{\frac{k}{\alpha}}) = (1-p)^{k-1}p$, $k \in \mathbb{N}$. Then, observe that

$$P(X \leq x) = \begin{cases} 0, & \text{if } x < r^{\frac{1}{\alpha}}, \\ 1 - (1-p)^{\lfloor \alpha \log_r x \rfloor} = 1 - \frac{\{r^{\alpha \log_r x}\}}{x^{\alpha}}, & \text{if } x \geq r^{\frac{1}{\alpha}}, \end{cases} \tag{2.1.23}$$

where $\lfloor y \rfloor$ is the largest integer smaller than or equal to y and $\{y\} = y - \lfloor y \rfloor$ is the fractional part of $y \in \mathbb{R}$. Observe that the pay off parameter $\alpha > 0$ is the tail exponent of the distribution. Observe that $E(X^{\alpha}) = \infty$ but $E(X^{\beta}) = \frac{p}{q^{\beta/\alpha} - q}$ is finite when $\beta \in (0, \alpha)$. So, if $\alpha > 2$, then the variance of X is finite.

Consider the case $\alpha \in (0, 2)$ and the cumulative earning $S_n = X_1 + X_2 + \ldots + X_n$ if the game is repeated n times. Csörgö and Simons [12] considered the problem when $\alpha = 1$ and $r = 2$. For this case $E(X) = \infty$ and so is $E(X^2)$. They showed that after suitable centering and normalization and along suitable subsequences, the partial sum S_n converges to a semi-stable distribution. For further reading on the generalized version of the St. Petersburg problem, see Csörgö and Kevei [10].

2.2 Numerical calculation of densities

In applications of semi-stable distributions below, we will need their critical values. These values can not be obtained in a direct way since semi-stable distribution functions and densities are not available in explicit forms (except in special cases). To obtain the critical values, we will calculate the density numerically from the characteristic function (Laplace transform) of a semi-stable distribution by using one of the inversion methods due to Abate and Whitt [1].

To apply the method of Abate and Whitt [1] for numerically calculating the density from the Laplace transform, one needs to make an assumption that the density function is supported by the positive real line. (If the density is not supported by the positive real line, the method is applied in practice to the density shifted to the right sufficiently so that most of the mass concentrates on the positive real line.) For such a density function f, consider its Laplace transformation defined by

$$\mu^*(s) = \int_0^\infty e^{-sx} f(x)dx \quad s \in \mathbb{C}, \Re(s) \geq 0, \tag{2.2.1}$$

where $\Re$ stands for the real part. If the characteristic function of f is denoted by $\widehat{\mu}(s)$, then

$$\mu^*(s) = \widehat{\mu}(-is). \tag{2.2.2}$$

As described by Abate and Whitt [1], a standard inversion formula can be used to express the density function f in terms of its Laplace transform $\mu^*(s)$ as: for any

$a > 0$,

$$f(s) = \frac{2e^{as}}{\pi} \int_0^\infty \Re(\mu^*(a+iu)) \cos(us) du, \quad s > 0. \tag{2.2.3}$$

Moreover, by applying the trapezoidal rule to the integral appearing in (2.2.3), one gets the approximate relation

$$f(s) \approx \frac{e^{\frac{A}{2}}}{2s} \mu^* \left(\frac{A}{2s} \right) + \frac{e^{\frac{A}{2}}}{s} \sum_{k=1}^\infty (-1)^k \Re \left(\mu^* \left(\frac{A + 2k\pi i}{2s} \right) \right), \tag{2.2.4}$$

where A is a tuning parameter taking positive values. The goodness of the approximation on the right-hand side depends on the choice of the value A. This is illustrated next through the numerical calculation of several densities of stable distributions. The Matlab code for numerical calculation is included in Appendix A.2.

The characteristic function of a stable distribution with exponent α is

$$\widehat{\mu}(t) = \exp\{it\eta - |ct|^\alpha (1 - i\beta \text{sign}(t)\Phi(t))\}, \quad t \in \mathbb{R}, \tag{2.2.5}$$

where sign(t) is the sign of t and

$$\Phi(t) = \begin{cases} \tan(\frac{\pi\alpha}{2}), & \text{if } \alpha \neq 1, \\ -\frac{2}{\pi} \log |t|, & \text{if } \alpha = 1 \end{cases} \tag{2.2.6}$$

(see, for example, Samorodnitsky and Taqqu [25]). In (2.2.5), η is the location parameter, c is the scale parameter and β is the skewness parameter taking values in $[-1, 1]$. The stability index α takes values in $(0, 2]$. When $1 < \alpha < 2$, the second moment and the higher order moments of the distribution are infinite. For $\alpha < 1$, all the moments of the distribution are infinite. In the case $\alpha = 2$, the stable distribution is a normal distribution with mean η and standard deviation c. We

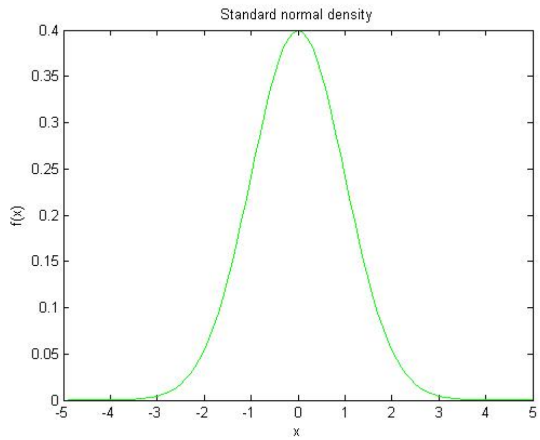

Figure 2.1: Plot of the standard normal distribution.

give a few examples next with the use of Abate and Whitt method of Section 2.2 to calculate numerically the density of a stable distribution.

Example 2.2.1. The stable distribution with $\alpha = 2$, $c = \frac{1}{\sqrt{2}}$, $\eta = 0$ and $\beta = 0$ is the standard normal distribution $N(0,1)$, having the characteristic function

$$\widehat{\mu}(t) = e^{-\frac{t^2}{2}}, \quad t \in \mathbb{R}.$$

Since the density f of $N(0,1)$ distribution is concentrated on the whole real line, in order to apply the method of Abate and Whitt [1], we shift the density by S to the right and numerically calculate the shifted density $f(x+S)$ between $\frac{S}{2}$ and $\frac{3S}{2}$. After the shifted density is obtained, it is brought back to the left by S to get the density of f within $-\frac{S}{2}$ to $\frac{S}{2}$. Note that the Laplace transform $f(x+S)$ is given by

$$\mu^*(t) = e^{\frac{t^2}{2} - St}, \quad t \in \mathbb{C}.$$

In Figure 2.1, we present the numerically calculated density of $N(0,1)$ by taking $S = 10$ and the tuning parameter $A = 9$. The numerically calculated density matches with the actual density of $N(0,1)$ on the interval $[-5, 5]$ – the two curves overlay in

Figure 2.1.

Example 2.2.2. Let f be the density of a stable distribution with the skewness parameter $\beta = 1$ and the location parameter $\eta = 0$. The log Laplace transformation is given by

$$\log \mu^*(t) = \begin{cases} -\frac{c^\alpha t^\alpha}{\cos(\frac{\pi\alpha}{2})}, & \text{if } \alpha \neq 1, \\ \frac{2ct}{\pi} \log t, & \text{if } \alpha = 1, \end{cases} \tag{2.2.7}$$

where $t \in \mathbb{C}, \Re(t) > 0$ and c is the scale parameter. It is known (e.g. Zolotarev [29], Samorodnitsky and Taqqu [25]) that the stable density is concentrated on the whole real line. But whereas its right tail is heavy-tailed, the left tail has an exponential decay. To numerically calculate the density f by using the method of Section 2.1, we shift f to the right by a margin of $S > 0$ by considering the density $f(x+S)$. After numerically calculating the density $f(x+S)$, we shift it to the left by a margin of S to get the density f.

Figure 2.2 presents the plots of the density approximation on the interval -9.9 to 60 for the choices of $A = 1$, 2, 3 and 9. We have used the value of $S = 10$ for the shift of the density. Figure 2.3 presents the plot of the actual density of the stable distribution (obtained using the R function *dstable* in the package *stabledist*) showing an agreement with the numerical calculation in Figure 2.2, when $A = 9$. Such larger values of A are also recommended by Abate and Whitt [1].

2.3 Semi-stable densities

We apply here the method of Abate and Whitt described in Section 2.2 to calculate numerically the densities of several semi-stable distributions. We are also interested in the goodness of the approximation of the partial sums (2.1.4) by the semi-stable

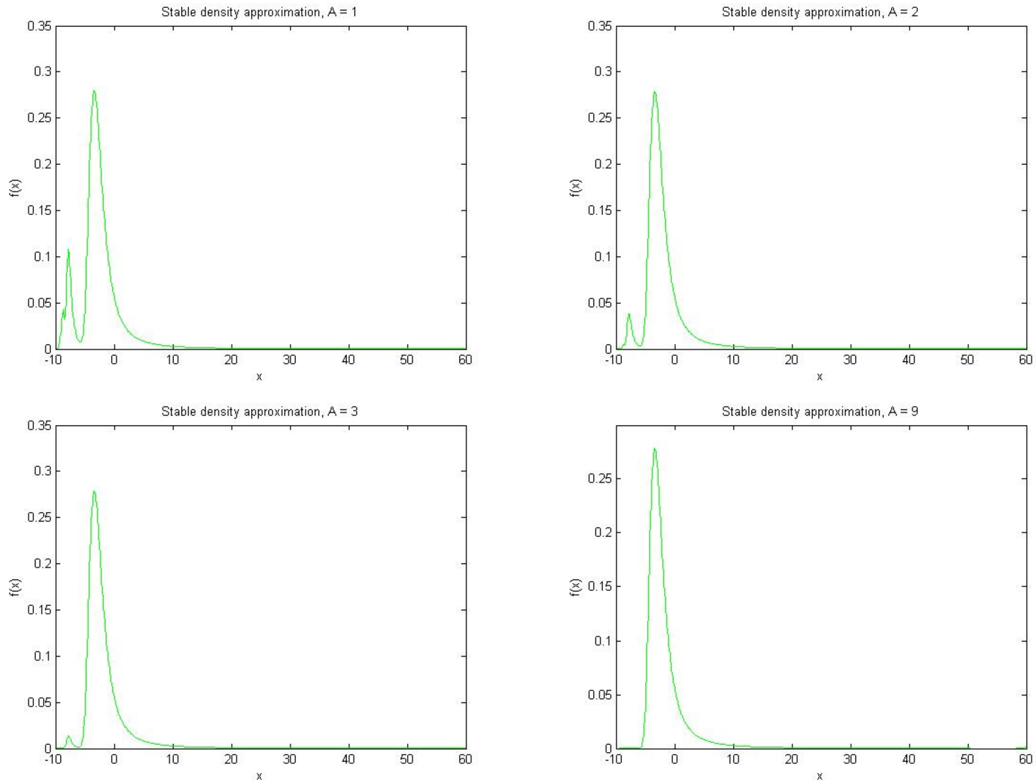

Figure 2.2: Stable density with $\alpha = 1.2$, $c = 1$, $\beta = 1$.

distribution appearing in the limit of their domain of attraction. This will be carried out by comparing the histograms of the partial sums and the corresponding densities of the semi-stable distributions. The findings of the section will serve as a guide in the applications of semi-stable laws to sampling (Chapter 4).

We thus begin with a distribution in the domain of attraction of a semi-stable law according to the theorems in Section 2.1 and characterize the characteristic function of the limiting semi-stable distribution. We shall follow the notation consistent with the sampling framework briefly discussed in Section 1.1 and further developed in Chapters 3 and 4.

Example 2.3.1. Let $X, X_1, X_2, \ldots$ be a sequence of i.i.d. random variables defined as

$$X = e^{\beta W_q},$$

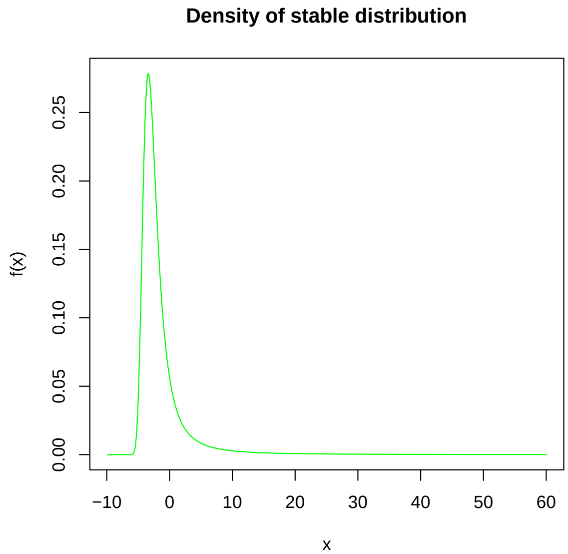

Figure 2.3: Plot of actual stable density with $\alpha = 1.2$, $c = 1$, $\beta = 1$.

where $\beta > 0$ and W_q follows a geometric distribution with parameter $s \in (0, 1)$, that is, its p.m.f. is

$$f_{W_q}(x) = s^{x-1}(1-s), \quad x = 1, 2, \ldots. \tag{2.3.1}$$

We shall apply Theorem 2.1.1 to show that, depending on the choices of s and β, the distribution F of X is in the domain of attraction of a semi-stable distribution. We shall also derive the sequences $\{A_{k_n}\}$, $\{B_{k_n}\}$ and $\{k_n\}$ for the convergence of the corresponding partial sums (2.1.4).

First, observe that

$$\begin{aligned} 1 - F(x) = P(X > x) &= P(e^{\beta W_q} > x) = P(W_q > \frac{1}{\beta}\log x) = \frac{1}{s} s^{\lceil \frac{1}{\beta}\log x \rceil_+} \\ &= \frac{1}{s} x^{-\frac{1}{\beta}\log\frac{1}{s}} s^{\lceil \frac{1}{\beta}\log x \rceil_+ - \frac{1}{\beta}\log x}, \end{aligned} \tag{2.3.2}$$

where $\lceil x \rceil_+$ denotes the smallest integer strictly larger than x. In view of (2.1.8), this suggests $\alpha = \frac{1}{\beta}\log\frac{1}{s}$. Now, consider

$$M_R(x) = e^{-\log(\frac{1}{s})(\lceil \frac{1}{\beta}\log x\rceil_+ - \frac{1}{\beta}\log x)},$$

which is a right-continuous function with a multiplicative period $e^{\beta} = c^{\frac{1}{\alpha}}$ with $c = \frac{1}{s}$. Taking $A_{k_n} = e^{\beta(n-1)}$, we have

$$\delta(x) = \frac{x}{e^{\beta(n-1)}}, \qquad \text{if } e^{\beta(n-1)} \leq x < e^{\beta n}.$$

Then, when $\alpha = \frac{1}{\beta}\log\frac{1}{s} < 2$, the right-hand side of (2.3.2) is exactly of the form (2.1.8) with $h_R(x) \equiv 0$ and $l^*(x) = \frac{1}{s}$. Hence, F is in the domain of attraction of a semi-stable law.

We derive the form of k_n by first defining the sequence $\tilde{A}_n$ as described following Theorem 2.1.1. According to (2.1.12), we need to have $\lim_{n\to\infty}\frac{n}{s}\tilde{A}_n^{-\alpha} = 1$. So, we define

$$\tilde{A}_n = (\frac{n}{s})^{\frac{1}{\alpha}}.$$

By (2.1.14), k_n are then defined as the natural numbers satisfying

$$(\frac{k_n}{s})^{\frac{1}{\alpha}} \leq e^{\beta(n-1)} < (\frac{k_{n+1}}{s})^{\frac{1}{\alpha}}.$$

Substituting $\alpha = \frac{1}{\beta}\log\frac{1}{s}$ above, we get

$$k_n \leq (\frac{1}{s})^{n-2} < k_{n+1},$$

so that $k_n = \lfloor(\frac{1}{s})^{n-2}\rfloor$.

The centering constants $\{B_{k_n}\}$ can be obtained from (2.1.15) and (2.1.16), involving the quantile function Q. We shall derive the asymptotic behavior of the quantity $\frac{B_{k_n}}{A_{k_n}} = \frac{k_n}{A_{k_n}} \int_{\frac{1}{k_n}}^{1-\frac{1}{k_n}} Q(t)dt$ for two different cases $0 < \alpha < 1$ and $1 < \alpha < 2$. The function $Q(t)$ is defined as the inverse of the distribution function $F(x) = P(e^{\beta W_q} \leq x)$. The function $F(x)$ has jumps at points $x = e^{\beta n}$ of sizes $(1-s)s^{n-1}, n \geq 1$. This means that the inverse function $Q(t)$ has jumps at points $t = 1 - s^m, m \geq 1$, of sizes $e^{\beta(m+1)} - e^{\beta m} = e^{\beta m}(e^{\beta} - 1)$. Moreover, $Q(t) = e^{\beta m}$ when $1 - s^{m-1} \leq t < 1 - s^m, m \geq 1$. Also note that $e^{\beta s} > 1$ when $0 < \alpha < 1$ and $e^{\beta s} < 1$ when $1 < \alpha < 2$.

Assume for simplicity that $\frac{1}{s}$ is an integer so that $k_n = (\frac{1}{s})^{n-2}$. Then,

$$\frac{k_n}{A_{k_n}} \int_{\frac{1}{k_n}}^{1-\frac{1}{k_n}} Q(t)dt = \frac{k_n}{A_{k_n}} \int_0^{1-\frac{1}{k_n}} Q(t)dt - \frac{k_n}{A_{k_n}} \int_0^{\frac{1}{k_n}} Q(t)dt.$$

Now,

$$\frac{k_n}{A_{k_n}} \int_0^{\frac{1}{k_n}} Q(t)dt \sim \frac{k_n}{A_{k_n}} \frac{1-s}{k_n} = \frac{1-s}{A_{k_n}} \to 0.$$

Moreover,

$$\begin{aligned}
\frac{k_n}{A_{k_n}} \int_0^{1-\frac{1}{k_n}} Q(t)dt &= \frac{k_n}{A_{k_n}} \sum_{m=1}^{n-2} e^{\beta m}(s^{m-1} - s^m) \\
&= e^{-\beta}(\frac{1}{e^{\beta}s})^{n-2} \sum_{m=1}^{n-2} e^{\beta}(1-s)(e^{\beta}s)^{m-1} \\
&= (1-s) \sum_{m=1}^{n-2} (e^{\beta}s)^{m+1-n} \\
&= (1-s)e^{-\beta}s^{-1} \frac{1-(e^{\beta}s)^{1-n}}{1-e^{-\beta}s^{-1}} \to \frac{1-s}{e^{\beta}s - 1}.
\end{aligned}$$

Hence, when $0 < \alpha < 1$, $\frac{k_n}{A_{k_n}} \int_{\frac{1}{k_n}}^{1-\frac{1}{k_n}} Q(t)dt \to \zeta = \frac{1-s}{e^{\beta}s-1}$. When $1 < \alpha < 2$, it can be shown that $\frac{k_n E(X) - B_{k_n}}{A_{k_n}}$ converges to $-\zeta$. Indeed, using the fact that $E(X) =$

$\int_0^1 Q(t)dt$, observe that

$$\begin{aligned}\frac{k_n EX - B_{k_n}}{A_{k_n}} &= \frac{k_n}{A_{k_n}} \int_0^{\frac{1}{k_n}} Q(t)dt + \frac{k_n}{A_{k_n}} \int_{1-\frac{1}{k_n}}^1 Q(t)dt \sim \frac{k_n}{A_{k_n}} \int_{1-\frac{1}{k_n}}^1 Q(t)dt \\ &= \frac{k_n}{A_{k_n}} \int_{1-s^{n-2}}^1 Q(t)dt = \frac{k_n}{A_{k_n}} \sum_{m=n-1}^{\infty} e^{\beta m}(s^{m-1} - s^m) \to \frac{1-s}{1-e^\beta s}.\end{aligned}$$

Hence, for $1 < \alpha < 2$, $\frac{k_n E(X) - B_{k_n}}{A_{k_n}} \to -\zeta = \frac{1-s}{1-e^\beta s}$.

For this example, the log of the characteristic function of the semi-stable distribution can be derived as follows. Note that the Lévy measure is characterized by the (distribution) function

$$R(x) = -e^{-\log(\frac{1}{s})(\lceil \frac{1}{\beta}\log x\rceil_+ - \frac{1}{\beta}\log x)} x^{-\frac{1}{\beta}\log\frac{1}{s}} = -e^{-\log(\frac{1}{s})(\lceil \frac{1}{\beta}\log x\rceil_+)} = -s^{\lceil \frac{1}{\beta}\log x\rceil_+}.$$

The function R is right-continuous with jumps at $e^{k\beta}$ taking values $-s^{k+1}$, where $k \in \mathbb{Z}$. Hence, by combining (2.1.2), (2.1.3) and (2.1.17), (2.1.18), (2.1.19), we get the log characteristic function as

$$\log \widehat{\mu}(t) = -i(\eta - \zeta)t + \sum_{k=-\infty}^{\infty} (e^{ite^{k\beta}} - 1 - \frac{e^{k\beta}}{1+e^{2k\beta}})(1-s)s^k, \tag{2.3.3}$$

where

$$\eta = \sum_{k=-\infty}^{0} \frac{e^{3k\beta}}{1+e^{2k\beta}} s^k(1-s) - \sum_{k=0}^{\infty} \frac{e^{k\beta}}{1+e^{2k\beta}} s^k(1-s). \tag{2.3.4}$$

We also note that the semi-stable distribution has a density of class C^∞ (the class of infinitely divisible functions). This follows from Proposition 28.3 in Sato [18], if for some $0 < \delta < 2$,

$$\liminf_{r \to 0} \frac{\int_{[0,r]} x^2 dR(x)}{r^{2-\delta}} > 0. \tag{2.3.5}$$

In the example considered here, the relation (2.3.5) holds with $\delta = \alpha$ since

$$\begin{aligned} \lim_{m \to -\infty} \frac{\int_{[0,e^{\beta m}]} x^2 dR(x)}{e^{\beta m(2-\alpha)}} &= \lim_{m \to -\infty} \frac{\sum_{k=-\infty}^{m} e^{2k\beta}(1-s)s^k}{e^{\beta m(2-\alpha)}} \\ &= \lim_{m \to -\infty} \frac{(1-s)s^m e^{\beta m \alpha}}{1-e^{-2\beta}s^{-1}} \to \frac{1-s}{1-e^{-2\beta}s^{-1}} \end{aligned}$$

since $se^{\beta\alpha} = 1$.

Using the characteristic function in (2.3.3) and the Abate and Whitt [1] method, we are going to compute numerically the density of the limiting semi-stable density for this example and also compare it to the histograms of the partial sums S_n. We fix $s = 0.25$ but vary the value of β such that $\alpha = \frac{1}{\beta} \log \frac{1}{s} < 2$. Figures 2.4–2.8 are associated with $\beta = 0.85, 0.9, 0.95, 1.05$ and 1.1, respectively. In each of these figures, four plots are presented corresponding to $n = 7, 8, 9, 10$ so that k_n takes values $1024, 4096, 16384$ and 65536. For each of these cases, we numerically calculate the density of the limiting semi-stable distribution (centered at 0; in red solid curve) and also plot the histogram of the empirical distribution of the partial sum S_n.

From the figures, the agreement between the numerically calculated density and the empirical histogram is very good for smaller values of $\beta = 0.85, 0.9$ and 0.95. As β increases and especially becomes larger than 1, the numerically calculated density deviates from the the empirical histogram. Note, however, that the agreement is quite good in the two tails in all the cases considered.

Finally, note that the approximation is expected to be worse as β increases. Indeed, as β becomes larger, the random variable $e^{\beta X}$ takes values separated by wider gaps. The partial sums of the variables $e^{\beta X_i}$ then should naturally require larger values of n and hence k_n to converge to the limiting semi-stable distribution.

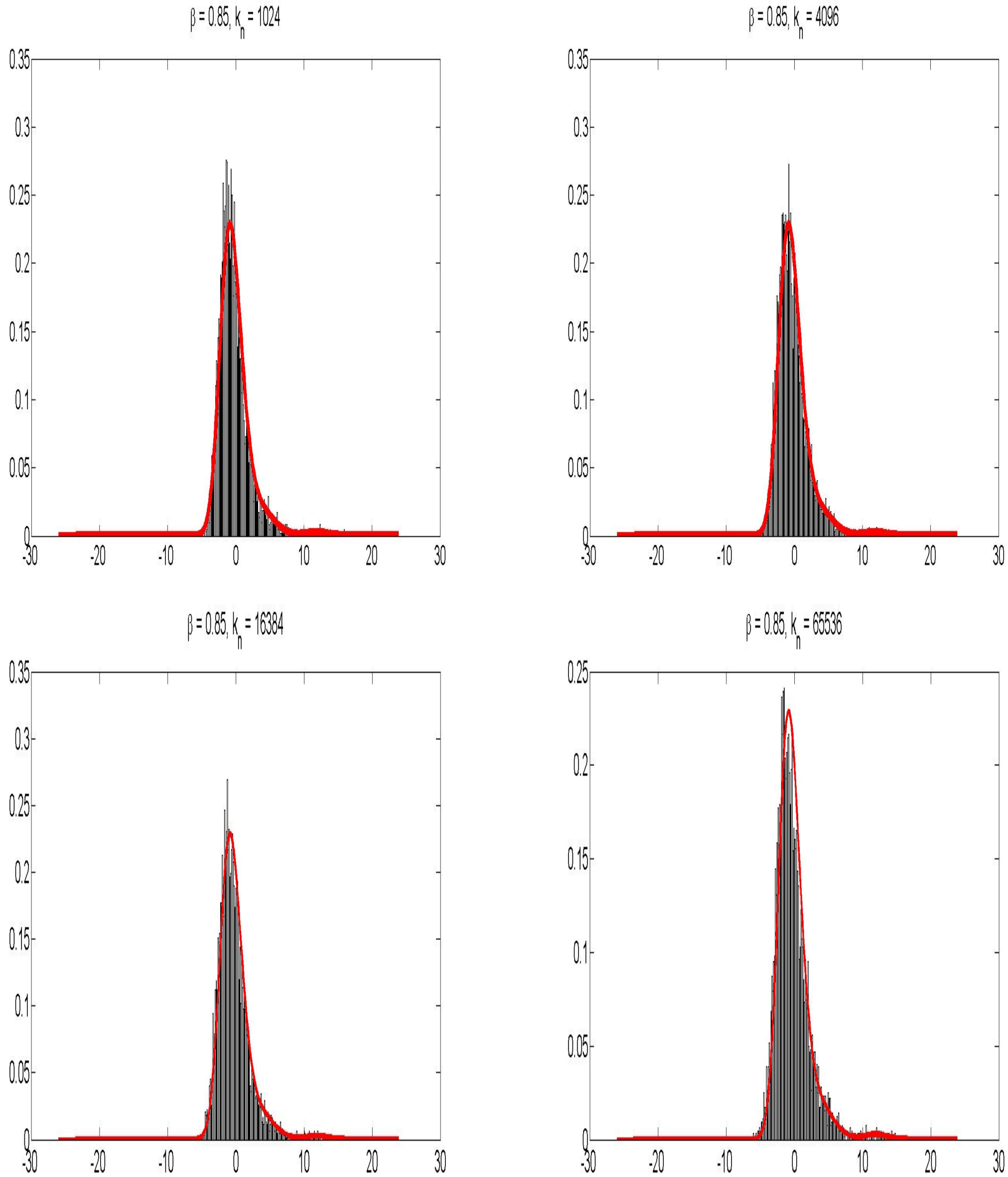

Figure 2.4: The empirical histogram against the actual density when $\beta = 0.85$.

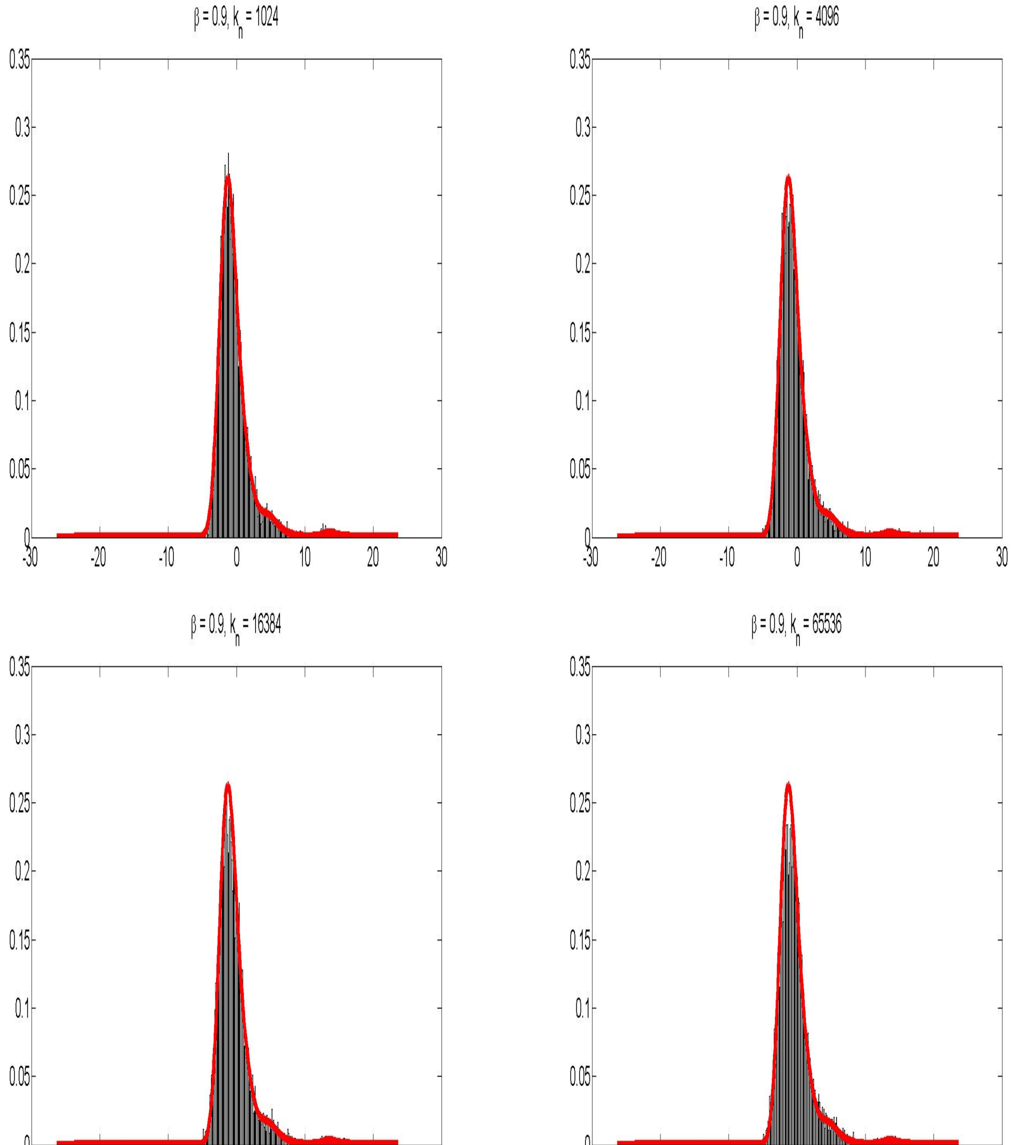

Figure 2.5: The empirical histogram against the actual density when $\beta = 0.9$.

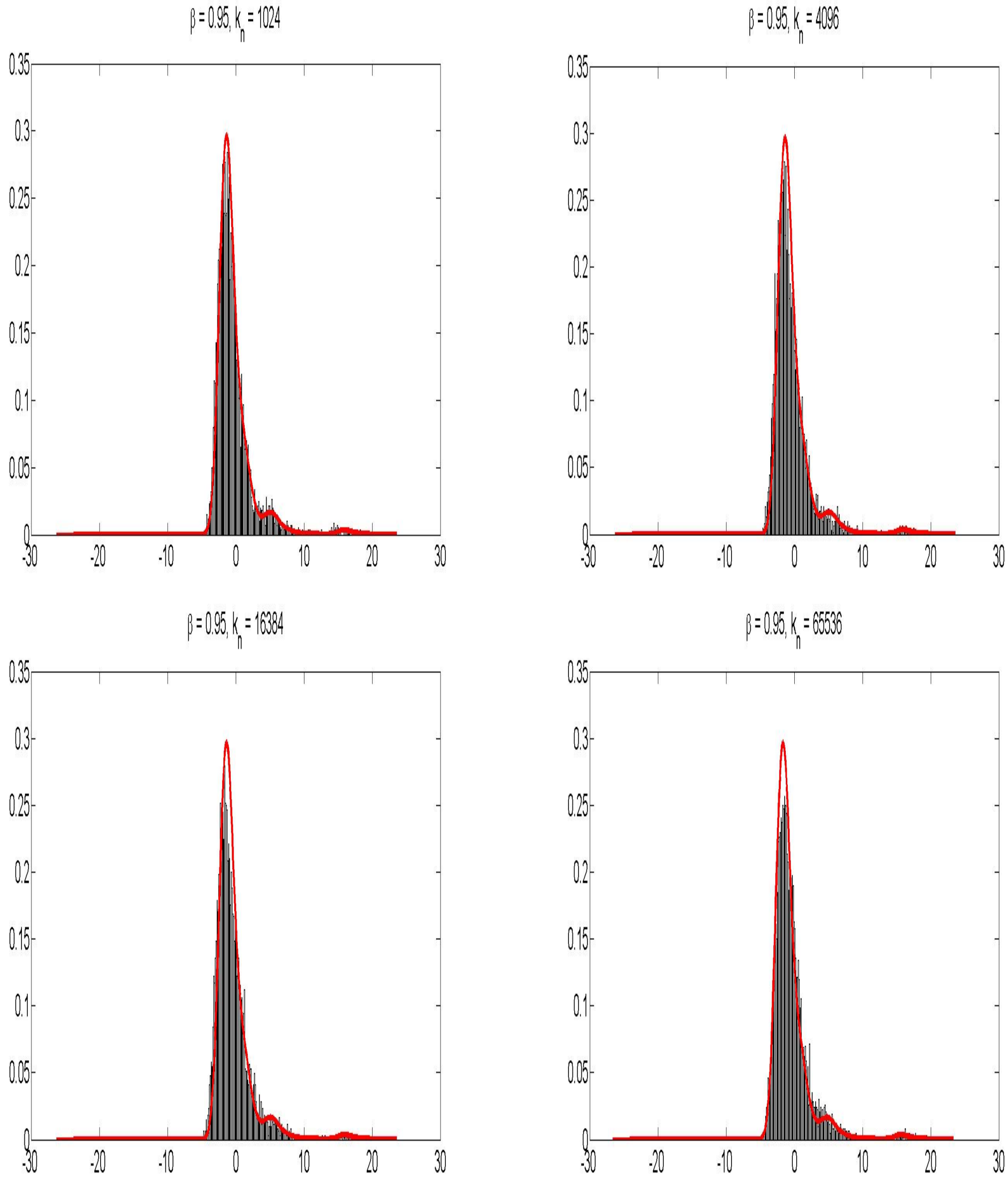

Figure 2.6: The empirical histogram against the actual density when $\beta = 0.95$.

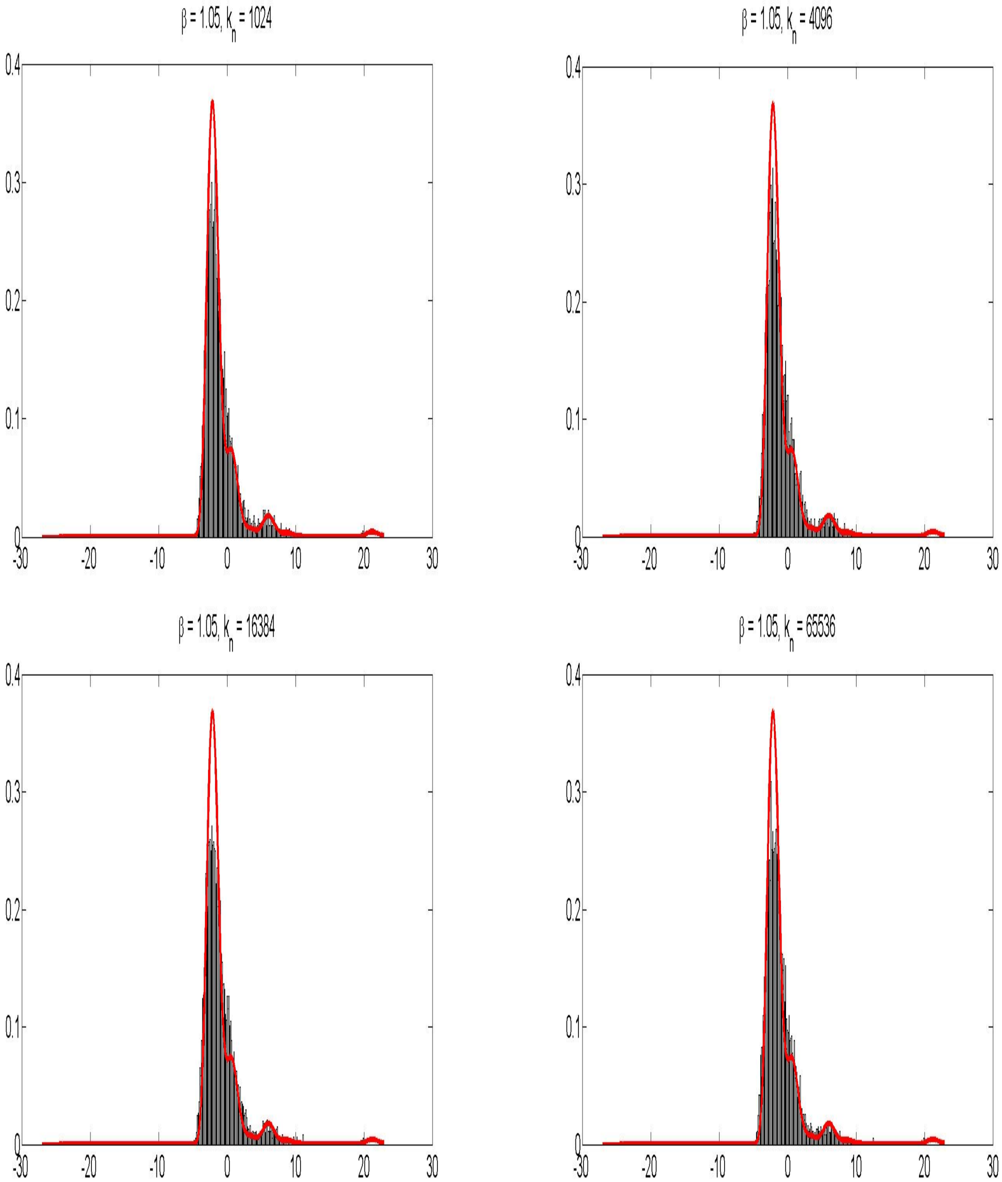

Figure 2.7: The empirical histogram against the actual density when $\beta = 1.05$.

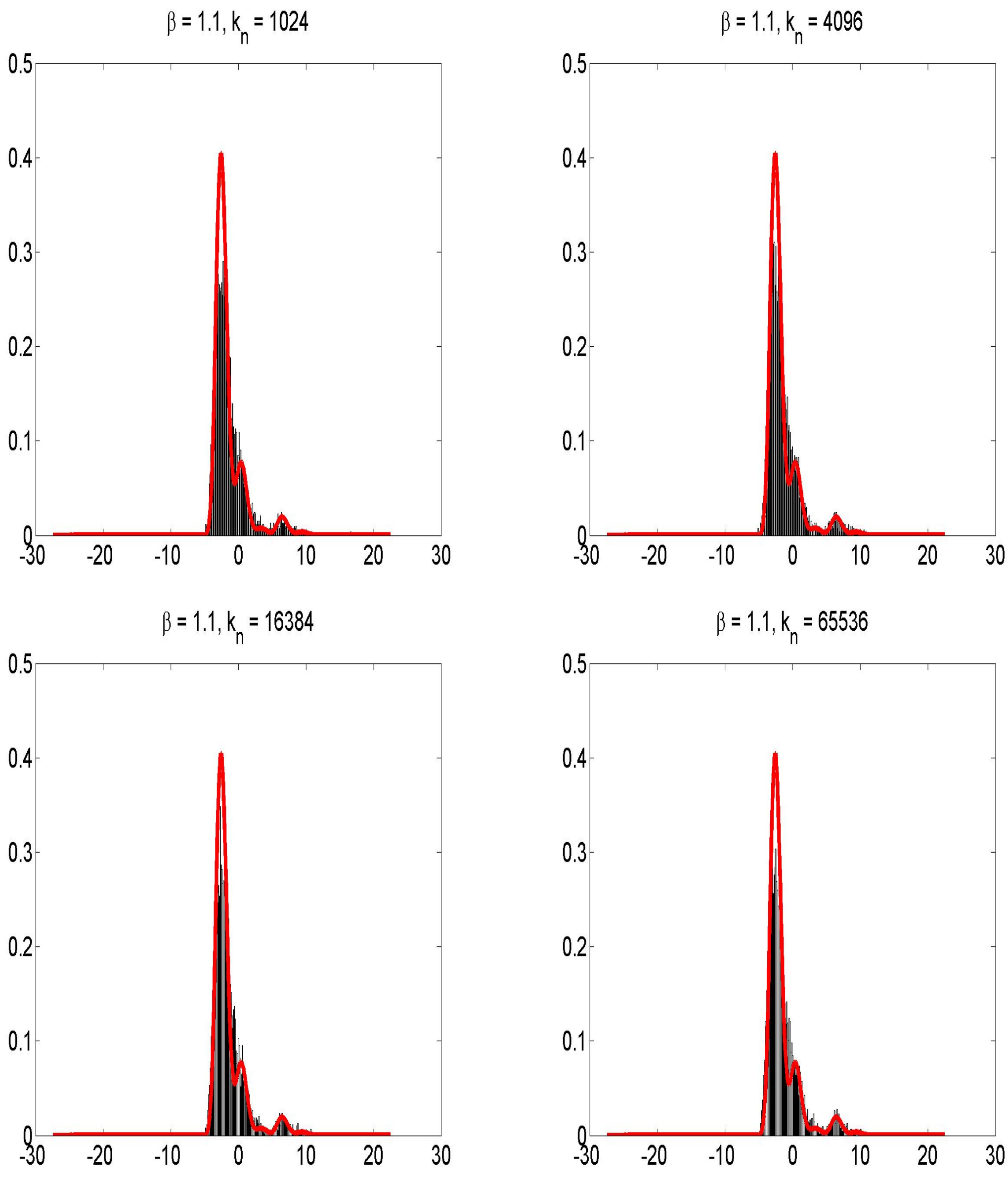

Figure 2.8: The empirical histogram against the actual density when $\beta = 1.1$.

CHAPTER 3: SEMI-STABLE DOMAINS OF ATTRACTION

This chapter contains our main results on the convergence to semi-stable distributions. In Section 3.1, we give a general practical result for a distribution function to be in the domain of attraction of semi-stable distribution. We explicitly derive the normalizing and centering constants, and the form of subsequence along which the convergence of partial sums to the semi-stable distribution takes place. We also consider the behavior of partial sums along all natural numbers instead of a suitable subsequence. In Section 3.2, we illustrate the results of Section 3.1 with concrete example. We also numerically calculate the density of the limiting semi-stable distribution for the example and assess how well it approximates the empirical histograms of partial sums in finite samples.

3.1 General results

The next theorem is the main result of this dissertation. We use the following notation throughout this dissertation:

$$\lceil x \rceil = \text{ the smallest integer larger than or equal to } x,$$

$$\lceil x \rceil_+ = \text{ the smallest integer strictly larger than } x.$$

For example, $\lceil 2.47 \rceil = \lceil 2.47 \rceil_+ = 3$ but $\lceil 3 \rceil = 3$ and $\lceil 3 \rceil_+ = 4$. The function $\lceil x \rceil_+$ is the right-continuous version of the function $\lceil x \rceil$. Also note that $\lceil x \rceil_+ = [x] + 1$, where $[x]$ is the integer part of x (that is, the largest integer smaller than or equal to x).

Theorem 3.1.1. *Let W_q be an integer-valued random variable taking values in $0, 1, 2, \ldots$ such that, for all $x > 0$,*

$$P\left(\frac{W_q}{2} \geq x, W_q \text{ is even}\right) = \sum_{n=\lceil x \rceil}^{\infty} P\left(\frac{W_q}{2} = n\right) = h_1(\lceil x \rceil) e^{-\nu \lceil x \rceil}, \tag{3.1.1}$$

$$P\left(\frac{W_q - 1}{2} \geq x, W_q \text{ is odd}\right) = \sum_{n=\lceil x \rceil}^{\infty} P\left(\frac{W_q - 1}{2} = n\right) = h_2(\lceil x \rceil) e^{-\nu \lceil x \rceil}, \tag{3.1.2}$$

where $\nu > 0$ and the functions h_1 and h_2 satisfy

$$\frac{h_2(x)}{h_1(x)} \to c_1, \text{ as } x \to \infty, \tag{3.1.3}$$

for some fixed $c_1 \geq 0$, and

$$\frac{h_1(ax)}{h_1(x)} \to 1 \text{ as } x \to \infty, a \to 1. \tag{3.1.4}$$

Let also

$$X = L(e^{W_q}) e^{\beta W_q} (-1)^{W_q}, \tag{3.1.5}$$

where $\beta > 0$ and L is a slowly varying function at ∞ such that $L(e^n)$ is ultimately monotonically increasing. Suppose that

$$\alpha := \frac{\nu}{2\beta} < 2. \tag{3.1.6}$$

Then, X is in the domain of attraction of a semi-stable distribution in the following sense. If $X, X_1, X_2, \ldots$ are i.i.d. random variables, then as $n \to \infty$, the partial sums

$$\frac{1}{A_{k_n}} \left\{ \sum_{j=1}^{k_n} X_j - B_{k_n} \right\} \tag{3.1.7}$$

converge to a semi-stable distribution with

$$k_n = \left\lceil \frac{e^{(n-1)\nu}}{h_1(n-1)} \right\rceil, \quad A_{k_n} = L(e^{2n-2})e^{2\beta(n-1)} \tag{3.1.8}$$

and B_{k_n} given by (2.1.15). The limiting semi-stable distribution is non-Gaussian, has location parameter given in (2.1.17) and is characterized by

$$\alpha = \frac{\nu}{2\beta}, \tag{3.1.9}$$

$$M_L(-x) = c_1 e^{-\nu([\frac{1}{2}+\frac{1}{2\beta}\log x]-\frac{1}{2\beta}\log x)}, \qquad M_R(x) = e^{-\nu(\lceil\frac{1}{2\beta}\log x\rceil_+ - \frac{1}{2\beta}\log x)}, \quad x > 0. \tag{3.1.10}$$

Before giving a proof of Theorem 3.1.1, let us explain its novelty, especially when compared to the available Theorems 2.1.1 and 2.1.2 on the domains of attraction of semi-stable distributions. Note that if L appearing in (3.1.5) is constant, then the convergence of the partial sums (3.1.7) to a semi-stable distribution can easily be deduced from Theorem 2.1.1. Taking $L(x) = 1$ without loss of generality,

$$\begin{aligned} \bar{F}(x) := 1 - F(x) &= P(e^{\beta W_q}(-1)^{W_q} > x) \\ &= P(\frac{W_q}{2} > \frac{1}{2\beta}\log x, \frac{W_q}{2} \text{ is an integer}) \\ &= h_1(\lceil\frac{1}{2\beta}\log x\rceil_+)e^{-\nu\lceil\frac{1}{2\beta}\log x\rceil_+} \\ &= x^{-\frac{\nu}{2\beta}} h_1(\lceil\frac{1}{2\beta}\log x\rceil_+)e^{-\nu(\lceil\frac{1}{2\beta}\log x\rceil_+ - \frac{1}{2\beta}\log x)} \\ &= x^{\frac{\nu}{2\beta}} l^*(x) M_R(\delta(x)), \end{aligned} \tag{3.1.11}$$

where $l^*(x) = h_1(\lceil\frac{1}{2\beta}\log x\rceil_+)$, $M_R(x)$ is given in (3.1.10) and $\delta(x) = \frac{x}{a(x)}$ with $a(x) =$

$e^{2\beta n}$ if $e^{2\beta n} \le x < e^{2\beta(n+1)}$. It can be shown that l^* is a slowly varying function, and hence that (3.1.11) is of the form (2.1.8). Similarly, it can be shown that the left tail of the distribution function of the random variable $e^{\beta W_q}(-1)^{W_q}$ can be expressed as (2.1.7).

Similarly, when the functions h_1 and h_2 appearing in (3.1.1) and (3.1.2) are constant, the convergence of the partial sums (3.1.7) to a semi-stable distribution can be deduced easily from Theorem 2.1.2. In this case, taking $h_1(x) \equiv h_2(x) \equiv 1$ for simplicity, consider first the quantile function Q corresponding to the right tail of F. We have

$$\begin{aligned} Q(1-s) &= \inf_y\{F(y) \ge 1-s\} \\ &= \inf_y\{P(L(e^{W_q})e^{\beta W_q}(-1)^{W_q} \le y, W_q \text{ is an even number}) \ge 1-s\} \\ &= \inf_y\{P(L(e^{2Z})e^{2\beta Z} \le y) \ge 1-s\}, \end{aligned} \tag{3.1.12}$$

where $Z = \frac{W_q}{2}$ and Z is an integer. Note from (3.1.1) that $P(Z \ge x) = e^{-\nu\lceil x\rceil}$ and hence $P(Z > x) = e^{-\nu\lceil x\rceil_+}$. Then, $P(Z \le x) = 1 - P(Z > x) = 1 - e^{-\nu\lceil x\rceil_+}$ and to solve (3.1.12), consider the equation

$$x = \inf_t\{1 - e^{-\nu\lceil t\rceil_+} \ge 1-s\}. \tag{3.1.13}$$

The inequality in (3.1.13) becomes

$$\lceil t\rceil_+ \ge \frac{1}{\nu}\log\frac{1}{s}, \tag{3.1.14}$$

leading to the solution

$$x = \lceil\frac{1}{\nu}\log\frac{1}{s}\rceil - 1. \tag{3.1.15}$$

Turning back to (3.1.12), we have $y = L(e^{2x})e^{2\beta x} = L(e^{-2}e^{\lceil \frac{1}{\nu}\log\frac{1}{s}\rceil})e^{-2\beta}e^{\lceil \frac{1}{\nu}\log\frac{1}{s}\rceil}$ which has the form (2.1.21). One can show similarly that the left tail of the quantile function Q is also of the form (2.1.20).

If one of the functions h_1 or h_2 is non-constant and L is also non-constant, then proving the convergence of the partial sums (3.1.7) to a semi-stable distribution is not so trivial. The main purpose of Theorem 3.1.1 is to show how this can be done using Theorem 2.1.1. We also characterize the normalizing constants A_{k_n}, centering constants B_{k_n} and the subsequence k_n along which the sequence of partial sums converges. We also present the explicit form of the log characteristic function for the limiting semi-stable distribution.

Proof of Theorem 3.1.1. The result will be proved by verifying the sufficient conditions (2.1.7)–(2.1.8) of Theorem 2.1.1. We break the proof into two cases dealing with (2.1.7) and (2.1.8) separately. The final part of the proof shows that the sequence k_n can be chosen as in (3.1.8)

Case 1 (showing (2.1.8)): Fix $x > 0$ large enough. In view of (3.1.5), we are interested in

$$\bar{F}(x) := 1 - F(x) = P\Big(L\big(e^{W_q}\big)e^{\beta W_q}(-1)^{W_q} > x\Big). \tag{3.1.16}$$

Let $Z_2 = \frac{W_q}{2}$. Note that (3.1.16) can be written as

$$\begin{aligned}\bar{F}(x) &= P\Big(L(e^{2Z_2})e^{2\beta Z_2} > x, Z_2 \text{ is integer}\Big) \\ &= P\Big(L(e^{2Z_2})e^{2\beta Z_2} > x\Big) \\ &= P\Big(Z_2 + \frac{1}{2\beta}\log L\big(e^{2Z_2}\big) > \frac{1}{2\beta}\log x\Big),\end{aligned} \tag{3.1.17}$$

where, in view of (3.1.1),

$$P(Z_2 \geq x) = h_1(\lceil x \rceil) e^{-\nu \lceil x \rceil}. \tag{3.1.18}$$

We next want to write $\bar{F}(x)$ in (3.1.17) as

$$\bar{F}(x) = P\left(Z_2 \geq g\left(\frac{1}{2\beta} \log x\right)\right) \tag{3.1.19}$$

for some function g.

There are many choices for g in (3.1.19). One natural choice is to take

$$g_0(y) = n, \text{ if } (n-1) + \frac{1}{2\beta} \log L\left(e^{2n-2}\right) \leq y < n + \frac{1}{2\beta} \log L\left(e^{2n}\right). \tag{3.1.20}$$

The function g_0, however, is not suitable for our purpose. We will use a function g_1 defined, for integer $n \geq 2$, as

$$g_1(y) = \begin{cases} n-1, & \text{if } n-1+\frac{1}{2\beta}\log L\left(e^{2n-2}\right) \leq y < n-1+\frac{1}{2\beta}\log L\left(e^{2n}\right), \\ y - \frac{1}{2\beta}\log L\left(e^{2n}\right), & \text{if } n-1+\frac{1}{2\beta}\log L\left(e^{2n}\right) \leq y < n+\frac{1}{2\beta}\log L\left(e^{2n}\right). \end{cases} \tag{3.1.21}$$

We will also use the function

$$g_2(y) = f^{-1}(y) = \inf\{z : f(z) \geq y\} \tag{3.1.22}$$

defined as an inverse of the function

$$f(z) = z + \frac{1}{2\beta} \log L\left(e^{2z}\right). \tag{3.1.23}$$

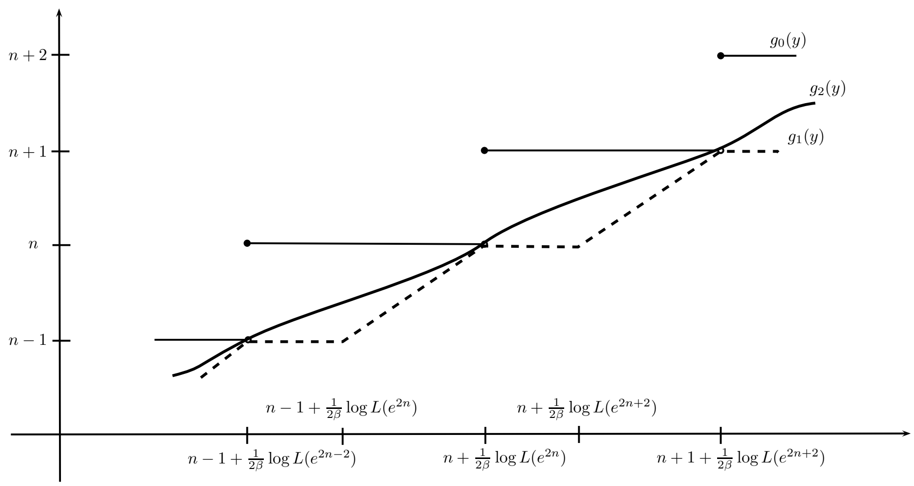

Figure 3.1: Plot of $g_0(y)$, $g_1(y)$ and $g_2(y)$.

Note that

$$\lceil g_0(y) \rceil = \lceil g_1(y) \rceil_+ = \lceil g_2(y) \rceil_+ = \lceil g(y) \rceil, \tag{3.1.24}$$

where g is any function satisfying (3.1.19). The functions g_0, g_1 and g_2 are plotted in Figure 3.1.

We shall use another function $\tilde{g}_1$ which modifies g_1 in the following way: for $n \geq 2$,

$$\tilde{g}_1(y) = y - \frac{1}{2\beta}\log L(e^{2n-2}), \quad \text{if } n-1+\frac{1}{2\beta}\log L(e^{2n-2}) \leq y < n + \frac{1}{2\beta}\log L(e^{2n}). \tag{3.1.25}$$

One relationship between the functions g_1 and $\tilde{g}_1$ can be found in Lemma A.1.1 in Appendix A.1, and will be used in the proof below. Note that $\tilde{g}_1(y)$ can be expressed as

$$\tilde{g}_1(y) = y - \tilde{g}_1^*(y), \tag{3.1.26}$$

where, for $n \geq 2$,

$$\tilde{g}_1^*(y) = \frac{1}{2\beta} \log L\left(e^{2n-2}\right), \qquad \text{if } n-1+\frac{1}{2\beta}\log L(e^{2n-2}) \leq y < n + \frac{1}{2\beta}\log L(e^{2n}). \tag{3.1.27}$$

See Lemma A.1.2 in Appendix A.1 for a property of $\tilde{g}_1^*$ which will be used in the proof below.

We need few properties of the function g_2. Since g_2 is the inverse of the function f, we have $e^{g_2(\log x)}$ as the inverse of $e^{f(\log x)}$. Indeed,

$$e^{g_2\left(\log e^{f(\log x)}\right)} = e^{g_2(f(\log x))} = e^{\log x} = x.$$

Note now from (3.1.23) that

$$e^{f(\log x)} = e^{\log x + \frac{1}{2\beta}\log L(x^2)} = x\left(L(x^2)\right)^{\frac{1}{2\beta}}.$$

Since $(L(x^2))^{\frac{1}{2\beta}}$ is a slowly varying function, $e^{f(\log x)}$ is a regularly varying function. So, by Theorem 1.5.13 of Bingham, Goldie and Teugels [3],

$$e^{g_2(\log x)} = xl(x),$$

where $l(x)$ is a slowly varying function. Hence,

$$g_2\left(\log x\right) = \log x + \log l(x) = \log x + g_2^*\left(\log x\right),$$

where

$$g_2^*\left(\log x\right) = \log l(x)$$

or replacing $\log x$ by y,

$$g_2(y) = y + g_2^*(y). \tag{3.1.28}$$

Note also that for any $A > 0$, we have

$$\begin{aligned} g_2^*(\log Ax) - g_2^*(\log x) &= \log l(Ax) - \log l(x) \\ &= \log \frac{l(Ax)}{l(x)} \to 0, \text{ as } x \to \infty. \end{aligned} \tag{3.1.29}$$

Continuing with (3.1.19) now, note that, by using (3.1.18) and (3.1.24),

$$\begin{aligned} \bar{F}(x) &= P\left(Z_2 \geq g\left(\frac{1}{2\beta}\log x\right)\right) \\ &= h_1\left(\lceil g(\frac{1}{2\beta}\log x)\rceil\right)e^{-\nu\lceil g(\frac{1}{2\beta}\log x)\rceil} \\ &= h_1\left(\lceil g_2(\frac{1}{2\beta}\log x)\rceil_+\right)e^{-\nu\lceil g_1(\frac{1}{2\beta}\log x)\rceil_+}. \end{aligned} \tag{3.1.30}$$

By using (3.1.26), note further that

$$\begin{aligned} \bar{F}(x) &= h_1\left(\lceil g_2(\frac{1}{2\beta}\log x)\rceil_+\right)e^{-\nu\tilde{g}_1(\frac{1}{2\beta}\log x)} \cdot \\ &\qquad e^{-\nu(g_1(\frac{1}{2\beta}\log x)-\tilde{g}_1(\frac{1}{2\beta}\log x))}e^{-\nu(\lceil g_1(\frac{1}{2\beta}\log x)\rceil_+ - g_1(\frac{1}{2\beta}\log x))} \\ &= h_1\left(\lceil g_2(\frac{1}{2\beta}\log x)\rceil_+\right)e^{-\nu(\frac{1}{2\beta}\log x - \tilde{g}_1^*(\frac{1}{2\beta}\log x))} \cdot \\ &\qquad e^{-\nu(g_1(\frac{1}{2\beta}\log x)-\tilde{g}_1(\frac{1}{2\beta}\log x))}e^{-\nu(\lceil g_1(\frac{1}{2\beta}\log x)\rceil_+ - g_1(\frac{1}{2\beta}\log x))} \\ &= x^{-\alpha}l_1^*(x)(M_R(\delta(x)) + h_R(x)), \end{aligned} \tag{3.1.31}$$

where $\alpha = \frac{\nu}{2\beta}$ as given in (3.1.81),

$$l_1^*(x) = h_1\left(\lceil g_2(\frac{1}{2\beta}\log x)\rceil_+\right)e^{\nu\tilde{g}_1^*(\frac{1}{2\beta}\log x)}e^{-\nu(g_1(\frac{1}{2\beta}\log x)-\tilde{g}_1(\frac{1}{2\beta}\log x))}, \tag{3.1.32}$$

$$M_R(\delta(x)) = e^{-\nu(\lceil\tilde{g}_1(\frac{1}{2\beta}\log x)\rceil_+ - \tilde{g}_1(\frac{1}{2\beta}\log x))} \tag{3.1.33}$$

and

$$h_R(x) = e^{-\nu(\lceil g_1(\frac{1}{2\beta}\log x)\rceil_+ - g_1(\frac{1}{2\beta}\log x))} - e^{-\nu(\lceil\tilde{g}_1(\frac{1}{2\beta}\log x)\rceil_+ - \tilde{g}_1(\frac{1}{2\beta}\log x))}. \tag{3.1.34}$$

We next show that the functions l_1^*, M_R and h_R satisfy the conditions of Theorem 2.1.1 with suitable choices of $\delta(x)$ and A_{k_n}.

By Lemma A.1.3 in Appendix A.1, $l_1^*(x)$ is a right-continuous slowly varying function and hence it satisfies the conditions of Theorem 2.1.1. For the function $M_R(\delta(x))$, note from (3.1.33) that

$$\begin{aligned} M_R(\delta(x)) &= e^{-\nu\left(\lceil\frac{2\beta\tilde{g}_1(\frac{1}{2\beta}\log x)}{2\beta}\rceil_+ - \frac{2\beta\tilde{g}_1(\frac{1}{2\beta}\log x)}{2\beta}\right)} \\ &= M_R\left(e^{2\beta\tilde{g}_1(\frac{1}{2\beta}\log x)}\right) \end{aligned} \tag{3.1.35}$$

with

$$M_R(x) = e^{-\nu(\lceil\frac{\log x}{2\beta}\rceil_+ - \frac{\log x}{2\beta})}. \tag{3.1.36}$$

The function $M_R(x)$ is periodic with multiplicative period $e^{2\beta}$, and is right-continuous as required in Theorem 2.1.1. Since the period $e^{2\beta}$ is also $c^{\frac{1}{\alpha}}$, this yields

$$c = e^{\nu}. \tag{3.1.37}$$

To choose $\delta(x)$, note from (3.1.35) that

$$M_R(\delta(x)) = M_R\left(e^{2\beta\tilde{g}_1(\frac{1}{2\beta}\log x)-2\beta(n-1)}\right),$$

for any $n \geq 1$, since M_R has multiplicative period $e^{2\beta}$. We can set

$$\delta(x) = e^{2\beta\tilde{g}_1(\frac{1}{2\beta}\log x)-2\beta(n-1)}, \quad \text{if } e^{2\beta(n-1)}L(e^{2n-2}) \leq x < e^{2n\beta}L(e^{2n}). \tag{3.1.38}$$

From (3.1.25), we have

$$\begin{aligned} \delta(x) &= e^{2\beta(\frac{1}{2\beta}\log x - \frac{1}{2\beta}\log L(e^{2n-2}))-2\beta(n-1)} \\ &= \frac{x}{e^{2\beta(n-1)}L(e^{2n-2})}, \quad \text{if } e^{2\beta(n-1)}L(e^{2n-2}) \leq x < e^{2n\beta}L(e^{2n}). \end{aligned} \tag{3.1.39}$$

Thus, $\delta(x)$ has the required form (2.1.10)–(2.1.11) with

$$A_{k_n} = e^{2\beta(n-1)}L(e^{2n-2}) \tag{3.1.40}$$

and

$$a(x) = e^{2\beta(n-1)}L(e^{2n-2}) = A_{k_n}, \quad \text{if } A_{k_n} \leq x < A_{k_{n+1}}. \tag{3.1.41}$$

Note also from (3.1.39) that

$$1 \leq \delta(x) < \frac{e^{2\beta n}L(e^{2n})}{e^{2\beta(n-1)}L(e^{2n-2})} = e^{2\beta}\frac{L(e^{2n})}{L(e^{-2}e^{2n})} \to e^{2\beta} = c^{\frac{1}{\alpha}},$$

so that $\delta(x) \in [1, c^{\frac{1}{\alpha}} + \epsilon]$ for large enough x when $\epsilon > 0$ is fixed.

To complete Case 1, we need to prove that $h_R(A_{k_n}x_0) \to 0$ as $n \to \infty$ for every continuity point x_0 of $M_R(x)$. The discontinuity points of M_R are

$$x = e^{2k\beta}, \quad k \in \mathbb{Z}. \tag{3.1.42}$$

To show $h_R(A_{k_n}x_0) \to 0$, note that, by Lemma A.1.1, it is enough to prove that $\tilde{h}_R(A_{k_n}x_0) \neq 0$ for finitely many values of n, where

$$\tilde{h}_R(x) = e^{-\nu\lceil g_1(\frac{1}{2\beta}\log x)\rceil_+} - e^{-\nu\lceil \tilde{g}_1(\frac{1}{2\beta}\log x)\rceil_+}.$$

This holds only if for some integer $m \geq 2$,

$$m + \log L(e^{2m-2}) \leq \frac{1}{2\beta}\log A_{k_n}x_0 < m + \log L(e^{2m}). \tag{3.1.43}$$

By Lemma A.1.4, (3.1.43) holds for infinitely many values of n only if $x_0 = e^{2r\beta}$, $r \in \mathbb{Z}$, which is a discontinuity point of $M_R(x)$ in (3.1.42). Hence, $h_R(A_{k_n}x_0) \to 0$ as $n \to \infty$ for every continuity point x_0 of $M_R(x)$.

<u>Case 2 (showing (2.1.7)):</u> In view of (3.1.5), we are now interested in

$$F_-(-x) \;=\; P\Big(L\big(e^{W_q}\big)e^{\beta W_q}(-1)^{W_q} < -x\Big). \tag{3.1.44}$$

Let $Z_2 = \frac{W_q}{2}$ as in Case 1. Note that (3.1.44) can be written as

$$\begin{aligned} F_{\text{-}}(-x) &= P\Big(L\big(e^{2Z_2}\big)e^{2\beta Z_2} > x, Z_2 - \frac{1}{2} \text{ is integer}\Big) \\ &= P\Big(L\big(ee^{2(Z_2-\frac{1}{2})}\big)e^{\beta}e^{2\beta(Z_2-\frac{1}{2})} > x, Z_2 - \frac{1}{2} \text{ is integer}\Big) \\ &= P\Big(L(ee^{2Z_1})e^{\beta}e^{2\beta Z_1} > x\Big) \\ &= P\Big(Z_1 + \frac{1}{2} + \frac{1}{2\beta}\log L(e^{2Z_1+1}) > \frac{1}{2\beta}\log x\Big), \end{aligned} \tag{3.1.45}$$

where, in view of (3.1.2),

$$P(Z_1 \geq x) = h_2(\lceil x \rceil)e^{-\nu\lceil x \rceil}. \tag{3.1.46}$$

Writing (3.1.45) as

$$F_{\text{-}}(-x) = P\Big(Z_1 + \frac{1}{2\beta}\log L(e^{2Z_1+1}) > \frac{1}{2\beta}\log x - \frac{1}{2}\Big),$$

the right-hand side has the form (3.1.16) where $L(e^{2Z_2})$ is replaced by $L(ee^{2Z_1})$ and $\frac{1}{2\beta}\log x$ is replaced by $\frac{1}{2\beta}\log x - \frac{1}{2}$. Thus, as in (3.1.19)–(3.1.20), one can write

$$F_{\text{-}}(-x) = P\Big(Z_1 \geq \tilde{g}(\frac{1}{2\beta}\log x - \frac{1}{2})\Big), \tag{3.1.47}$$

where

$$\tilde{g}(y) = n, \text{ if } n - 1 + \frac{1}{2\beta}\log L(ee^{2n-2}) \leq y < n + \frac{1}{2\beta}\log L(ee^{2n}). \tag{3.1.48}$$

The expression (3.1.47) can also be written as

$$F_{\text{-}}(-x) = P\left(Z_1 \geq \tilde{g}_0(\frac{1}{2\beta}\log x)\right), \tag{3.1.49}$$

where $\tilde{g}_0(y) = \tilde{g}(y - \frac{1}{2})$ or, for $n \geq 2$,

$$\tilde{g}_0(y) = n, \quad \text{ if } n - \frac{1}{2} + \frac{1}{2\beta}\log L(e^{2n-1}) \leq y < n + \frac{1}{2} + \frac{1}{2\beta}\log L(e^{2n+1}). \tag{3.1.50}$$

We want to work with the intervals $[n - 1 + \frac{1}{2\beta}\log L(e^{2n-2}), n + \frac{1}{2\beta}\log L(e^{2n}))$ appearing in Case 1, and use the results of that case. Note that, on the interval $[n - 1 + \frac{1}{2\beta}\log L(e^{2n-2}), n + \frac{1}{2\beta}\log L(e^{2n}))$, the function $\tilde{g}_0$ takes values as

$$\tilde{g}_0(y) = \begin{cases} n-1, & \text{if } n - 1 + \frac{1}{2\beta}\log L(e^{2n-2}) \leq y < n - \frac{1}{2} + \frac{1}{2\beta}\log L(e^{2n-1}), \\ n, & \text{if } n - \frac{1}{2} + \frac{1}{2\beta}\log L(e^{2n-1}) \leq y < n + \frac{1}{2\beta}\log L(e^{2n}). \end{cases} \tag{3.1.51}$$

Defining

$$I_0(y) = \begin{cases} -1, & \text{if } n - 1 + \frac{1}{2\beta}\log L(e^{2n-2}) \leq y < n - \frac{1}{2} + \frac{1}{2\beta}\log L(e^{2n-1}), \\ 0, & \text{if } n - \frac{1}{2} + \frac{1}{2\beta}\log L(e^{2n-1}) \leq y < n + \frac{1}{2\beta}\log L(e^{2n}), \end{cases} \tag{3.1.52}$$

and combing (3.1.20), (3.1.51) and (3.1.52), we have

$$\tilde{g}_0(y) = g_0(y) + I_0(y). \tag{3.1.53}$$

Continuing with (3.1.49), note further that, by using (3.1.46) and (3.1.53),

$$\begin{aligned} F_{\text{-}}(-x) &= h_2\Big(\tilde{g}_0(\frac{1}{2\beta}\log x)\Big)e^{-\nu\tilde{g}_0(\frac{1}{2\beta}\log x)} \\ &= e^{-\nu I_0(\frac{1}{2\beta}\log\, x)}h_2\Big(g_0(\frac{1}{2\beta}\log\, x)+I_0(\frac{1}{2\beta}\log\, x)\Big)e^{-\nu g_0(\frac{1}{2\beta}\log\, x)}. \end{aligned} \tag{3.1.54}$$

We want to write $F_{\text{-}}(-x)$ as in (2.1.7) of Theorem 2.1.1 (where by Lemma A.1.5, we can take a slowly varying function l_2^* which is asymptotically equivalent to l_1^*). We need the notation for the intervals appearing in (3.1.51)–(3.1.52), namely, for $n \geq 1$,

$$\begin{aligned} D_n &= [n-1+\frac{1}{2\beta}\log\, L(e^{2n-2}), n-\frac{1}{2}+\frac{1}{2\beta}\log\, L(e^{2n-1})), \\ E_n &= [n-\frac{1}{2}+\frac{1}{2\beta}\log\, L(e^{2n-1}), n+\frac{1}{2\beta}\log\, L(e^{2n})). \end{aligned}$$

We also need a similar notation without the slowly varying function L, that is, for $n \geq 1$,

$$D_n^{'} = [n-1, n-\frac{1}{2}), \qquad E_n^{'} = [n-\frac{1}{2}, n).$$

Set also

$$D = \bigcup_{n=1}^{\infty} D_n, \quad E = \bigcup_{n=1}^{\infty} E_n, \quad D^{'} = \bigcup_{n=1}^{\infty} D_n^{'}, \quad E^{'} = \bigcup_{n=1}^{\infty} E_n^{'}. \tag{3.1.55}$$

As in (3.1.31), we can now write (3.1.54) as

$$F_{\text{-}}(-x) \;=\; x^{-\alpha}\frac{h_2(g_0(\frac{1}{2\beta}\log\, x)+I_0(\frac{1}{2\beta}\log\, x))}{c_1h_1(g_0(\frac{1}{2\beta}\log\, x))}l_1^*(x)c_1e^{-\nu I_0(\frac{1}{2\beta}\log x)}e^{-\nu(\lceil g_1(\frac{1}{2\beta}\log\, x)\rceil_+ - g_1(\frac{1}{2\beta}\log\, x))},$$

where $\alpha = \frac{\nu}{2\beta}$ and $l_1^*(x)$ is given in (3.1.32). This can also be written as

$$F_-(-x) = x^{-\alpha} l_2^*(x)(M_L(-\delta(x)) + h_L(x)),$$

where

$$l_2^*(x) = \frac{h_2(g_0(\frac{1}{2\beta}\log x) + I_0(\frac{1}{2\beta}\log x))}{c_1 h_1(g_0(\frac{1}{2\beta}\log x))} l_1^*(x), \tag{3.1.56}$$

$$M_L(-\delta(x)) = c_1 e^{-\nu([\frac{1}{2}+\tilde{g}_1(\frac{1}{2\beta}\log x)]-\tilde{g}_1(\frac{1}{2\beta}\log x))}, \tag{3.1.57}$$

$$h_L(x) = c_1 e^{-\nu I_0(\frac{1}{2\beta}\log x)} e^{-\nu(\lceil g_1(\frac{1}{2\beta}\log x)\rceil_+ - g_1(\frac{1}{2\beta}\log x))} - c_1 e^{-\nu([\frac{1}{2}+\tilde{g}_1(\frac{1}{2\beta}\log x)]-\tilde{g}_1(\frac{1}{2\beta}\log x))}. \tag{3.1.58}$$

By using (3.1.3)–(3.1.4), we have

$$\frac{h_2(g_0(\frac{1}{2\beta}\log x) + I_0(\frac{1}{2\beta}\log x))}{c_1 h_1(g_0(\frac{1}{2\beta}\log x))} \to 1, \text{ as } x \to \infty.$$

Hence, $\frac{l_2^*(x)}{l_1^*(x)} \to 1$, as $x \to \infty$, that is, $l_2^*(x)$ and $l_1^*(x)$ are two asymptotically equivalent functions. By the definition of I_0 and using Lemma A.1.3, $l_2^*(x)$ is right-continuous and slowly varying.

The function $\delta(x)$ appearing in (3.1.57) is the same as in (3.1.38)–(3.1.39) of Case 1, while the function $M_L(-x)$ is defined as

$$M_L(-x) = c_1 e^{-\nu([\frac{1}{2}+\frac{1}{2\beta}\log x]-\frac{1}{2\beta}\log x)}, \quad x > 0. \tag{3.1.59}$$

It is left-continuous when $x > 0$, and also periodic with multiplicative period $e^{2\beta} = c^{\frac{1}{\alpha}}$. Thus, $M_L(x)$ for $x < 0$ is left-continuous as required in Theorem 2.1.1. The discontinuity points of $M_L(-x)$ are

$$x = e^{\beta(2k+1)}, \quad k \in \mathbb{Z}. \tag{3.1.60}$$

To conclude the proof of Case 2, we need to show that $h_L(A_{k_n}x_0) \to 0$ as $n \to \infty$ for every continuity point x_0 of $M_L(-x)$, that is, x_0 different from (3.1.60). For this, we rewrite $h_L(x)$ as follows. Observe that

$$e^{-\nu I_0(y)} = e^{\nu} 1_D(y) + 1_E(y)$$

and

$$(e^{\nu} 1_{D'}(y) + 1_{E'}(y)) e^{-\nu(\lceil y \rceil_+ - y)} = e^{-\nu([\frac{1}{2}+y]-y)},$$

where after taking the logs, using $\lceil y \rceil_+ = [y]+1$ and simplification, the last identity is equivalent to $[y]1_{D'}(y) + ([y]+1)1_{E'}(y) = [\frac{1}{2}+y]$ and can be seen easily by drawing a picture. By using these identities and (3.1.58), we can write

$$\begin{aligned} c_1^{-1} h_L(x) &= (e^{\nu} 1_D(\frac{1}{2\beta}\log x) + 1_E(\frac{1}{2\beta}\log x)) e^{-\nu(\lceil g_1(\frac{1}{2\beta}\log x)\rceil_+ - g_1(\frac{1}{2\beta}\log x))} \\ &\qquad - e^{-\nu([\frac{1}{2}+\tilde{g}_1(\frac{1}{2\beta}\log x)] - \tilde{g}_1(\frac{1}{2\beta}\log x))} \\ &= h_{1,L}(x) e^{-\nu(\lceil g_1(\frac{1}{2\beta}\log x)\rceil_+ - g_1(\frac{1}{2\beta}\log x))} + h_{2,L}(x), \end{aligned}$$

where

$$h_{1,L}(x) = e^{\nu} 1_D(\frac{1}{2\beta}\log x) + 1_E(\frac{1}{2\beta}\log x) - e^{\nu} 1_{D'}(g_1(\frac{1}{2\beta}\log x)) - 1_{E'}(g_1(\frac{1}{2\beta}\log x)),$$

$$h_{2,L}(x) = e^{-\nu([\frac{1}{2}+g_1(\frac{1}{2\beta}\log x)] - g_1(\frac{1}{2\beta}\log x))} - e^{-\nu([\frac{1}{2}+\tilde{g}_1(\frac{1}{2\beta}\log x)] - \tilde{g}_1(\frac{1}{2\beta}\log x))}.$$

It is therefore enough to show that $h_{1,L}(A_{k_n}x_0) \to 0$ and $h_{2,L}(A_{k_n}x_0) \to 0$, as $n \to \infty$. From (3.1.21), (3.1.25) and (3.1.55), $h_{1,L}(A_{k_n}x_0) \neq 0$ if, for some integer $m \geq 1$,

$$m - \frac{1}{2} + \log L(e^{2m-1}) \leq \frac{1}{2\beta} \log A_{k_n}x_0 < m - \frac{1}{2} + \log L(e^{2m}). \tag{3.1.61}$$

(To see this, partition $[m - 1 + \frac{1}{2\beta}\log L(e^{2m-2}), m + \frac{1}{2\beta}\log L(e^{2m}))$ into four subintervals $[m - 1 + \frac{1}{2\beta}\log L(e^{2m-2}), m - 1 + \frac{1}{2\beta}\log L(e^{2m}))$, $[m - 1 + \frac{1}{2\beta}\log L(e^{2m}), m - \frac{1}{2} + \frac{1}{2\beta}\log L(e^{2m-1}))$, $[m - \frac{1}{2} + \frac{1}{2\beta}\log L(e^{2m-1}), m - \frac{1}{2} + \frac{1}{2\beta}\log L(e^{2m}))$, $[m - \frac{1}{2} + \frac{1}{2\beta}\log L(e^{2m}), m + \frac{1}{2\beta}\log L(e^{2m}))$ and check that the function is nonzero only on the third subinterval as given in (3.1.61).) By Lemma A.1.4, (3.1.61) holds for infinitely many values of n only if $x_0 = e^{\beta(2r+1)}$ which is a discontinuity point of $M_L(-x)$ in (3.1.60). To show $h_{2,L}(A_{k_n}x_0) \to 0$, note that, by Lemma A.1.1, it is enough to prove that $\tilde{h}_{2,L}(A_{k_n}x_0) \neq 0$ for finitely many values of n, where

$$\tilde{h}_{2,L}(x) = e^{-\nu[\frac{1}{2} + g_1(\frac{1}{2\beta}\log x)]} - e^{-\nu[\frac{1}{2} + \tilde{g}_1(\frac{1}{2\beta}\log x)]}.$$

By using (3.1.21) and (3.1.25), the relation $\tilde{h}_{2,L}(A_{k_n}x_0) = 0$ holds only if, for some integer $m \geq 1$,

$$m - \frac{1}{2} + \log L(e^{2m-2}) \leq \frac{1}{2\beta} \log A_{k_n}x_0 < m - \frac{1}{2} + \log L(e^{2m}). \tag{3.1.62}$$

(To see this, draw a plot of $g_1(y)$ and $\tilde{g}_1(y)$ for y in $[m - 1 + \frac{1}{2\beta}\log L(e^{2m-2}), m - \frac{1}{2}\log L(e^{2m}))$, and note that $\tilde{g}_1(y) = m - \frac{1}{2}$ at $y = m - \frac{1}{2} + \frac{1}{2\beta}\log L(e^{2m-2})$ and $g_1(y) = m - \frac{1}{2} + \frac{1}{2\beta}\log L(e^{2m})$.) By Lemma A.1.4, (3.1.62) holds for infinitely many values of n only if $x_0 = e^{\beta(2r+1)}$ which is a discontinuity point of $M_L(-x)$ in (3.1.60). Hence, $h_L(A_{k_n}x_0) \to 0$ as $n \to \infty$ for every continuity point x_0 of $M_L(-x)$.

Deriving subsequence k_n: We conclude the proof of the theorem by showing that k_n is given by (3.1.8). In view of the discussion following Theorem 2.1.1, we want to choose a sequence $\tilde{A}_n$ satisfying (2.1.12)–(2.1.13) such that k_n given by (3.1.8) now satisfies (2.1.14). We define such sequence $\tilde{A}_n$ as

$$\begin{aligned} \log \tilde{A}_n &= 2\beta(m-1) + \log L(e^{2m-2}) \\ &+ \frac{(\log n - \log k_m)(2\beta + \log L(e^{2m}) - \log L(e^{2m-2}))}{\log k_{m+1} - \log k_m} \\ &\qquad \text{if } k_m \le n < k_{m+1}, \quad m \ge 1. \end{aligned} \tag{3.1.63}$$

The sequence $\tilde{A}_n$ satisfies (2.1.13). For example, if $k_m \le n < k_{m+1} - 1$, the last limit in (2.1.13) follows from

$$\log \tilde{A}_{n+1} - \log \tilde{A}_n = \frac{(\log n - \log (n+1))(2\beta + \log L(e^{2m}) - \log L(e^{2m-2}))}{\log k_{m+1} - \log k_m} \to 0.$$

If $n = k_{m+1} - 1$, the limit follows from

$$\begin{aligned} \log \tilde{A}_{n+1} - \log \tilde{A}_n &= 2\beta + \log L(e^{2m}) - \log L(e^{2m-2}) \\ &- \frac{(\log (k_{m+1} - 1) - \log k_m)(2\beta + \log L(e^{2m}) - \log L(e^{2m-2}))}{\log k_{m+1} - \log k_m} \to 0 \end{aligned}$$

since $\log L(e^{2m}) - \log L(e^{2m-2}) \to 0$, and

$$\frac{\log (k_{m+1} - 1) - \log k_m}{\log k_{m+1} - \log k_m} \to 1.$$

Next we show (2.1.12), that is, $n\tilde{A}_n^{-\alpha} l_1^*(\tilde{A}_n) \to 1$, as $n \to \infty$, where $\alpha = \frac{\nu}{2\beta}$ and l_1^*

is as defined in (3.1.32). When $k_m \le n < k_{m+1}$, observe that

$$\begin{aligned}
& \log n\tilde{A}_n^{-\alpha} l_1^*(\tilde{A}_n) = \log \frac{n l_1^*(\tilde{A}_n)}{\tilde{A}_n^{\nu/2\beta}} \\
= \;& \log \frac{n l_1^*(\tilde{A}_n)}{e^{(m-1)\nu} L(e^{2m-2})^{\nu/2\beta}} + \frac{\nu + \frac{\nu}{2\beta}\log\, L(e^{2m}) - \frac{\nu}{2\beta}\log\, L(e^{2m-2})}{\log\, k_{m+1} - \log\, k_m} \log\left(\frac{k_m}{n}\right) \\
\sim \;& \log n + \log \frac{l_1^*(\tilde{A}_n)}{h_1(m-1) L(e^{2m-2})^{\nu/2\beta}} - \log k_m \\
& + \frac{\nu + \frac{\nu}{2\beta}\log\, L(e^{2m}) - \frac{\nu}{2\beta}\log\, L(e^{2m-2})}{\log\, k_{m+1} - \log\, k_m} \log\left(\frac{k_m}{n}\right). \qquad (3.1.64)
\end{aligned}$$

Now observe that as $n \to \infty$, we have $m \to \infty$, and thus $\frac{k_m}{n}$ is bounded and

$$\frac{\nu + \frac{\nu}{2\beta}\log\, L(e^{2m}) - \frac{\nu}{2\beta}\log\, L(e^{2m-2})}{\log\, k_{m+1} - \log\, k_m} \to 1.$$

Thus, (3.1.64) is asymptotically equivalent to

$$\log \frac{l_1^*(\tilde{A}_n)}{h_1(m-1) L(e^{2m-2})^{\nu/2\beta}}. \qquad (3.1.65)$$

By the relation (A.1.4) in Appendix A.1, $l_1^*(\tilde{A}_n) \sim h_1(g_2(\frac{1}{2\beta}\log\, \tilde{A}_n)) e^{\nu \tilde{g}_1^*(\frac{1}{2\beta}\log\, \tilde{A}_n)}$ and hence (3.1.64) is also asymptotically equivalent to

$$\log \frac{h_1(g_2(\frac{1}{2\beta}\log\, \tilde{A}_n)) e^{\nu \tilde{g}_1^*(\frac{1}{2\beta}\log\, \tilde{A}_n)}}{h_1(m-1) L(e^{2m-2})^{\nu/2\beta}}. \qquad (3.1.66)$$

Since $k_m \le n < k_{m+1}$, we have

$$2\beta(m-1) + \log L(e^{2m-2}) \le \log \tilde{A}_n < 2\beta m + \log L(e^{2m})$$

and, by (3.1.27), $\frac{e^{\nu \tilde{g}_1^*(\frac{1}{2\beta}\log \tilde{A}_n)}}{L(e^{2m-2})^{\nu/2\beta}} = 1$. Hence, (3.1.66) simplifies to $\log \frac{h_1(m-1+\kappa)}{h_1(m-1)}$, where $0 \le$

$\kappa < 1$. But as $n \to \infty$, we have $m \to \infty$ and thus $\frac{h_1(m-1+\kappa)}{h_1(m-1)} \to 1$ by using (3.1.4). This proves that $\log n\tilde{A}_n^{-\alpha} l_1^*(\tilde{A}_n) \to 0$ and thus $n\tilde{A}_n^{-\alpha} l_1^*(\tilde{A}_n) \to 1$, as $n \to \infty$.

Finally, we show that k_n defined in (3.1.8) satisfies (2.1.14). Define $a_n = A_{k_n} = e^{2\beta(n-1)} L(e^{2n-2})$. Hence,

$$\log a_n = \log A_{k_n} = 2\beta(n-1) + \log L(e^{2n-2}).$$

Now observe that $\tilde{A}_{k_n} = a_n$ and thus (2.1.14) is satisfied. $\square$

The partial sums (3.1.7) involve centering constants B_{k_n} defined in (2.1.15). As in the stable case, one can expect to replace B_{k_n} by $k_n EX$ when $1 < \alpha < 2$, and to show the convergence of (3.1.7) without B_{k_n} when $0 < \alpha < 1$. The next result shows that this is indeed the case.

Proposition 3.1.1. *Suppose that the assumptions of Theorem 3.1.1 hold. Let*

$$\begin{aligned} \zeta &= -\frac{1-e^{-\nu}}{1-e^{2\beta-\nu}} - e^{\beta(2\lceil \frac{1}{\nu}\log c_1 \rceil - 1)}(c_1 e^{-\nu(\lceil \frac{1}{\nu}\log c_1 \rceil - 1)} - 1) \\ &\quad + c_1 \frac{(1-e^{-\nu})e^{\nu-\beta}}{1-e^{2\beta-\nu}} e^{(2\beta-\nu)\lceil \frac{1}{\nu}\log c_1 \rceil}. \end{aligned} \tag{3.1.67}$$

If $0 < \alpha < 1$, *then*

$$\frac{B_{k_n}}{A_{k_n}} \to \zeta, \quad \frac{1}{A_{k_n}} \sum_{j=1}^{k_n} X_j \xrightarrow{d} Y + \zeta$$

and if $1 < \alpha < 2$, *then*

$$\frac{k_n EX - B_{k_n}}{A_{k_n}} \to -\zeta, \quad \frac{1}{A_{k_n}} \left\{ \sum_{j=1}^{k_n} X_j - k_n EX \right\} \xrightarrow{d} Y + \zeta,$$

*where Y follows the semi-stable law characterized by (*3.1.8*) and (3.1.10).*

Proof. Case $0 < \alpha < 1$: It is enough to show the convergence of $\frac{B_{k_n}}{A_{k_n}} = \frac{k_n}{A_{k_n}} \int_{\frac{1}{k_n}}^{1-\frac{1}{k_n}} Q(s)ds$ to ζ. For fixed s_1 and s_2, write

$$\frac{k_n}{A_{k_n}} \int_{\frac{1}{k_n}}^{1-\frac{1}{k_n}} Q(s)ds = \frac{k_n}{A_{k_n}} \int_{\frac{1}{k_n}}^{s_1} Q(s)ds + \frac{k_n}{A_{k_n}} \int_{s_1}^{s_2} Q(s)ds + \frac{k_n}{A_{k_n}} \int_{s_2}^{1-\frac{1}{k_n}} Q(s)ds. \tag{3.1.68}$$

Observe first that, for fixed s_1 and s_2, the second term in (3.1.68) converges to zero. Indeed, this follows from the fact that $\frac{k_n}{A_{k_n}} \to 0$. For the latter convergence, note from (3.1.8) that

$$\frac{k_n}{A_{k_n}} \sim \frac{e^{(n-1)\nu}}{h_1(n-1)} \frac{1}{L(e^{2n-2})e^{2\beta(n-1)}}. \tag{3.1.69}$$

For arbitrarily small $\delta > 0$, by using Potter's bounds for L and Lemma A.1.6 for h_1, the right-hand side of (3.1.69) is bounded by $Ce^{(\nu-2\beta+\delta)(n-1)} \to 0$, as long as $\nu - 2\beta + \delta < 0$.

Consider now the third term in (3.1.68), involving the function $Q(s)$ for values of s close to 1. The function $Q(s)$ is defined as the inverse of the distribution function $F(x) = P(L(e^{W_q})e^{\beta W_q}(-1)^{W_q} \leq x)$. Since we are interested in $Q(s)$ for s close to 1, it is enough to look at the function for $x > 0$. For $x > 0$, the function $F(x)$ has jumps at points $x = L(e^{2n})e^{2\beta n}$ of size

$$P(W_q = 2n) = P(\frac{W_q}{2} \geq n, W_q \text{ is even}) - P(\frac{W_q}{2} \geq n+1, W_q \text{ is even}).$$

This means that, for s close to 1, the inverse function $Q(s)$ has jumps at points $s = 1 - P(\frac{W_q}{2} \geq n, W_q \text{ is even})$ of size $L(e^{2n})e^{2\beta n} - L(e^{2n-2})e^{2\beta(n-1)}$. Moreover, $Q(s) = L(e^{2n})e^{2\beta n}$ when $1 - P(\frac{W_q}{2} \geq n, W_q \text{ is even}) \leq s < 1 - P(\frac{W_q}{2} \geq n+1, W_q \text{ is even})$. (If this step is unclear, the reader may want to draw a picture.) Note that the jump points satisfy

$$1 - s = P(\frac{W_q}{2} \geq n, W_q \text{ is even}) = h_1(n)e^{-\nu n}$$

by (3.1.1).

Assuming for simplicity that $\frac{e^{\nu(n-1)}}{h_1(n-1)}$ are integers so that $k_n = \frac{e^{\nu(n-1)}}{h_1(n-1)}$ and taking $s_2 = 1 - h_1(n_1)e^{-\nu n_1}$, we can write,

$$\begin{aligned}
& \frac{k_n}{A_{k_n}} \int_{s_2}^{1-h_1(n-1)e^{-\nu(n-1)}} Q(s)ds \\
= \;& \frac{k_n}{A_{k_n}} \sum_{m=n_1}^{n-2} L(e^{2m})e^{2\beta m}(h_1(m)e^{-\nu m} - h_1(m+1)e^{-\nu(m+1)}) \\
= \;& \frac{e^{\nu(n-1)}}{h_1(n-1)e^{2\beta(n-1)}L(e^{2n-2})} \sum_{m=n_1}^{n-2} L(e^{2m})e^{2\beta m}h_1(m)e^{-\nu m}(1 - \frac{h_1(m+1)}{h_1(m)}e^{-\nu}) \\
=: \;& I_1 + I_2,
\end{aligned}$$

where, for fixed K,

$$I_1 = \frac{e^{\nu(n-1)}}{h_1(n-1)e^{2\beta(n-1)}L(e^{2n-2})} \sum_{m=n_1}^{n-K} L(e^{2m})e^{2\beta m}h_1(m)e^{-\nu m}(1 - \frac{h_1(m+1)}{h_1(m)}e^{-\nu}),$$

$$I_2 = \frac{e^{\nu(n-1)}}{h_1(n-1)e^{2\beta(n-1)}L(e^{2n-2})} \sum_{m=n-K}^{n-2} L(e^{2m})e^{2\beta m}h_1(m)e^{-\nu m}(1 - \frac{h_1(m+1)}{h_1(m)}e^{-\nu}).$$

For the term I_2, note that, after changing m to $n-j$ in the sum,

$$I_2 = e^{2\beta-\nu} \sum_{j=2}^{K} \frac{L(e^{2(n-j)})}{L(e^{2(n-1)})} e^{-(2\beta-\nu)j} \frac{h_1(n-j)}{h_1(n-1)}(1 - \frac{h_1(n-j+1)}{h_1(n-j)}e^{-\nu}).$$

By using (3.1.4), we get that

$$I_2 \to e^{2\beta-\nu}(1-e^{-\nu}) \sum_{j=2}^{K} e^{-(2\beta-\nu)j} = (1-e^{-\nu})\frac{e^{\nu-2\beta}}{1-e^{\nu-2\beta}}(1-e^{-(K-1)(2\beta-\nu)}), \quad (3.1.70)$$

as $n \to \infty$. For the term I_1, we have similarly

$$I_1 = e^{2\beta-\nu} \sum_{j=K}^{n-n_1} \frac{L(e^{2(n-j)})}{L(e^{2(n-2)})} e^{-(2\beta-\nu)j} \frac{h_1(n-j)}{h_1(n-1)} (1 - \frac{h_1(n-j+1)}{h_1(n-j)} e^{-\nu}).$$

For arbitrarily small $\delta > 0$, by using Potter's bounds and Lemma A.1.6, we can write

$$|I_1| \leq C \sum_{j=K}^{n-n_1} e^{-(2\beta-\nu-\delta)j}. \tag{3.1.71}$$

When $2\beta - \nu - \delta > 0$, the last bound is arbitrarily small for large enough K. Together with (3.1.70), this shows that

$$\frac{k_n}{A_{k_n}} \int_{s_2}^{1-h_1(n-1)e^{-(n-1)\nu}} Q(s)ds = I_1 + I_2 \to (1 - e^{-\nu}) \frac{e^{\nu-2\beta}}{1 - e^{\nu-2\beta}} = -\frac{1 - e^{-\nu}}{1 - e^{2\beta-\nu}},$$

as $n \to \infty$.

Consider now the first term in (3.1.68), involving the function $Q(s)$ for values of s close to 0. Here we need to examine the function $F(x) = P(L(e^{W_q})e^{\beta W_q}(-1)^{W_q} \leq x)$ for $x < 0$. For $x < 0$, the function $F(x)$ has jumps at $x = -L(e^{2n+1})e^{\beta(2n+1)}$ of size

$$P(W_q = 2n+1) = P(\frac{W_q - 1}{2} \geq n, W_q \text{ is odd}) - P(\frac{W_q - 1}{2} \geq n+1, W_q \text{ is odd}).$$

Moreover, $Q(s) = -L(e^{2n+1})e^{\beta(2n+1)}$ when $P(\frac{W_q-1}{2} \geq n+1, W_q \text{ is odd}) < s \leq P(\frac{W_q-1}{2} \geq n, W_q \text{ is odd})$. Note that, by (3.1.2), the jump points satisfy

$$s = P(\frac{W_q - 1}{2} \geq n, W_q \text{ is odd}) = h_2(n)e^{-\nu n}.$$

Write the first term in (3.1.68) as

$$\frac{k_n}{A_{k_n}}\int_{h_1(n-1)e^{-\nu(n-1)}}^{h_2(l(n)-1)e^{-\nu(l(n)-1)}} Q(s)ds + \frac{k_n}{A_{k_n}}\int_{h_2(l(n)-1)e^{-\nu(l(n)-1)}}^{s_1} Q(s)ds =: I_1^* + I_2^*, \quad (3.1.72)$$

where $l(n)$ is the integer such that

$$h_2(l(n))e^{-\nu l(n)} \leq h_1(n-1)e^{-\nu(n-1)} < h_2(l(n)-1)e^{-\nu(l(n)-1)}$$

or

$$h_2(l(n))e^{-\nu l(n)} \leq h_2(n-1)e^{-\nu(n-1+\frac{1}{\nu}\log\frac{h_2(n-1)}{h_1(n-1)})} < h_2(l(n)-1)e^{-\nu(l(n)-1)}.$$

Note that, when $\frac{h_2(x)}{h_1(x)} \to c_1$ and $\frac{1}{\nu}\log c_1$ is not an integer, or when $\frac{1}{\nu}\log c_1$ is an integer and $\frac{h_2(x)}{h_1(x)} \uparrow c_1$, for large values of n one can take $l(n) = n-1+\lceil\frac{1}{\nu}\log c_1\rceil$. Indeed, this follows from

$$e^{-\nu} < e^{-\nu(\lceil\frac{1}{\nu}\log\frac{h_2(n-1)}{h_1(n-1)}\rceil - \frac{1}{\nu}\log\frac{h_2(n-1)}{h_1(n-1)})} \leq 1 \quad (3.1.73)$$

and the fact that

$$\frac{h_2(n-1+\lceil\frac{1}{\nu}\log\frac{h_2(n-1)}{h_1(n-1)}\rceil)}{h_2(n-1)} \to 1, \quad (3.1.74)$$

as $n \to \infty$.

Now, taking $s_1 = h_2(n_2)e^{-\nu n_2}$, we can write I_2^* in (3.1.72) as

$$I_2^* = -\frac{k_n}{A_{k_n}}\sum_{m=n_2}^{l(n)-2} L(e^{2m+1})e^{\beta(2m+1)}(h_2(m)e^{-\nu m} - h_2(m+1)e^{-\nu(m+1)}).$$

Following a similar calculation as done for the third term in (3.1.68), we get, as

$n \to \infty$,

$$I_2^* \to -c_1 \frac{(1-e^{-\nu})e^{2(\nu-2\beta)}e^{\beta}}{1-e^{\nu-2\beta}} e^{-(\nu-2\beta)\lceil \frac{1}{\nu}\log c_1 \rceil} = c_1 \frac{(1-e^{-\nu})e^{\nu-\beta}}{1-e^{2\beta-\nu}} e^{(2\beta-\nu)\lceil \frac{1}{\nu}\log c_1 \rceil}.$$

One can write I_1^* in (3.1.72) as

$$\begin{aligned} I_1^* &= -\frac{e^{\nu(n-1)}}{h_1(n-1)e^{2\beta(n-1)}L(e^{2n-2})} L(e^{2l(n)-1})e^{\beta(2l(n)-1)} \cdot \\ &\qquad (h_2(l(n)-1)e^{-\nu(l(n)-1)} - h_1(n-1)e^{-\nu(n-1)}). \end{aligned}$$

It can be seen that

$$I_1^* \to -e^{\beta(2\lceil \frac{1}{\nu}\log c_1 \rceil - 1)}(c_1 e^{-\nu(\lceil \frac{1}{\nu}\log c_1 \rceil - 1)} - 1),$$

as $n \to \infty$.

Now we consider the case when $\frac{1}{\nu}\log c_1$ is an integer and $\frac{h_2(x)}{h_1(x)} \downarrow c_1$. We want to find $l(n)$ such that (3.1.73) holds. Hence, we want

$$\lim_{n\to\infty} \frac{h_2(n-1)}{h_2(l(n))} e^{-\nu(n-1+\frac{1}{\nu}\log\frac{h_2(n-1)}{h_1(n-1)} - l(n))} \geq 1.$$

Take $l(n) = n - 2 + \lceil \frac{1}{\nu}\log \frac{h_2(n-1)}{h_1(n-1)} \rceil$. Then, $\lim_{n\to\infty} \frac{h_2(n-1)}{h_1(l(n))} \to 1$. Now,

$$\lim_{n\to\infty} e^{-\nu(n-1+\frac{1}{\nu}\log\frac{h_2(n-1)}{h_1(n-1)} - n + 2 - \lceil \frac{1}{\nu}\log\frac{h_2(n-1)}{h_1(n-1)} \rceil)} = e^{-\nu(1+\frac{1}{\nu}\log c_1 - \frac{1}{\nu}\log c_1 - 1)} = e^0 = 1.$$

We also need

$$\frac{h_2(l(n)-1)}{h_2(n-1)} e^{-\nu(l(n)-1-n+1-\frac{1}{\nu}\log\frac{h_2(n-1)}{h_1(n-1)})} > 1$$

for large n. For this, observe that $\frac{h_2(l(n)-1)}{h_2(n-1)} \to 1$ and

$$\lim_{n\to\infty} e^{-\nu(l(n)-1-n+1-\frac{1}{\nu}\log\frac{h_2(n-1)}{h_1(n-1)})} = \lim_{n\to\infty} e^{-\nu(n-3+\lceil\frac{1}{\nu}\log\frac{h_2(n-1)}{h_1(n-1)}\rceil-n+1-\frac{1}{\nu}\log\frac{h_2(n-1)}{h_1(n-1)})}$$
$$= \lim_{n\to\infty} e^{-\nu(-2+1+\frac{1}{\nu}\log c_1-\frac{1}{\nu}\log c_1)} = e^{-\nu}.$$

Hence, when $\frac{1}{\nu}\log c_1$ is an integer, we have $\frac{h_2(x)}{h_1(x)} \downarrow c_1$ and $\lceil\frac{1}{\nu}\log\frac{h_2(x)}{h_1(x)}\rceil \downarrow \frac{1}{\nu}\log c_1 + 1$, and as in the previous calculations,

$$I_1^* \to -e^{\beta(2\lceil\frac{1}{\nu}\log c_1\rceil-1)}(c_1 e^{-\nu(\lceil\frac{1}{\nu}\log c_1\rceil-1)} - 1)$$

and

$$I_2^* \to c_1\frac{(1-e^{-\nu})e^{\nu-\beta}}{1-e^{2\beta-\nu}}e^{(2\beta-\nu)\lceil\frac{1}{\nu}\log c_1\rceil}.$$

Finally, gathering the results above, we deduce the convergence to the constant ζ given by (3.1.67).

<u>Case $1<\alpha<2$:</u> It is enough to show the convergence of $\frac{k_n EX - B_{k_n}}{A_{k_n}}$ to $-\zeta$. Using the fact that $EX = \int_0^1 Q(s)ds$, observe that

$$\frac{k_n EX - B_{k_n}}{A_{k_n}} = \frac{k_n}{A_{k_n}}\int_0^{\frac{1}{k_n}} Q(s)ds + \frac{k_n}{A_{k_n}}\int_{1-\frac{1}{k_n}}^1 Q(s)ds.$$

For simplicity, we assume that $\frac{e^{(n-1)\nu}}{h_1(n-1)}$ is an integer. To evaluate $\int_{1-\frac{1}{k_n}}^1 Q(s)ds$, one follows a similar procedure as in the case $0<\alpha<1$ to obtain

$$\begin{aligned}&\frac{k_n}{A_{k_n}}\int_{1-h_1(n-1)e^{-\nu(n-1)}}^1 Q(s)ds\\ =\;&\frac{k_n}{A_{k_n}}\sum_{m=n-1}^{\infty} L(e^{2m})e^{2\beta m}(h_1(m)e^{-\nu m} - h_1(m+1)e^{-\nu(m+1)}) =: \tilde{I}_1 + \tilde{I}_2,\end{aligned}$$

where, for fixed K,

$$\tilde{I}_1 = \frac{e^{\nu(n-1)}}{h_1(n-1)e^{2\beta(n-1)}L(e^{2n-2})}\sum_{m=n-1}^{n+K} L(e^{2m})e^{2\beta m}h_1(m)e^{-\nu m}(1-\frac{h_1(m+1)}{h_1(m)}e^{-\nu}),$$

$$\tilde{I}_2 = \frac{e^{\nu(n-1)}}{h_1(n-1)e^{2\beta(n-1)}L(e^{2n-2})}\sum_{m=n+K}^{\infty} L(e^{2m})e^{2\beta m}h_1(m)e^{-\nu m}(1-\frac{h_1(m+1)}{h_1(m)}e^{-\nu}).$$

Similar to the case $0 < \alpha < 1$, one can show that

$$\frac{k_n}{A_{k_n}}\int_{1-h_1(n-1)e^{-(n-1)\nu}}^{1} Q(s)ds = \tilde{I}_1 + \tilde{I}_2 \to \frac{1-e^{-\nu}}{1-e^{2\beta-\nu}}.$$

Similarly, one can write

$$\begin{aligned}\int_0^{\frac{1}{k_n}} Q(s)ds &= \frac{k_n}{A_{k_n}}\int_0^{h_2(l(n)-1)e^{-\nu(l(n)-1)}} Q(s)ds - \frac{k_n}{A_{k_n}}\int_{h_1(n-1)e^{-\nu(n-1)}}^{h_2(l(n)-1)e^{-\nu(l(n)-1)}} Q(s)ds \\ &=: \tilde{I}_2^* - \tilde{I}_1^*.\end{aligned}$$

As shown in the case $0 < \alpha < 1$, we again use two different representations of $l(n)$ for two different cases. Note that $\tilde{I}_1^*$ is exactly I_1^* considered in that case.

Observe that

$$\tilde{I}_2^* = -\frac{k_n}{A_{k_n}}\sum_{m=l(n)-1}^{\infty} L(e^{2m+1})e^{(2m+1)\beta}(h_2(m)e^{-\nu m} - h_2(m+1)e^{-\nu(m+1)}).$$

As $n \to \infty$,

$$\tilde{I}_2^* \to -c_1\frac{(1-e^{-\nu})e^{\nu-\beta}}{1-e^{2\beta-\nu}}e^{(2\beta-\nu)\lceil\frac{1}{\nu}\log c_1\rceil}$$

and, from the case $0 < \alpha < 1$,

$$\tilde{I}_1^* \to -e^{\beta(2\lceil\frac{1}{\nu}\log c_1\rceil - 1)}(c_1 e^{-\nu\lceil\frac{1}{\nu}\log c_1\rceil} - 1).$$

Finally, gathering the results above, we deduce the convergence to $-\zeta$ where ζ is given by (3.1.67). $\square$

We state a corollary to this theorem where the random variable X can only take positive values.

Corollary 3.1.1. *Let W_q be an integer-valued random variable taking values in $0, 1, 2, \ldots$ such that, for all $x > 0$,*

$$P(W_q \geq x) = \sum_{n=\lceil x\rceil}^{\infty} P(W_q = n) = h_1(\lceil x\rceil)e^{-\nu\lceil x\rceil}, \tag{3.1.75}$$

where $\nu > 0$ and the function h_1 satisfies

$$\frac{h_1(ax)}{h_1(x)} \to 1 \text{ as } x \to \infty, a \to 1. \tag{3.1.76}$$

Let also

$$X = L(e^{W_q})e^{\beta W_q}, \tag{3.1.77}$$

where $\beta > 0$ and L is a slowly varying function at ∞ such that $L(e^n)$ is ultimately monotonically increasing. Suppose that

$$\alpha := \frac{\nu}{\beta} < 2 \tag{3.1.78}$$

Then, X is in the domain of attraction of a semi-stable distribution in the following *sense. If $X, X_1, X_2, \ldots$ are i.i.d. random variables, then as $n \to \infty$, the partial sums*

$$\frac{1}{A_{k_n}} \left\{ \sum_{j=1}^{k_n} X_j - B_{k_n} \right\} \tag{3.1.79}$$

converge to a semi-stable distribution with

$$k_n = \left\lceil \frac{e^{(n-1)\nu}}{h_1(n-1)} \right\rceil, \quad A_{k_n} = L(e^{n-1}) e^{(n-1)\beta} \tag{3.1.80}$$

and B_{k_n} given by (2.1.15). The limiting semi-stable distribution is non-Gaussian and is characterized by

$$\alpha = \frac{\nu}{\beta}, \tag{3.1.81}$$

$$M_L(-x) \equiv 0 \qquad M_R(x) = e^{-\nu(\lceil \frac{1}{\beta} \log x \rceil_+ - \frac{1}{\beta} \log x)}, \quad x > 0. \tag{3.1.82}$$

The proof of the corollary is similar to the proof of Theorem 3.1.1, and is omitted.

Theorem 3.1.1 concerns the partial sums $\sum_{j=1}^{n} X_j$ along a subsequence k_n of n. The following result describes the behavior of the partial sums across all n. The result is a direct consequence of Lemma 5 of Meerschaert and Scheffler [22]. Recall that a collection of random variables $\{Y_n\}_{n \geq 1}$ is called *stochastically compact* if every subsequence $\{n'\}$ has a further subsequence $\{n''\} \subset \{n'\}$ for which $\{Y_{n''}\}$ converges in distribution. The following notation will also be used. For a semi-stable distribution τ with characteristic function $\psi(t)$, τ^λ will denote the semi-stable distribution with the characteristic function $\psi(t)^\lambda$.

Proposition 3.1.2. *Let $X, X_1, X_2, \dots$ be i.i.d. random variables such that*

$$\frac{1}{A_{k_n}}\Big\{\sum_{j=1}^{k_n} X_j - B_{k_n}\Big\} \xrightarrow{d} Y,$$

where Y follows a semi-stable distribution τ with $0 < \alpha < 2$ and k_n,A_{k_n},B_{k_n} are given in (2.1.5), (2.1.6) and (2.1.15). Then, there exist a_n and b_n such that a_n is regularly varying with index $\frac{1}{\alpha}$, $a_{k_n} = A_{k_n}$ and $a_n^{-1}(X_1 + X_2 + \dots + X_n) - b_n$ is stochastically compact, with every limit point of the form $\lambda^{-\frac{1}{\alpha}}\tau^{\lambda}$ for some $\lambda \in [1, c]$. Moreover, one can take

$$a_n = \lambda_n^{\frac{1}{\alpha}} A_{k_{p_n}} \qquad \text{and} \qquad b_n = \lambda_n^{1-\frac{1}{\alpha}} \frac{B_{k_{p_n}}}{A_{k_{p_n}}},$$

where $\lambda_n = \frac{n}{k_{p_n}}$ and p_n, k_{p_n} are chosen so that $k_{p_n} \le n < k_{p_{n+1}}$ for every $n \ge 1$.

Proof. The proposition follows directly from Lemma 5 and its proof in Meerschaert and Scheffler [22]. □

Corollary 3.1.2. *Under the assumptions of Proposition 3.1.2,*

$$\limsup_n P(a_n^{-1}(X_1 + X_2 + \dots + X_n) - b_n > x) \le \sup_{1\le\lambda\le c} P(Y_\lambda > x) \tag{3.1.83}$$

and

$$\limsup_n P(a_n^{-1}(X_1 + X_2 + \dots + X_n) - b_n < x) \le \sup_{1\le\lambda\le c} P(Y_\lambda < x), \tag{3.1.84}$$

where Y_λ has the distribution of the form $\lambda^{-\frac{1}{\alpha}}\tau^{\lambda}$.

Proof. Along a subsequence $\{n(k)\}$ of $\{n\}$, we have

$$\begin{aligned} &\limsup_n P(a_n^{-1}(X_1 + X_2 + \ldots + X_n) - b_n > x) \\ = \; &\lim_k P(a_{n(k)}^{-1}(X_1 + X_2 + \ldots + X_{n(k)}) - b_{n(k)} > x). \end{aligned} \tag{3.1.85}$$

Now, by Proposition 3.1.2, there exists a further subsequence $\{n(k_m)\}$ of $\{n(k)\}$ such that

$$\lim_m P(a_{n(k_m)}^{-1}(X_1 + X_2 + \ldots + X_{n(k_m)}) - b_{n(k_m)} > x) = P(Y_\lambda > x), \tag{3.1.86}$$

where Y_λ follows the distribution $\lambda^{-\frac{1}{\alpha}}\tau^\lambda$. The relation (3.1.86) holds for all x as long as the semi-stable distribution τ^λ is continuous. By Huff [17], the continuity of τ^λ is equivalent to $\int_{-\infty}^0 dL_\lambda(x) + \int_0^\infty dR_\lambda(x) = \infty$, where L_λ and R_λ define the Lévy measure of τ^λ. By the definition of τ^λ, $L_\lambda = \lambda L$ and $R_\lambda = \lambda R$. Denote the multiplicative period of $M_L(x)$ and $M_R(x)$ by $p > 1$. Then, after the change of variables $x = p^k y$ in the integrals below,

$$\begin{aligned} \int_{-\infty}^0 dL(x) + \int_0^\infty dR(x) &= \sum_{k=-\infty}^{\infty} \int_{-p^{k+1}}^{-p^k} d\frac{M_L(x)}{|x|^\alpha} + \sum_{k=-\infty}^{\infty} \int_{p^k}^{p^{k+1}} d\frac{(-M_R(x))}{x^\alpha} \\ = \sum_{k=-\infty}^{\infty} p^{-k\alpha} \int_{-p}^{-1} d\frac{M_L(y)}{|y|^\alpha} &+ \sum_{k=-\infty}^{\infty} p^{-k\alpha} \int_1^p d\frac{(-M_R(y))}{y^\alpha} = \infty, \end{aligned}$$

unless $M_L \equiv 0$ and $M_R \equiv 0$. Combining (3.1.85) and (3.1.86), we have (3.1.83) for all $x \in \mathbb{R}$. The relation (3.1.84) can be obtained similarly. □

3.2 Special case and numerical illustrations

In this section, we consider a special case of Theorem 3.1.1, supposing that W_q is a geometric random variable with parameter $1 - s$ and $X = e^{\beta W_q}(-1)^{W_q}$. We also explore numerically how well the limiting semi-stable distribution approximates the partial sums, where the limiting semi-stable distribution is computed by using the method of Abate and Whitt [1] described in Sections 2.2–2.3. The presentation here is similar to that of Example 2.3.1 where the random variable $X = e^{\beta W_q}$ was considered. The key difference is that X was supported by the positive real line in Example 2.3.1, whereas here the support is the entire real line.

Example 3.2.1. Let W_q be a random variable following a geometric distribution with success probability $1-s$ where $s \in (0,1)$, that is,

$$f_{W_q}(x) = s^{x-1}(1-s), \quad x = 1, 2, \ldots. \tag{3.2.1}$$

Let $X, X_1, X_2, \ldots$ be a sequence of i.i.d. random variables where

$$X = (-1)^{W_q} e^{\beta W_q}.$$

We will use Theorem 3.1.1 to show that the distribution F of X is in the domain of attraction of a semi-stable distribution. For the convergence of the partial sums (3.1.7), we shall also derive the normalizing constants A_{k_n}, centering constants B_{k_n} and the subsequence k_n along which the convergence takes place.

From Theorem 3.1.1, k_n and A_{k_n} can be obtained having the constant ν and the functions $h_1(x)$ and $h_2(x)$. Observe that

$$P(\frac{W_q}{2} \geq x, W_q \text{ is even}) = \sum_{n=\lceil x \rceil}^{\infty} s^{2n-1}(1-s) = \frac{1}{s(s+1)} e^{-\lceil x \rceil (\log \frac{1}{s^2})} \tag{3.2.2}$$

and

$$P(\frac{W_q - 1}{2} \geq x, W_q \text{ is odd}) = \sum_{n=\lceil x \rceil}^{\infty} s^{2n}(1-s) = \frac{1}{s+1} e^{-\lceil x \rceil (\log \frac{1}{s^2})}. \tag{3.2.3}$$

Hence, we have $\nu = \log \frac{1}{s^2}, h_1(x) = \frac{1}{s(s+1)}$, $h_2(x) = \frac{1}{s+1}$ and then

$$k_n = \lceil \frac{s+1}{s^{2n-3}} \rceil, \quad A_{k_n} = e^{(2n-2)\beta}. \tag{3.2.4}$$

For the distribution to be attracted to a semi-stable distribution, we also need to have $\alpha < 2$ where $\alpha = \frac{1}{\beta} \log \frac{1}{s}$. We make a further assumption that $E(X)$ is finite. This assumption holds when $1 < \alpha < 2$ or, equivalently,

$$\frac{1}{2} \log \frac{1}{s} < \beta < \log \frac{1}{s}. \tag{3.2.5}$$

The limiting semi-stable distribution has the spectral functions

$$\begin{aligned} M_L(-x) &= s^{2[\frac{1}{2} + \frac{1}{2\beta} \log x] + 1 - \frac{1}{\beta} \log x}, \\ M_R(x) &= s^{2\lceil \frac{1}{2\beta} \log x \rceil_+ - \frac{1}{\beta} \log x}. \end{aligned}$$

The centering constant ζ appearing in (3.1.66) of Proposition 3.1.1 is

$$\zeta = -\frac{1-s^2}{1-s^2 e^{2\beta}} - e^{-\beta}(s^{-1} - 1) + \frac{s(1-s^2)}{(1-s^2 e^{2\beta}) s e^{\beta}}, \tag{3.2.6}$$

where we used the fact that $c_1 = s$ and $\frac{1}{\nu} \log c_1 = -\frac{1}{2}$. The log characteristic function of the limiting semi-stable distribution $Y + \zeta$ in Proposition 3.1.1 is then

$$\begin{aligned}\log \hat{\mu}(t) &= i\zeta t + it(\Theta(\psi_1) - \Theta(\psi_2)) \\ &+ \sum_{k=-\infty}^{\infty} (e^{-ite^{(2k+1)\beta}} - 1 + it\frac{e^{(2k+1)\beta}}{1 + e^{(4k+2)\beta}})(1 - s^2)s^{2k+1} \\ &+ \sum_{k=-\infty}^{\infty} (e^{-ite^{2k\beta}} - 1 + it\frac{e^{2k\beta}}{1 + e^{4k\beta}})(1 - s^2)s^{2k}, \end{aligned} \tag{3.2.7}$$

where

$$\begin{aligned}\Theta(\psi_1) &= -\sum_{k=-1}^{\infty} \frac{e^{(2k+3)\beta}}{1 + e^{(4k+6)\beta}} s^{2k+3}(1 - s^2) - \frac{e^{-2\beta}}{1 + e^{-2\beta}}(1 - s) + \frac{e^{-3\beta}}{1 + e^{-2\beta}}(s^{-1} - 1) \\ &+ \sum_{k=-\infty}^{-3} \frac{e^{(6k+9)\beta}}{1 + e^{(4k+6)\beta}} s^{2k+3}(1 - s^2)\end{aligned}$$

and

$$\Theta(\psi_2) = -\sum_{k=0}^{\infty} \frac{e^{2k\beta}}{1 + e^{4k\beta}}(1 - s^2)s^{2k} + \sum_{k=-\infty}^{0} \frac{e^{6k\beta}}{1 + e^{4k\beta}}(1 - s^2)s^{2k}.$$

We next want to calculate numerically the density of the limiting semi-stable distribution having the characteristic function (3.2.7). We use the Abate and Whitt [1] method described in Sections 2.2–2.3. Write the limiting semi-stable random variable $Y + \xi$ as

$$T = U^+ + U^- + \xi,$$

where the characteristic functions of U^+ and U^- are given by

$$\sum_{k=-\infty}^{\infty} (e^{-ite^{2k\beta}} - 1 + it\frac{e^{2k\beta}}{1 + e^{4k\beta}})(1 - s^2)s^{2k}$$

and

$$\sum_{k=-\infty}^{\infty}(e^{-ite^{(2k+1)\beta}}-1+it\frac{e^{(2k+1)\beta}}{1+e^{(4k+2)\beta}})(1-s^2)s^{2k+1},$$

and $\xi=\Theta(\psi_2)-\Theta(\psi_1)-\zeta$. The reason for splitting T into U^+ and U^- is that the Lévy spectral function $M_R(x)$ of U^+ is supported only on the positive real line. Similarly, the Lévy spectral function $M_L(x)$ of U^- is supported only on the negative real line. By Theorem 3 of Watanabe and Yamamuro [27], the positive tail of U^+ is heavy-tailed with the exponent α while the negative tail decays exponentially fast. Similarly, the negative tail of U^- is heavy-tailed with the exponent α while the positive tail decays exponentially fast. This is the reason why it is easier to numerically calculate the densities of U^+ and U^- separately (using the Abate and Whitt method by shifting densities to be concentrated mostly on the positive real line) and then convolute the two densities to get the density of T after shifting it by the constant ξ.

Figures 3.2 and 3.3 compare the limiting semi-stable density and the empirical histogram of the partial sum S_n. We fix $s=0.25$ but vary the value of β. In Figure 3.2, we see that the numerically calculated density approximates well the right tails of the histogram, though the numerically calculated density is not necessarily a good approximation of the histogram in the central section and in the left tail. As β moves further up from 1, the approximation of the central part and the left tail by the limiting semi-stable density is worse. In Figure 3.3, we present analogous plots with smaller values of k_n. Similar conclusions can be drawn, though the approximation improves for larger values of k_n.

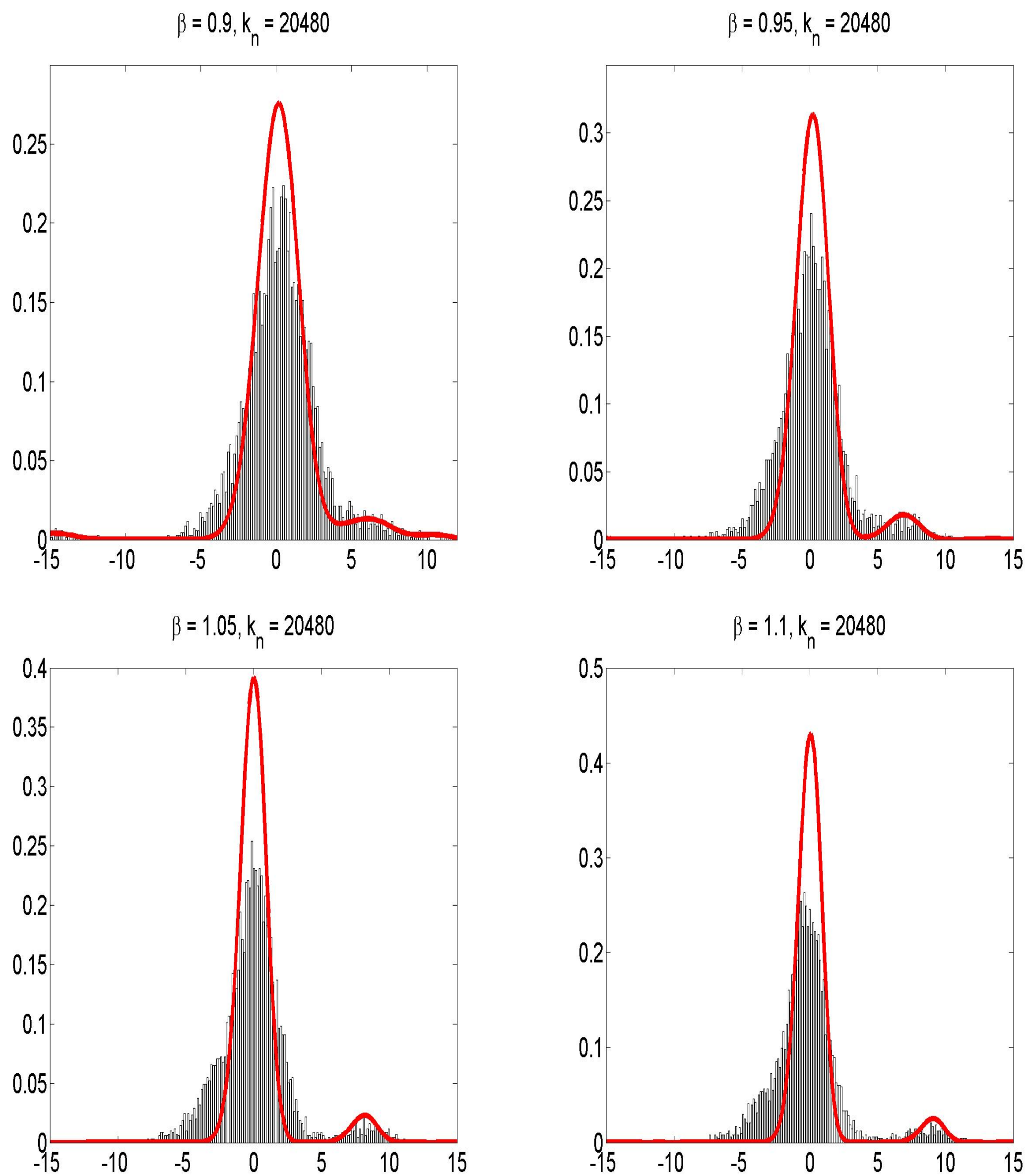

Figure 3.2: The empirical histogram against the actual density.

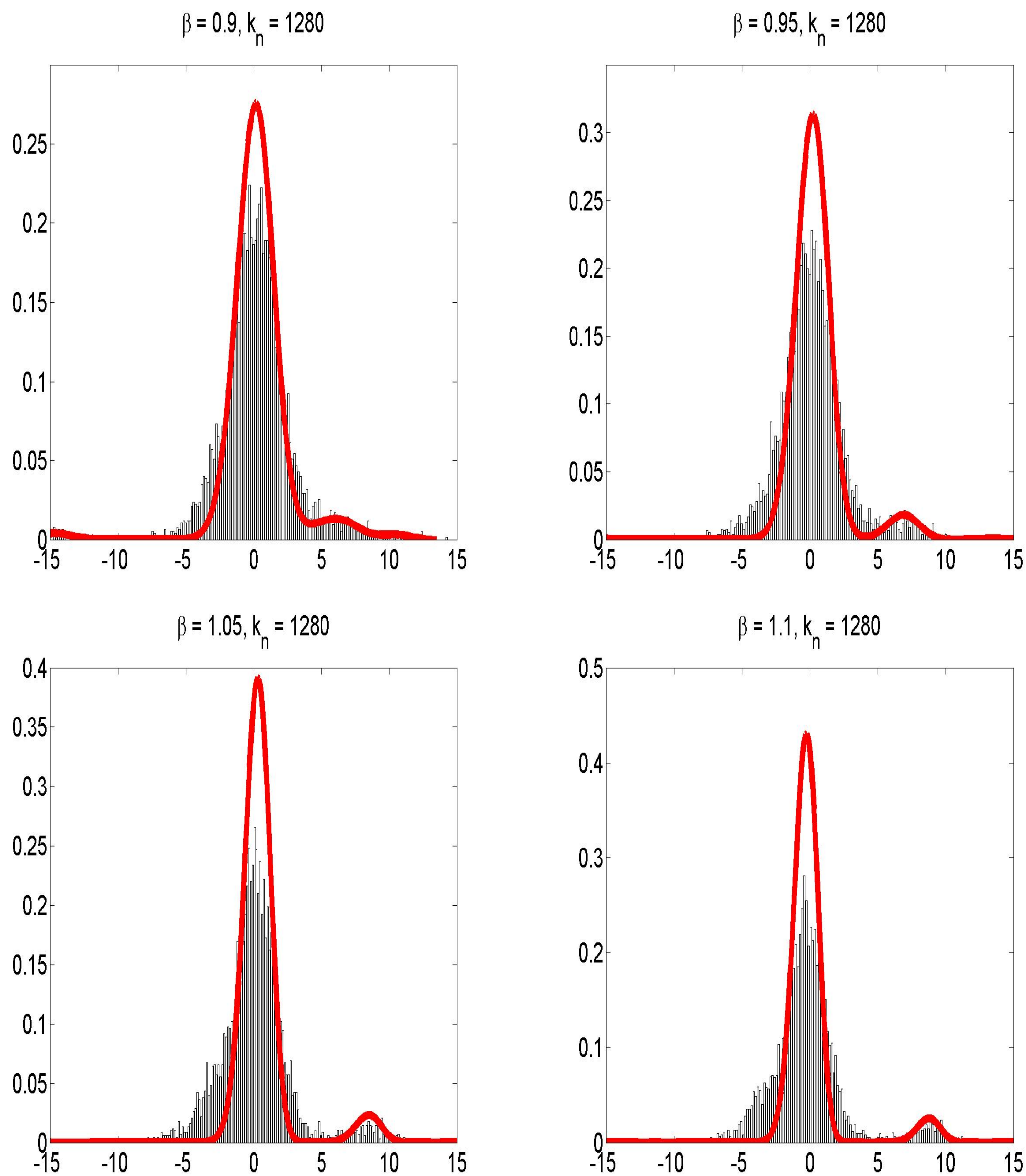

Figure 3.3: The empirical histogram against the actual density.

CHAPTER 4: APPLICATION TO SAMPLING

In this chapter, we turn back to the context of sampling of finite point processes and are interested, more specifically, in the asymptotic behavior of the estimator $\widehat{f}_W(w)$ in (1.1.4)–(1.1.5) of the p.m.f. $f_W(w)$ of flow sizes. In Section 4.1, we apply the general results on convergence of the partial sums to a semi-stable distribution in Section 3.1 to the estimator $\widehat{f}_W(w)$. In particular, we derive conservative confidence intervals for the p.m.f. $f_W(w)$ based on the estimator $\widehat{f}_W(w)$. In Section 4.2, we illustrate the results of the previous section on two examples of p.m.f.'s $f_W(w)$. Finally, in Section 4.3, we investigate the performance of the introduced confidence intervals through a small simulation study.

4.1 General results

The following result restates Theorem 3.1.1 and Proposition 3.1.1 for the non-parametric estimator $\widehat{f}_W(w)$ of $f_W(w)$ given in (1.1.4) or (1.1.11)–(1.1.12).

Theorem 4.1.1. *Suppose conditions (3.1.1)–(3.1.4) hold and k_n is given in (3.1.8). Let*

$$\alpha = \frac{\nu}{2\log(q^{-1}-1)}. \tag{4.1.1}$$

If $\alpha \in (1,2)$, then

$$d_N(\widehat{f}(w) - f(w)) \xrightarrow{d} (-1)^{-w}(Y+\zeta),$$

and if $\alpha \in (0,1)$, then

$$d_N\widehat{f}(w) \xrightarrow{d} (-1)^{-w}(Y+\zeta),$$

along the sample sizes $N = k_n$, where $d_N = \frac{k_n}{A_{k_n}}$ with

$$A_{k_n} = \binom{2n-2}{w}(1-q)^{-w}\left(q^{-1}-1\right)^{2n-2}, \tag{4.1.2}$$

and ζ is defined in (3.1.67) and Y is a semi-stable distribution characterized by (3.1.10) with

$$\beta = \log(q^{-1}-1). \tag{4.1.3}$$

Proof. In view of (1.1.11)–(1.1.12), we are interested in the distribution of

$$X = \binom{W_q}{w}(-1)^{W_q-w}\frac{(1-q)^{W_q-w}}{q^{W_q}}1_{\{W_q\geq w\}},$$

where $w > 0$ is fixed and W_q follows a p.m.f. satisfying (3.1.1)–(3.1.4). For $W_q > w$ large enough, one can write $(-1)^w X = L(e^{W_q})e^{\beta W_q}(-1)^{W_q}$ as given in Theorem 3.1.1 with

$$L(x) = \binom{\log x}{w}(1-q)^{-w} = (1-q)^{-w}\frac{\prod_{i=0}^{w-1}(\log x - i)}{w!} \tag{4.1.4}$$

and $\beta = \log\frac{1-q}{q} = \log(q^{-1}-1)$. Observe that $L(x)$ is an ultimately increasing slowly varying function. Hence, when $\alpha \in (1,2)$, by using (1.1.11)–(1.1.12) and applying Theorem 3.1.1 and Proposition 3.1.1,

$$\frac{k_n}{A_{k_n}}(\widehat{f}_W(w) - f_W(w)) = d_N(\widehat{f}_W(w) - f_W(w))$$

converges to a semi-stable distribution $(-1)^{-w}(Y+\zeta)$ with α in (4.1.1) and A_{k_n} in (4.1.2). When $\alpha \in (0,1)$,

$$\frac{k_n}{A_{k_n}}\widehat{f}_W(w) = d_N\widehat{f}_W(w)$$

converges to a semi-stable distribution $(-1)^{-w}(Y+\zeta)$ with α in (4.1.1) and A_{k_n} in

(4.1.2). $\square$

The next result provides a conservative confidence interval for $f(w)$ based on $\widehat{f}(w)$ when $1 < \alpha < 2$.

Proposition 4.1.1. *Under the assumptions and notation of Theorem 4.1.1, suppose $\alpha \in (1,2)$. For $\gamma \in (0,1)$, set*

$$\mathcal{C} = [\widehat{f}_W(w) - \tilde{b}_N x_{1-\frac{\gamma}{2}}, \widehat{f}_W(w) - \tilde{b}_N x_{\frac{\gamma}{2}}], \tag{4.1.5}$$

where

$$\tilde{b}_N = N^{\frac{1}{\alpha}-1} A_{k_{p_N}} k_{p_N}^{-\frac{1}{\alpha}} \tag{4.1.6}$$

with p_N such that $k_{p_N} \leq N < k_{p_{N+1}}$ and

$$\sup_{1\leq\lambda\leq c} P(Y_\lambda^\zeta < x_{\frac{\gamma}{2}}) = \frac{\gamma}{2}, \quad \sup_{1\leq\lambda\leq c} P(Y_\lambda^\zeta > x_{1-\frac{\gamma}{2}}) = \frac{\gamma}{2}, \tag{4.1.7}$$

where Y_λ^ζ has the distribution of the form $\lambda^{-\frac{1}{\alpha}}\tau^\lambda$ and τ is the distribution of $Y + \zeta$. Then,

$$\liminf_{N\to\infty} P(f_W(w) \in \mathcal{C}) \geq 1 - \gamma, \tag{4.1.8}$$

that is, $\mathcal{C}$ is a conservative $100(1-\gamma)\%$ confidence interval for $f_W(w)$.

Proof. When $\alpha \in (1,2)$, by using Corollary 3.1.2 and Theorem 4.1.1, we get

$$\begin{aligned}
&\limsup_{N\to\infty} P(\frac{N\lambda_N^{-\frac{1}{\alpha}}}{A_{k_{p_N}}}\widehat{f}_W(w) - \lambda_N^{1-\frac{1}{\alpha}}\frac{k_{p_N}}{A_{k_{p_N}}} f_W(w) < x_{\frac{\gamma}{2}}) \leq \sup_{1\leq\lambda\leq c} P(Y_\lambda^\eta < x_{\frac{\gamma}{2}}) = \frac{\gamma}{2} \\
\Leftrightarrow &\limsup_{N\to\infty} P(\frac{N}{\lambda_N k_{p_N}}\widehat{f}_W(w) - \frac{\lambda_N^{\frac{1}{\alpha}-1} A_{k_{p_N}}}{k_{p_N}} x_{\frac{\gamma}{2}} < f_W(w)) \leq \frac{\gamma}{2}.
\end{aligned}$$

Using $\lambda_N = \frac{N}{k_{p_N}}$, we get

$$\limsup_{N\to\infty} P(\widehat{f}_W(w) - N^{\frac{1}{\alpha}-1} A_{k_{p_N}} k_{p_N}^{-\frac{1}{\alpha}} x_{\frac{\gamma}{2}} < f_W(w)) \leq \frac{\gamma}{2}. \tag{4.1.9}$$

Similarly for the right tail, we get

$$\begin{aligned}
&\limsup_{N\to\infty} P(\frac{N\lambda_N^{-\frac{1}{\alpha}}}{A_{k_{p_N}}}\widehat{f}_W(w) - \lambda_N^{1-\frac{1}{\alpha}}\frac{k_{p_N}}{A_{k_{p_N}}} f_W(w) > x_{1-\frac{\gamma}{2}}) \leq \sup_{1\leq\lambda\leq c} P(Y_\lambda^\eta > x_{1-\frac{\gamma}{2}}) = \frac{\gamma}{2} \\
&\Leftrightarrow \limsup_{N\to\infty} P(\frac{N}{\lambda_N k_{p_N}}\widehat{f}_W(w) - \frac{\lambda_N^{\frac{1}{\alpha}-1} A_{k_{p_N}}}{k_{p_N}} x_{1-\frac{\gamma}{2}} > f_W(w)) \leq \frac{\gamma}{2} \\
&\Leftrightarrow \limsup_{N\to\infty} P(\widehat{f}_W(w) - N^{\frac{1}{\alpha}-1} A_{k_{p_N}} k_{p_N}^{-\frac{1}{\alpha}} x_{1-\frac{\gamma}{2}} > f_W(w)) \leq \frac{\gamma}{2}. \qquad (4.1.10)
\end{aligned}$$

Combining (4.1.9) and (4.1.10), we get (4.1.8). □

4.2 Examples

In this section, we consider two examples illustrating Theorem 4.1.1.

Example 4.2.1. Consider the case where W follows a geometric distribution, that is, $f_W(w) = c^{w-1}(1-c)$, $w = 1, 2, 3, \ldots$ and $0 < c < 1$. Substituting this into (1.1.1) leads to

$$f_{W_q}(s) = \sum_{w=s}^{\infty} \binom{w}{s} q^s (1-q)^{w-s} c^{w-1}(1-c). \tag{4.2.1}$$

When $s = 0$, we get

$$f_{W_q}(0) = \sum_{w=1}^{\infty} (1-q)^w c^{w-1}(1-c) = \frac{(1-q)(1-c)}{1-c(1-q)}. \tag{4.2.2}$$

When $s \geq 1$, on the other hand, we have

$$\begin{aligned} f_{W_q}(s) &= q^s c^{s-1}(1-c)\sum_{w=s}^{\infty}\binom{w}{s}(c(1-q))^{w-s} \\ &= \frac{q^s c^{s-1}(1-c)}{(1-c(1-q))^{s+1}} = \frac{c_q}{c}c_q^{s-1}(1-c_q), \end{aligned} \tag{4.2.3}$$

where $c_q = \frac{qc}{1-c(1-q)}$, by using the identity $\sum_{w=s}^{\infty}\binom{w}{s}x^{w-s} = \sum_{r=0}^{\infty}\binom{s+r}{r}x^r = (1-x)^{-(s+1)}$. Hence, for $x \geq 1$,

$$P\left(\frac{W_q}{2} \geq x, W_q \text{ is even}\right) = \sum_{s=\lceil x\rceil}^{\infty}\frac{c_q}{c}c_q^{2s-1}(1-c_q) = \frac{1}{c}\frac{c_q^{2\lceil x\rceil}}{1+c_q} \tag{4.2.4}$$

and

$$P\left(\frac{W_q-1}{2} \geq x, W_q \text{ is odd}\right) = \sum_{s=\lceil x\rceil}^{\infty}\frac{c_q}{c}c_q^{2s}(1-c_q) = \frac{c_q}{c}\frac{c_q^{2\lceil x\rceil}}{1+c_q}. \tag{4.2.5}$$

Thus, the conditions (3.1.1)–(3.1.4) in Theorem 3.1.1 are satisfied with $\nu = 2\log\frac{1}{c_q}$, $h_1(\lceil x\rceil) = \frac{1}{c(1+c_q)}$, $h_2(\lceil x\rceil) = \frac{c_q}{c(1+c_q)}$ with $\frac{h_2(x)}{h_1(x)} = c_q$. By using the expression of β in (4.1.3), the parameter α appearing in (3.1.81) or (4.1.1) is given by

$$\alpha = \frac{\log\frac{1}{c_q}}{\log(q^{-1}-1)} = \frac{\log\frac{1-c(1-q)}{cq}}{\log(q^{-1}-1)}.$$

Note that $c_q < 1$ and hence $\log\frac{1}{c_q} > 0$. Then, $\alpha > 0$ is possible only when $q \in (0, 0.5)$. In particular, for $q \in (0, 0.5)$,

$$1 < \alpha < 2 \Leftrightarrow \frac{q}{1-q} < c < \frac{1}{2(1-q)}, \tag{4.2.6}$$

$$0 < \alpha < 1 \Leftrightarrow \frac{1}{2(1-q)} < c < 1. \tag{4.2.7}$$

Theorem 4.1.1 can now be applied in these two cases with

$$A_{k_n} = \binom{2n-2}{w}(1-q)^{-w}\big(q^{-1}-1\big)^{2n-2} \text{ and } k_n = \left\lceil \frac{c(1+c_q)}{c_q^{2n-2}} \right\rceil.$$

Remark. Under (4.2.6) or (4.2.7), and $q \in (0, 0.5)$, the limit of $\widehat{f}(w)$ involves a semi-stable distribution. On the other hand, as proved in Antunes and Pipiras [2], $\widehat{f}(w)$ is asymptotically normal if $R_{q,w} < \infty$, where $R_{q,w}$ is given in (1.1.9). This condition obviously holds when $q \in (0.5, 1)$ (and also for $q = 0.5$ by recalling from Example 4.2.1 above that $f_{W_q}(s) \sim C c_q^s$ as $s \to \infty$). To understand when $R_{q,w} < \infty$ for $q \in (0, 0.5)$, observe that

$$\begin{aligned} R_{q,w} &= \sum_{k=w}^{\infty} f_W(k)(1-q)^{k-2w}\binom{k}{w}\sum_{s=w}^{k}\binom{s}{w}\binom{k-w}{s-w}\left(\frac{1}{q}-1\right)^s \\ &= \sum_{s=w}^{\infty}\binom{s}{w}(q^{-1}-1)^s\sum_{k=s}^{\infty} c^{k-1}(1-c)(1-q)^{k-2w}\binom{k}{w}\binom{k-w}{s-w}. \end{aligned} \quad (4.2.8)$$

Since

$$\binom{k}{w}\binom{k-w}{s-w} = \frac{k!}{w!(k-w)!}\frac{(k-w)!}{(s-w)!(k-s)!} = \frac{k!}{(k-s)!s!}\frac{s!}{w!(s-w)!} = \binom{k}{s}\binom{s}{w},$$

we have

$$
\begin{aligned}
R_{q,w} &= (1-c)\sum_{s=w}^{\infty}\binom{s}{w}^2(q^{-1}-1)^s\sum_{k=s}^{\infty}\binom{k}{s}c^{k-1}(1-q)^{k-2w}\\
&= (1-c)\sum_{s=w}^{\infty}\binom{s}{w}^2(q^{-1}-1)^s\sum_{k=s}^{\infty}\binom{k}{s}(c(1-q))^{k-s}c^{s-1}(1-q)^{s-2w}\\
&= (1-c)\sum_{s=w}^{\infty}\binom{s}{w}^2(q^{-1}-1)^s c^{s-1}(1-q)^{s-2w}\sum_{k=s}^{\infty}\binom{k}{k-s}(c(1-q))^{k-s}\\
&= (1-c)\sum_{s=w}^{\infty}\binom{s}{w}^2(q^{-1}-1)^s c^{s-1}(1-q)^{s-2w}(1-c(1-q))^{-(s+1)}\\
&= \left(\frac{1-c}{c}\right)\sum_{s=w}^{\infty}\binom{s}{w}^2(q^{-1}-1)^s(c(1-q))^s(1-q)^{-2w}(1-c(1-q))^{-(s+1)}\\
&= d_w\sum_{s=w}^{\infty}\binom{s}{w}^2\left((q^{-1}-1)\frac{c(1-q)}{1-c(1-q)}\right)^s, \qquad (4.2.9)
\end{aligned}
$$

where $d_w = \left(\frac{1-c}{c}\right)(1-q)^{-2w}\frac{1}{1-c(1-q)}$. Thus, $R_{q,w} < \infty$ if and only if

$$
(q^{-1}-1)\frac{c(1-q)}{1-c(1-q)} < 1 \Leftrightarrow c < \frac{q}{1-q}. \qquad (4.2.10)
$$

Apart from the boundary cases $c = \frac{q}{1-q}$ and $c = \frac{1}{2(1-q)}$, the ranges of c given in (4.2.6), (4.2.7) and (4.2.10) now cover the whole permissible interval $c \in (0,1)$.

Example 4.2.2. Consider the case where W follows a negative binomial distribution, that is, $f_W(w) = \binom{w-1}{r-1}c^{w-r}(1-c)^r$, $w = r, r+1, \ldots$, $0 < c < 1$. We first compute $f_{W_q}(s)$. One can write $W = G_1 + G_2 + \ldots + G_r$, where $G_1, G_2, \ldots, G_r$ are i.i.d. geometric random variables with p.m.f. $f_{G_1}(w) = c^{w-1}(1-c), w \geq 1$, and hence $W_q = G_1' + G_2' + \ldots + G_r'$, where $G_1', G_2', \ldots, G_r'$ are i.i.d. random variables following the distribution given in (4.2.2)–(4.2.3). Hence,

$$
f_{W_q}(0) \;=\; \left\{\frac{(1-q)(1-c)}{1-c(1-q)}\right\}^r. \qquad (4.2.11)
$$

For $s \geq 1$, we have

$$f_{W_q}(s) = \sum_{i_1, i_2, \ldots, i_r \geq 0, i_1+i_2+\ldots+i_r = s} P(G_1^{'} = i_1) P(G_2^{'} = i_2) \ldots P(G_r^{'} = i_r).$$

To evaluate this quantity, let

$$p_j^r = \sum_{i_{j+1}, i_{j+2}, \ldots, i_r \geq 1, i_{j+1}+i_{j+2}+\ldots+i_r = s} P(G_{j+1}^{'} = i_{j+1}) P(G_{j+2}^{'} = i_{j+2}) \ldots P(G_r^{'} = i_r), \tag{4.2.12}$$

for $0 \leq j < r$. Then, by using (4.2.2),

$$f_{W_q}(s) = \sum_{j=0}^{r-1} \binom{r}{j} \left\{ \frac{(1-q)(1-c)}{1-c(1-q)} \right\}^j p_j^r.$$

Now, by using (4.2.3),

$$\begin{aligned} p_j^r &= \left(\frac{c_q(1-c_q)}{c} \right)^{r-j} c_q^{s-(r-j)} \sum_{i_{j+1}, i_{j+2}, \ldots, i_r \geq 1, i_{j+1}+i_{j+2}+\ldots+i_r = s} 1 \\ &= \left(\frac{1-c_q}{c} \right)^{r-j} c_q^s \binom{s-1}{r-j-1}. \end{aligned}$$

Hence, for $s \geq 1$,

$$\begin{aligned} f_{W_q}(s) &= c_q^s \sum_{j=0}^{r-1} \left\{ \frac{(1-q)(1-c)}{1-c(1-q)} \right\}^j \left(\frac{1-c_q}{c} \right)^{r-j} \binom{r}{j} \binom{s-1}{r-j-1} \\ &= c_q^{s-1} p^*(s), \end{aligned}$$

where $p^*(s)$ is a polynomial given as

$$p^*(s) = \sum_{i=1}^{r-1} a_i^* s^i.$$

This implies that for $x > 1$,

$$P\left(\frac{W_q}{2} \geq x, W_q \text{ is even}\right) = \sum_{s=\lceil x \rceil}^{\infty} c_q^{2s-1} p^*(2s) = c_q^{2\lceil x \rceil} \sum_{s=\lceil x \rceil}^{\infty} c_q^{2s-2\lceil x \rceil -1} p^*(2s) \quad (4.2.13)$$

and

$$P\left(\frac{W_q - 1}{2} \geq x, W_q \text{ is odd}\right) = \sum_{s=\lceil x \rceil}^{\infty} c_q^{2s} p^*(2s+1) = c_q^{2\lceil x \rceil} \sum_{s=\lceil x \rceil}^{\infty} c_q^{2s-2\lceil x \rceil} p^*(2s+1). \quad (4.2.14)$$

Thus the conditions (3.1.1)–(3.1.2) in Theorem 3.1.1 are satisfied with $\nu = 2\log\frac{1}{c_q}$, $h_1(x) = \sum_{k=0}^{\infty} c_q^{2k-1} p^*(2x+2k)$, $h_2(x) = \sum_{k=0}^{\infty} c_q^{2k} p^*(2x+1+2k)$. The conditions (3.1.3)–(3.1.4) also hold with $c_1 = c_q$. The parameter α appearing in (3.1.81) is given by

$$\alpha = \frac{\log\frac{1}{c_q}}{\log(q^{-1}-1)} = \frac{\log\frac{1-c(1-q)}{cq}}{\log(q^{-1}-1)}.$$

Note that $c_q < 1$ and hence $\log\frac{1}{c_q} > 0$. Then, $\alpha > 0$ is possible only when $q \in (0, 0.5)$. In particular, for $q \in (0, 0.5)$, the two cases (4.2.6)–(4.2.7) can be considered. Theorem 4.1.1 can now be applied in these two cases with

$$A_{k_n} = \binom{2n-2}{w}(1-q)^{-w}\left(q^{-1}-1\right)^{2n-2} \text{ and } k_n = \left\lceil \frac{1}{c_q^{2n-2} h_1(n-1)} \right\rceil.$$

4.3 Performance of confidence intervals

In this section, we are interested in assessing the performance of the conservative confidence interval for $f(w)$ derived in Proposition 4.1.1. We focus on the 95% confidence intervals. As considered in Example 4.2.1, we assume that W follows a geometric distribution with parameter $(1-c)$ where $0 < c < 1$. For the purpose of numerical calculation of the limiting semi-stable distribution obtained in Example

4.2.1, we take the sampling probability $q = 0.2689$ and $c = 0.5$. These two parameters control the values of all the other parameters such as α, β, ν. For this particular case, the values are $\beta = 1.002$, $\nu = 3.1031$ and $\alpha = 1.5512$. Note that $\widehat{f}_W$ is a consistent estimator of f_W when $1 < \alpha < 2$. Note also that the values of q and c are chosen in such a way that β is close to 1. The main reason for choosing β close 1 is that in this case, the numerically calculated limiting semi-stable distribution approximates the empirical distribution well (see Sections 2.3 and 3.2).

The parameter λ appearing in Proposition 4.1.1 can take values in $[1, e^{\nu}] = [1, 22.2676]$ for this particular problem. We calculate numerically the semi-stable densities for various values of λ within the stated interval and then for each of these densities, we calculate critical values of $x^{\lambda}_{0.025}$ and $x^{\lambda}_{0.975}$ as stated in Proposition 4.1.1. In Figure 4.1, we give a plot of the numerically calculated semi-stable densities for four different values of λ. The critical values $x^{\lambda}_{0.025}$, $x^{\lambda}_{0.975}$ for a number of values λ

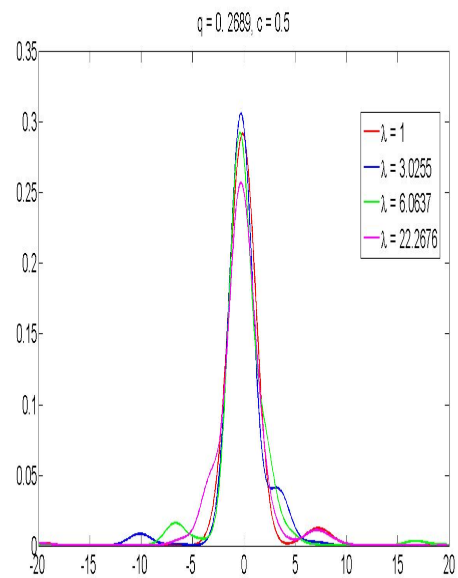

Figure 4.1: Numerically calculated semi-stable distributions

are given in Table 4.1. From Table 4.1 and after using (4.1.7), we get the critical values $x_{0.025} = -8.8$ and $x_{0.975} = 8.4$.

Table 4.1: Critical values $x^{\lambda}_{0.025}$ and $x^{\lambda}_{0.0975}$ for various values of λ

λ	$x^{\lambda}_{0.025}$	$x^{\lambda}_{0.0975}$
1	-3	7.1
2.0127	-3.4	5.2
3.0255	-8.8	4.5
4.0382	-8.1	4.2
5.0510	-7.2	4.4
6.0637	-6.6	4.5
7.0765	-6.3	4.5
8.0892	-5.9	4.5
9.1019	-5.7	4.7
10.1147	-5.5	5
11.1274	-5.3	5.5
12.1402	-5.2	6.8
13.1529	-5.1	8.3
14.1657	-5	8.4
15.1784	-4.9	8.3
16.1911	-4.9	8.1
17.2039	-4.8	8
18.2166	-4.8	7.8
19.2294	-4.7	7.6
20.2421	-4.7	7.4
21.2548	-4.8	7.2
22.2676	-4.8	6.9

In order to calculate the confidence interval for $f_W(w)$, we still need to calculate $\tilde{b}_N$ appearing in (4.1.6). We consider the sample sizes $N = 7000, 8000, 9000, 10000$ and fix w in $f_W(w)$ to be 2. Then, p_N appearing in (4.1.6) is 4 under the condition that $k_{p_N} \leq N < k_{p_N+1}$. In particular, $k_{p_N} = 6691$ and $A_{k_{p_N}} = 6059.1$. The value of $\tilde{b}_N$ for each value of N can now be calculated.

Finally, Table 4.2 concerns the performance of the constructed confidence intervals for finite samples $N = 7000, 8000, 9000, 10000$. The success rate is reported based on the number of times $\widehat{f}_W(w)$ falls in the confidence interval in 100 replications. For each replication, the estimator $\widehat{f}_W(w)$ is computed using the expression (1.1.4)–(1.1.5). The reported performance of the confidence intervals is very good.

Table 4.2: Sample size and success rate

N	success rate
7000	96
8000	95
9000	94
10000	95

CHAPTER 5: EXTENSIONS AND FUTURE DIRECTIONS

In this chapter, we discuss briefly several extensions of our results and also raise several open problems for future directions.

In Chapter 4, we studied the asymptotic behavior of the estimator $\widehat{f}_W(w)$ of the distribution of flow sizes for *fixed* w. One interesting extension is to consider the joint behavior of $\widehat{f}_W(w)$ across *multiple* w's. We provide next several comments on this extension.

For simplicity, consider two w's, w_1 and w_2 satisfying $w_1 < w_2$. Suppose also that $\alpha \in (1,2)$, where α is given in (4.1.1). Then, by Theorem 4.1.1, for $j = 1, 2$

$$d_{j,N}(\widehat{f}_W(w_j) - f_W(w_j)) \xrightarrow{d} (-1)^{w_j}(Y + \zeta), \tag{5.0.1}$$

where the convergence takes place along the subsequence $N = k_n$ as in Theorem 4.1.1 and $d_{j,N} = \frac{k_n}{A_{j,k_n}}$ with

$$A_{j,k_n} = \binom{2n-2}{w_j}(1-q)^{-w_j}\left(q^{-1}-1\right)^{2n-2}. \tag{5.0.2}$$

Since $w_1 < w_2$, note that

$$\frac{\binom{2n-2}{w_1}}{\binom{2n-2}{w_2}} \to 0,$$

as $n \to \infty$ and hence that

$$\frac{d_{2,N}}{d_{1,N}} \to 0, \tag{5.0.3}$$

as $N \to \infty$. The relations (5.0.1) and (5.0.3) imply that

$$d_{2,N} \begin{pmatrix} \widehat{f}_W(w_1) - f_W(w_1) \\ \widehat{f}_W(w_2) - f_W(w_2) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} 0 \\ (-1)^{-w_2}(Y + \zeta) \end{pmatrix}. \tag{5.0.4}$$

The joint convergence (5.0.4), however, is not particularly interesting since the first entry in the vector on the left-hand side of (5.0.4) has a too small normalization.

Instead of (5.0.4), another natural possibility is to consider the asymptotic of

$$Y_N = \begin{pmatrix} Y_{1,N} \\ Y_{2,N} \end{pmatrix} := \begin{pmatrix} d_{1,N}(\widehat{f}_W(w_1) - f_W(w_1)) \\ d_{2,N}(\widehat{f}_W(w_2) - f_W(w_2)) \end{pmatrix}, \tag{5.0.5}$$

so that the first entry $Y_{1,N}$ has now the correct normalization. Dealing with (5.0.5), however, is expected to be different from the univariate case.

Indeed, note that the convergence of Y_N in distribution is equivalent to the convergence in distribution of linear combinations

$$\theta_1 Y_{1,N} + \theta_2 Y_{2,N} =: Z_{\theta_1,\theta_2,N} - EZ_{\theta_1,\theta_2,N} \tag{5.0.6}$$

for any $\theta_1, \theta_2 \in \mathbb{R}$. Observe from (1.1.12) that

$$Z_{\theta_1,\theta_2,N} = \sum_{k=1}^{N} (\theta_1 \frac{X_{1,k}}{A_{1,N}} + \theta_2 \frac{X_{2,k}}{A_{2,N}}), \tag{5.0.7}$$

where

$$X_{j,k} = \binom{W_{q,k}}{w_j} (-1)^{W_{q,k}-w_j} \frac{(1-q)^{W_{q,k}-w_j}}{q^{W_{q,k}}} 1_{\{W_{q,k} \geq w_j\}}. \tag{5.0.8}$$

Hence,

$$Z_{\theta_1,\theta_2,N} = \sum_{k=1}^{N} e^{\beta W_{q,k}} (-1)^{W_{q,k}-w_j} \mathcal{L}_{\theta_1,\theta_2,N}(e^{W_{q,k}}), \tag{5.0.9}$$

where $\beta = \log(q^{-1} - 1)$ and

$$\mathcal{L}_{\theta_1,\theta_2,N}(x) = \sum_{j=1}^{2} \frac{\theta_j}{A_{j,N}} \binom{\log x}{w_j} (1-q)^{w_j} 1_{\{\log x \geq w_j\}}. \tag{5.0.10}$$

The setting (5.0.9)–(5.0.10) appears similar to that of Theorem 2.1.1. For example, one can show that the function $\mathcal{L}_{\theta_1,\theta_2,N}$ is an ultimately monotone, slowly varying function (after a possible change of sign to make it positive). The key difference, however, is that $\mathcal{L}_{\theta_1,\theta_2,N}$ also depends on N. The sequence $Z_{\theta_1,\theta_2,N}$ thus can not be dealt with using the available results in the domains of attraction of semi-stable distributions. Since $Z_{\theta_1,\theta_2,N}$ involves a triangular array of i.i.d. random variables, it is nevertheless expected to be dealt with using the classical results on the convergence of partial sums of such arrays.

A related direction for future work is to study the domains of attraction of multivariate semi-stable distributions. To the best of our knowledge, a characterization of such domains is not available in the multivariate case, in contrast to the one-dimensional context (Theorems 2.1.1 and 2.1.2). The next step would then be to provide sufficient conditions as we did in Theorem 3.1.1 and Corollary 3.1.1 for the univariate case.

Among perhaps less ambitious open problems, we finally note the following. It would be interesting to understand how close the sufficient conditions of Theorem 3.1.1 and Corollary 3.1.1 are to being necessary (for a distribution to be in the domain of attraction of a semi-stable law). It would also be important to shed light on why approximations of partial sums by semi-stable distributions deteriorates as the constant c in (2.1.5) increases (Sections 2.2 and 3.2) and how this can be remedied.

APPENDIX A: AUXILIARY RESULTS AND CODE

A.1 Auxiliary results

We state and prove here a number of auxiliary results used in Section 3.1.

Lemma A.1.1. *Let g_1 and $\tilde{g}_1$ be defined in (3.1.21) and (3.1.25), respectively. Then, $\tilde{g}_1(y) - g_1(y) \to 0,$ as $y \to \infty$.*

Proof. For $n \geq 2$, if

$$n - 1 + \frac{1}{2\beta} \log L(e^{2n-2}) \leq y < n - 1 + \frac{1}{2\beta} \log L\big(e^{2n}\big),$$

then

$$0 \leq \tilde{g}_1(y) - g_1(y) < \frac{1}{2\beta} \log \frac{L\left(e^{2n}\right)}{L\left(e^{2n-2}\right)} \to 0, \text{ as } y \to \infty \quad (n \to \infty), \tag{A.1.1}$$

since L is a slowly varying function. If

$$n - 1 + \frac{1}{2\beta} \log L(e^{2n}) \leq y < n + \frac{1}{2\beta} \log L\big(e^{2n}\big),$$

then similarly

$$\tilde{g}_1(y) - g_1(y) = \frac{1}{2\beta} \log \frac{L\left(e^{2n}\right)}{L\left(e^{2n-2}\right)} \to 0, \text{ as } y \to \infty \quad (n \to \infty). \quad \square \tag{A.1.2}$$

Lemma A.1.2. *Let $\tilde{g}_1^*$ be defined in (3.1.27). Then, for any $A > 0$,*

$$\tilde{g}_1^*\left(\log Ax\right) - \tilde{g}_1^*\left(\log x\right) \to 0, \text{ as } x \to \infty.$$

Proof. Suppose without loss of generality that $A > 1$. First, note that

$$\begin{aligned} \tilde{g}_1^*(\log Ax) - \tilde{g}_1^*(\log x) &= \frac{1}{2\beta}\left(\log L\left(e^{2n_{Ax}-2}\right) - \log L\left(e^{2n_x-2}\right)\right) \\ &= \frac{1}{2\beta}\log\frac{L\left(e^{n_{Ax}-2}\right)}{L\left(e^{n_x-2}\right)}, \\ &= \frac{1}{2\beta}\log\frac{L\left(e^{2n_{Ax}-2n_x}e^{2n_x-2}\right)}{L\left(e^{n_x-2}\right)}, \end{aligned} \tag{A.1.3}$$

where, for y ($= x$ or Ax),

$$n_y - 1 + \frac{1}{2\beta}\log L(e^{2n_y-2}) \leq \log y < n_y + \frac{1}{2\beta}\log L(e^{2n_y}).$$

Observe that $n_{Ax} - n_x$ takes only positive integer values, and that

$$0 \leq n_{Ax} - n_x \leq \lceil \log A \rceil.$$

Hence, by Theorem 1.2.1 of Bingham, Goldie and Teugels [3],

$$\frac{L\left(e^{n_{Ax}-n_x}e^{n_x-1}\right)}{L\left(e^{n_x-1}\right)} \to 1, \text{ as } e^{n_x-1} \to \infty \quad (\text{or } x \to \infty).$$

This yields the result. $\square$

Lemma A.1.3. *The function $l_1^*(x)$ defined in (3.1.32) is right-continuous and slowly varying at ∞.*

Proof. To show that $l_1^*(x)$ is slowly varying, write

$$l_1^*(x) = \frac{h_1\left(\lceil g_2(\frac{1}{2\beta}\log x)\rceil_+\right)}{h_1\left(g_2(\frac{1}{2\beta}\log x)\right)} h_1\left(g_2(\frac{1}{2\beta}\log x)\right) e^{\nu\tilde{g}_1^*(\frac{1}{2\beta}\log x)} e^{-\nu(g_1(\frac{1}{2\beta}\log x) - \tilde{g}_1(\frac{1}{2\beta}\log x))}.$$

Note that

$$\frac{h_1\left(\lceil g_2(\frac{1}{2\beta}\log x)\rceil_+\right)}{h_1\left(g_2(\frac{1}{2\beta}\log x)\right)} = \frac{h_1\left(\frac{\lceil g_2(\frac{1}{2\beta}\log x)\rceil_+}{g_2(\frac{1}{2\beta}\log x)} g_2(\frac{1}{2\beta}\log x)\right)}{h_1\left(g_2(\frac{1}{2\beta}\log x)\right)} \to 1$$

by using (3.1.4), since $g_2(\frac{1}{2\beta}\log x) \to \infty$ and

$$\frac{\lceil g_2(\frac{1}{2\beta}\log x)\rceil_+}{g_2(\frac{1}{2\beta}\log x)} \to 1, \qquad \text{as } x \to \infty.$$

By Lemma A.1.1, we also have

$$e^{-\nu(g_1(\frac{1}{2\beta}\log x) - \tilde{g}_1(\frac{1}{2\beta}\log x))} \to 1, \qquad \text{as } x \to \infty.$$

Hence, $l_1^*(x)$ is asymptotically equivalent to

$$h_1\left(g_2(\frac{1}{2\beta}\log x)\right) e^{\nu \tilde{g}_1^*(\frac{1}{2\beta}\log x)}. \tag{A.1.4}$$

It is enough to show that the function (A.1.4) is slowly varying. By using Lemma A.1.2, we have

$$\frac{e^{\nu \tilde{g}_1^*(\frac{1}{2\beta}\log Ax)}}{e^{\nu \tilde{g}_1^*(\frac{1}{2\beta}\log x)}} \to 1, \qquad \text{as } x \to \infty. \tag{A.1.5}$$

It remains to show that $h_1(g_2(\frac{1}{2\beta}\log x))$ is a slowly varying function. For $A > 0$,

$$\frac{h_1(g_2(\frac{1}{2\beta}\log Ax))}{h_1(g_2(\frac{1}{2\beta}\log x))} = \frac{h_1\left(\frac{g_2(\frac{1}{2\beta}\log Ax)}{g_2(\frac{1}{2\beta}\log x)} g_2(\frac{1}{2\beta}\log x)\right)}{h_1(g_2(\frac{1}{2\beta}\log x))}. \tag{A.1.6}$$

Now, by using (3.1.28),

$$
\begin{aligned}
\frac{g_2(\frac{1}{2\beta}\log Ax)}{g_2(\frac{1}{2\beta}\log x)} &= \frac{\frac{1}{2\beta}\log Ax + g_2^*(\frac{1}{2\beta}\log Ax)}{\frac{1}{2\beta}\log x + g_2^*(\frac{1}{2\beta}\log x)} \\
&= 1 + \frac{\frac{1}{2\beta}\log Ax + g_2^*(\frac{1}{2\beta}\log Ax) - \frac{1}{2\beta}\log x - g_2^*(\frac{1}{2\beta}\log x)}{\frac{1}{2\beta}\log x + g_2^*(\frac{1}{2\beta}\log x)} \\
&= 1 + \frac{\frac{1}{2\beta}\log A + g_2^*(\frac{1}{2\beta}\log Ax) - g_2^*(\frac{1}{2\beta}\log x)}{g_2(\frac{1}{2\beta}\log x)} \to 1,
\end{aligned}
$$

since $g_2(\frac{1}{2\beta}\log x) \to \infty$ and by using (3.1.29), $g_2^*(\frac{1}{2\beta}\log Ax) - g_2^*(\frac{1}{2\beta}\log x) \to 0$. Thus, by using (3.1.4) and (A.1.6), we have

$$
\frac{h_1(g_2(\frac{1}{2\beta}\log Ax))}{h_1(g_2(\frac{1}{2\beta}\log x))} \to 1, \qquad \text{as } x \to \infty.
$$

This completes the proof that $l_1^*(x)$ is a slowly varying function.

The function $l_1^*(x)$ is right-continuous since $h_1(x)$ can be defined to be continuous, g_2 is continuous (as the inverse of a continuous increasing function) and g_1, $\tilde{g}_1$ and $\tilde{g}_1^*$ are right-continuous functions. $\square$

Lemma A.1.4. *Let L be a slowly varying function. Then, for any fixed $x_0 \neq e^{2\beta(r+1-b_1)}$,$r \in \mathbb{Z}$, $\beta > 0$, there are only finitely many integer values of n for which*

$$
m - b_1 + \frac{1}{2\beta}\log L(e^{2m-b_2}) \leq \frac{1}{2\beta}\log(A_{k_n}x_0) < m - b_1 + \frac{1}{2\beta}\log L(e^{2m-b_3}), \quad \text{(A.1.7)}
$$

where $A_{k_n} = e^{(n-1)2\beta}L(e^{2n-2})$, m takes positive integer values, b_1, b_2 and b_3 are fixed positive constants with $b_2 > b_3$.

Proof. Suppose $m = n + r_n$, where r_n is a sequence of integers. We first show that if (A.1.7) is satisfied for infinitely many values of n, then $\sup_{n\geq 1}|r_n| < \infty$. Arguing by contradiction, for example, assume $r_n \to \infty$ as $n \to \infty$. From (A.1.7), we need to have

$$e^{2\beta(r_n+1-b_1)}\frac{L(e^{2n+2r_n-b_2})}{L(e^{2n-2})} \le x_0 < e^{2\beta(r_n+1-b_1)}\frac{L(e^{2n+2r_n-b_3})}{L(e^{2n-2})}. \tag{A.1.8}$$

A standard argument using Potter's bounds for L shows that $e^{2\beta(r_n+1-b_1)}\frac{L(e^{2n+2r_n-b})}{L(e^{2n-2})} \to \infty$, ($b = b_2$ or b_3) when $r_n \to \infty$. Since x_0 is fixed, this leads to a contradiction. A similar argument can be applied when $r_n \to -\infty$.

Next we show that m is necessarily of the form $m = n + r$ where r is a fixed integer for large enough n. We prove this by contradiction. First observe that r_n can only take finitely many integer values. Now if r_n has a subsequence $r_{n_k} \to r$, then letting $n \to \infty$ in (A.1.8), we have $e^{2\beta(r+1-b_1)} = x_0$. Thus, r is determined uniquely and since r_n are integers, we have that $r_n = r$ for large enough n.

Finally, if $m = n + r$, then (A.1.7) cannot hold for infinitely many values of n unless $x_0 = e^{2\beta(r+1-b_1)}$. This proves the lemma. $\square$

Lemma A.1.5. *Let (2.1.7)–(2.1.8) hold for a random variable X with $l^*(x)$ replaced by a right-continuous slowly varying function $l_1^*(x)$ in (2.1.7). Then, $l^*(x)$ in (2.1.8) can be replaced by another right-continuous function $l_2^*(x)$ if $\frac{l_2^*(x)}{l_1^*(x)} \to 1$ as $x \to \infty$.*

Proof. Observe that

$$\begin{aligned}
1 - F(x) &= x^{-\alpha} l_2^*(x)(M_R(\delta(x)) + h_R(x)) \\
&= x^{-\alpha} l_1^*(x)\Big(M_R(\delta(x)) + h_R(x) + (\frac{l_2^*(x)}{l_1^*(x)} - 1)(M_R(\delta(x)) + h_R(x))\Big) \\
&= x^{-\alpha} l_1^*(x)(M_R(\delta(x)) + h_R(x) + \tilde{h}_R(x)),
\end{aligned} \tag{A.1.9}$$

where

$$\tilde{h}_R(x) = (\frac{l_2^*(x)}{l_1^*(x)} - 1)(M_R(\delta(x)) + h_R(x)). \tag{A.1.10}$$

Since $\frac{l_2^*(x)}{l_1^*(x)} \to 1$ as $x \to \infty$, M_R is a bounded periodic function from (2.1.2) and $h_R(A_{k_n}x) \to 0$, as $n \to \infty$, we have $\tilde{h}_R(A_{k_n}x) \to 0$ for every continuity point x of $M_R(x)$. Hence, in (A.1.9), one can take the new error function to be $h_R(x) + \tilde{h}_R(x)$. Hence the result is proved. $\square$

Lemma A.1.6. *Let h_1 be the function defined in Theorem 3.1.1 and satisfying (3.1.4). For every $\delta > 0$, there is M_δ such that, for all $n > M_\delta$,*

$$h_1(M_\delta + 1)e^{M_\delta+1}e^{\delta n} < h_1(n) < \frac{h_1(M_\delta + 1)}{e^{\delta(M_\delta+1)}}e^{\delta n}.$$

Proof. Fix any $\delta = \delta_0 \in (0, 1)$. By using (3.1.4), there exists M_{δ_0} such that for all $m > M_{\delta_0}$, $1 - \delta_0 < \frac{h_1(m+1)}{h_1(m)} < 1 + \delta_0$. Take any $n > M_{\delta_0}$. Then,

$$\begin{aligned} h_1(n) &= \frac{h_1(n)}{h_1(n-1)}\frac{h_1(n-1)}{h_1(n-2)}\cdots\frac{h_1(M_{\delta_0}+2)}{h_1(M_{\delta_0}+1)}h_1(M_{\delta_0}+1) \\ &< h_1(M_{\delta_0}+1)(1+\delta_0)^{n-M_{\delta_0}-1} < h_1(M_{\delta_0}+1)e^{\delta_0(n-M_{\delta_0}-1)}. \end{aligned}$$

Similarly,

$$h_1(n) > h_1(M_{\delta_0}+1)(1-\delta_0)^{n-M_{\delta_0}-1} > h_1(M_{\delta_0}+1)e^{-\delta_0(n-M_{\delta_0}-1)}. \quad \square$$

A.2 Code for numerical calculation

We use seven different Matlab functions to numerically calculate the densities of the semi-stable distributions appearing in this dissertation. The first function **laplace_function_exp_positive.m** calculates the exponent of the Laplace transform of the density with the Lévy measure M_R in (3.1.10) having positive support.

```
function [f] = laplace_function_exp_positive(x,y,nu,beta,q,lambda);
eta = (1 - exp((-1)*nu))*(1/(1 - exp(2*beta - nu)))*(1 - (q)*exp(beta + (nu
     - 2*beta)*ceil((1/nu)*log(q)))) - exp((2*ceil((1/nu)*log(q)) - 1)*beta
    )*(1 - (q)*exp((-1)*nu*ceil((1/nu)*log(q))));
z = sqrt((-1)); %definition of complex i.
multi = x+z*y;
theta_psi=0; %initialization of thetapsi
k = [1:1:60];
theta_psi = sum(lambda * (-1)*(exp(k*2*beta)./(1+exp(4*k*beta)).*exp((-1)*
    nu*k)*(1 - exp((-1)*nu)) - exp(-6*k*beta)./(1+exp(-4*k*beta)).*exp(nu*k
    )*(1 - exp((-1)*nu))));
if ((y~=0))
    indi=0;
else
    indi=1;
end
theta_psi = theta_psi + eta*lambda*indi;
k = [1:1:60];
part1 = lambda*(exp((-1)*multi*exp(2*k*beta)) - 1 + multi*exp(2*k*beta)
    ./(1+exp(4.*k*beta)));
part2 = (1 - exp((-1)*nu))*exp((-1)*nu.*k);
sum1=0;
sum2=0;
sum1 = sum(part1.*part2);
barrier = floor((-1)*1/(2*beta)*log((10^(-10))/multi));
k = [1:1:barrier]; %The sum is obtained when k is negative
part1 = lambda*(exp((-1)*multi*exp((-1)*2*k*beta)) - 1 + multi*exp((-1)*2*k
    *beta)./(1+exp(4.*(-1)*k*beta)));
```

```
part2 = (1 - exp((-1)*nu))*exp((-1)*nu*(-1)*k);
sum2 = sum(part1.*part2);
k = [(barrier+1):1:60];
part1 = lambda*(((-1)*(multi*exp((-1)*6*k*beta)./(1+exp(4*(-1)*k*beta))))+(
    multi^2)*exp(4*(-1)*k*beta)/2);
part2 = (1 - exp((-1)*nu))*exp((-1)*nu*(-1)*k);
sum2 = sum2 + sum(part1.*part2);
f = sum1+sum2 - theta_psi*multi;
```

The second function **laplace_function_positive.m** calculates the Laplace transform of a density by making an appropriate location shift on the transformation.

```
function [f] = laplace_function_positive(x,y,nu,beta,q,shift,lambda) ;
result=0;
result = laplace_function_exp_positive(x,y,nu,beta,q,lambda) ;
z = sqrt(-1);
multi = x+z*y;
f = real(exp(result-shift*multi));
```

We use the third function **euler_ILT_positive.m** to calculate the inverse Laplace transform for a density function with the support on the positive half-axis. This implements the Abate and Whitt method described in Section 2.2.

```
function [density_positive] = euler_ILT_positive(T,nu,beta,shift,prob,
    lambda);
tail_coeffs = zeros(1,12);
% ---- input of extra tail coefficients -------
```

```
tail_coeffs(1) = 1; tail_coeffs(2) = 11; tail_coeffs(3) = 55;
tail_coeffs(4) = 165; tail_coeffs(5) = 330; tail_coeffs(6) = 462;
tail_coeffs(7) = 462; tail_coeffs(8) = 330; tail_coeffs(9) = 165;
tail_coeffs(10) = 55; tail_coeffs(11) = 11; tail_coeffs(12) = 1;
tail_values = zeros(1,12); %initiation of the tail_values which will later
   be calculated using the tail coefficients
A = 3; %initiation of the tuning parameter
Ntr = 100; %initiation of the number of loops
U = exp(A/2)/T;
X = A/(2*T);
H = pi/T;
second_part=0;
first_part = laplace_function_positive(X,0,nu,beta,q,shift,lambda)/2;
for N=1:Ntr
   Y = N*H ;
   second_part = second_part + (-1)^N*laplace_function_positive(X,Y,nu,
      beta,q,shift,lambda);
end;
tail_values(1) = first_part+second_part;
for K=1:12
   N = Ntr+K ;
   Y = N*H ;
   tail_values(K+1) = tail_values(K) + (-1)^N*laplace_function_positive(X,
      Y,nu,beta,q,shift,lambda);
end;
Avgsu = 0;
for J=1:12
   Avgsu = Avgsu + tail_coeffs(J)*tail_values(J);
end;
```

```
density_positive = U*Avgsu/2048;
```

The fourth function **laplace_function_exp_negative.m** calculates the exponent of the Laplace transform with Lévy measure M_L in (3.1.10) having negative support.

```
function [f] = laplace_function_exp_negative(x,y,nu,beta,q,lambda);
eta = (1 - exp((-1)*nu))*(1/(1 - exp(2*beta - nu)))*(1 - (q)*exp(beta + (nu
     - 2*beta)*ceil((1/nu)*log(q)))) - exp((2*ceil((1/nu)*log(q)) - 1)*beta
    )*(1 - (q)*exp((-1)*nu*ceil((1/nu)*log(q))));
z = sqrt((-1)); %definition of complex i.
multi = x+z*y;
theta_psi=0; %initialization of thetapsi
k = [1:1:60];
theta_psi = sum(lambda*q*((-1)*exp((2*(k-2)+3)*beta)./(1+exp((4*(k-2)+6)*
    beta)).*exp((-1)*nu*(k-1))*(1 - exp((-1)*nu))));
k = [3:1:60]; %summation of negative part
theta_psi = theta_psi + sum(lambda*q*exp((-6*k+9)*beta)./(1+exp((-4*k+6)*
    beta)).*exp((-1)*nu*((-1)*k+1))*(1 - exp((-1)*nu)));
theta_psi = theta_psi - lambda*exp((-1)*beta)/(1 + exp((-2)*beta))*(1 - (1/
    q)) + lambda*exp((-3)*beta)/(1 + exp((-2)*beta))*((q)/exp(1) - 1);


if ((y~=0))
    indi=0;
else
    indi=1;
end
theta_psi = theta_psi + eta*lambda*indi;
k = [1:1:60]; %The sum is obtained when k is positive
```

```
part1 = lambda*prob*(exp((-1)*multi*exp((2*k+1)*beta)) - 1 + multi*exp((2*k
    +1)*beta)/(1+exp((4*k+2)*beta)));
part2 = (1 - exp((-1)*nu))*exp((-1)*nu*k);
sum1 = sum(part1.*part2);
barrier = floor((-1)*1/(2*beta)*log((10^(-10))/multi)) - 2;
k = [1:1:barrier];
part1 = lambda*q*(exp((-1)*multi*exp((2*(-1)*k+1)*beta)) - 1 + multi*exp
    ((2*(-1)*k+1)*beta)./(1+exp((4*(-1)*k+2)*beta)));
part2 = (1 - exp((-1)*nu)).*exp((-1)*nu*(-1)*k);
sum2 = sum(part1.*part2);
k = [barrier:1:60];
part1 = lambda*q*(((multi*exp((-1)*(6*k*beta+3))./(1+exp((4*(-1)*k+2)*beta)
    )))+(multi^2)*exp((4*(-1)*k+2)*beta)/2);
part2 = (1 - exp((-1)*nu)).*exp((-1)*nu*(-1)*k);
sum2 = sum2 + sum(part1.*part2);
f = sum1+sum2 - theta_psi*multi;
```

The fifth function **laplace_function_negative.m** calculates the Laplace transform of a density by making an appropriate location shift on the transformation.

```
function [f] = laplace_function_negative(x,y,nu,beta,q,shift,lambda) ;
result=0;
result = laplace_function_exp_negative(x,y,nu,beta,q,lambda) ;
z = sqrt(-1);
multi = x+z*y;
f = real(exp(result-shift*multi));
```

We use the sixth function **euler_ILT_negative.m** to calculate the inverse Laplace transform for a density function with the support on the negative half-axis.

```
function [density_negative] = euler_ILT_negative(T,nu,beta,shift,prob,
    lambda);
tail_coeffs = zeros(1,12);
% ---- input of extra tail coefficients -------
tail_coeffs(1) = 1; tail_coeffs(2) = 11; tail_coeffs(3) = 55;
tail_coeffs(4) = 165; tail_coeffs(5) = 330; tail_coeffs(6) = 462;
tail_coeffs(7) = 462; tail_coeffs(8) = 330; tail_coeffs(9) = 165;
tail_coeffs(10) = 55; tail_coeffs(11) = 11; tail_coeffs(12) = 1;
tail_values = zeros(1,12); %initiation of tail values using tail
    coefficients
A = 3; %initiation of the tuning parameter
Ntr = 100; % initiation of the number of loops
U = exp(A/2)/T;
X = A/(2*T);
H = pi/T;
second_part=0;
first_part = laplace_function_negative(X,0,nu,beta,q,shift,lambda)/2;
for N=1:Ntr
    Y = N*H ;
    second_part = second_part + (-1)^N*laplace_function_negative(X,Y,nu,
        beta,q,shift,lambda);
end;
tail_values(1) = first_part+second_part;
for K=1:12
    N = Ntr+K ;
    Y = N*H ;
```

```
    tail_values(K+1) = tail_values(K) + (-1)^N*laplace_function_negative(X,
        Y,nu,beta,q,shift,lambda);
end;
Avgsu = 0;
for J=1:12
    Avgsu = Avgsu + tail_coeffs(J)*tail_values(J);
end;
density_negative = U*Avgsu/2048;
```

Finally, we give the main function **euler_ILT_expr.m** that is used to convolve the calculated densities for the Lévy measures with the negative and positive supports.

```
clc;
clear all
q = 0.25; %initiation of the sampling probability
beta = 1; %initiation of the beta variable
nu = log(1/(prob^2));
alpha = log(1/prob)/beta;
multi = 0; %initial value of A/2t in our code
lambda = 1;
A=3; %The value of the parameter A
shift = 10;
density_positive = [] ;
for t=.1:.1:30
    Fun_new = euler_ILT_positive(t,nu,beta,shift,q,lambda) ;
    density_positive = [density_positive Fun_new] ;
end ;
density_negative = [] ;
```

```
for t=.1:.1:30
    Fun_new = euler_ILT_negative(t,nu,beta,shift,q,lambda) ;
    density_negative = [density_negative Fun_new] ;
end ;
area_positive = 0;
for (i = 1:299)
   if ((density_positive(i) > 0) & (density_positive(i+1) > 0))
   area_positive = area_positive + 0.05*(density_positive(i) +
       density_positive(i+1));
    end
end
area_negative = 0;
for (i = 1:299)
   if ((density_negative(i) > 0) & (density_negative(i+1) > 0))
   area_negative = area_negative + 0.05*(density_negative(i) +
       density_negative(i+1));
    end
end
convolution = zeros(1,600);
for (l = (-300:300))
   for (k = (max(l,0):300))
   if (((k - l) <= 300) & ((k+1) <= 300) & ((k-l)>0))
   convolution(l+301) = convolution(l+301) + 0.1*density_positive(k+1)*
       density_negative(k-l);
    end
    end
end
total_area = 0;
for (i = 1:599)
```

```
if (convolution(i) > 0 & convolution(i+1) > 0)

total_area = total_area + 0.05*(convolution(i) + convolution(i+1));

end

end
```

BIBLIOGRAPHY

[1] J. Abate and W. Whitt. The Fourier-series method for inverting transforms of probability distributions. *Queueing Systems. Theory and Applications*, 10(1-2):5–87, 1992.

[2] N. Antunes and V. Pipiras. Probabilistic sampling of finite renewal processes. *Bernoulli*, 17(4):1285–1326, 2011.

[3] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular Variation*, volume 27 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 1987.

[4] N. Bouzar. Discrete semi-stable distributions. *Annals of the Institute of Statistical Mathematics*, 56(3):497–510, 2004.

[5] S. Csörgö. A probabilistic approach to domains of partial attraction. *Advances in Applied Mathematics*, 11:282–327, 1990.

[6] S. Csörgö. An approximation to infinitely divisible laws. *Acta Mathematica Hungarica*, 68:221–238, 1995.

[7] S. Csörgö. The St. Petersburg paradox. *Polygon*, 5:19–79, 1995.

[8] S. Csörgö and R. Dodunekova. Limit theorems for St. Petersburg game, sums, trimmed sums and extremes. *Progress in Probability*, 23:285–315, 1991.

[9] S. Csörgö, E. Haeusler, and D. M. Mason. A probabilitistic approach to the asymptotic distributions of the sums of independent indentically distributed random variables. *Advances in Applied Mathematics*, 9:259–333, 1988.

[10] S. Csörgö and P. Kevei. Merging asymptotic expansions for cooperative gamblers in generalized St. Petersburg games. *Acta Mathematica Hungarica*, 121:119–156, 2008.

[11] S. Csörgö and Z. Megyesi. Merging to semistable laws. *Theory of Probability and its Applications*, 47(1):17–33, 2003.

[12] S. Csörgö and G. Simons. On Steinhaus' resolution of the St. Petersburg paradox. *Probability and Mathematical Statistics*, 14:157–171, 1993.

[13] S. Csörgö and G. Simons. A strong law of large numbers for trimmed sums, with applications to generalized St. Petersburg games. *Statistics and Probability Letters*, 26:65–73, 1996.

[14] N. Duffield, C. Lund, and M. Thorup. Estimating flow distributions from sampled flow statistics. *IEEE/ACM Transactions on Networking*, 13(5):933–946, 2005.

[15] V. Grinevich and Y. S. Khokhlov. Domains of attraction of semistable laws. *Theory of Probability and its Applications*, 40(2):361–366, 1995.

[16] N. Hohn and D. Veitch. Inverting sampled traffic. *IEEE/ACM Transactions on Networking*, 14(1):68–80, 2006.

[17] B. W. Huff. On the continuity of infinitely divisible distributions. *Sankhyā (Statistics). The Indian Journal of Statistics. Series A*, 34:443–446, 1972.

[18] K. iti Sato. *Lévy Processes and Infinitely Divisible Distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. Translated from the 1990 Japanese original, Revised by the author.

[19] V. M. Kruglov. On an extension of the class of stable distributions. *Theory of Probability and its Applications*, 17:723–732, 1972.

[20] M. Maejima. Semistable distributions. In *Lévy Processes*, pages 169–183. Birkhäuser Boston, Boston, MA, 2001.

[21] M. M. Meerschaert and H. P. Scheffler. Semistable Lévy motion. *Fractional Calculus and Applied Analysis. An International Journal for Theory and Applications*, 5(1):27–54, 2002.

[22] M. M. Meesrchaert and H. P. Scheffler. A simple robust estimation method for the thickness of heavy tails. *Journal of Statistical Planning and Inference*, 71(1-2):19–34, 1998.

[23] Z. Megyesi. A probabilistic approach to semistable laws and their domains of partial attraction. *Acta Universitatis Szegediensis. Acta Scientiarum Mathematicarum*, 66(1-2):403–434, 2000.

[24] R. N. Pillai. Semistable laws as limit distributions. *The Annals of Mathematical Statistics*, 42:780–783, 1971.

[25] G. Samorodintsky and M. S. Taqqu. *Stable Non-Gaussian Random Processes*. Stochastic Modelling. Chapman & Hall, New York, 1994. Stochastic models with infinite variance.

[26] R. Shimizu. On the domain of partial attraction of semi-stable distributions. *Annals of the Institute of Statistical Mathematics*, 22(2):245–255, 1970.

[27] T. Watanabe and K. Yamamuro. Tail behaviors of semi-stable distribution. *Journal of Mathematical Analysis and Applications*, 393:108–121, 2013.

[28] L. Yang and G. Michailidis. Sampled based estimation of network traffic flow characteristics. In *INFOCOM 2007. 26th IEEE International Conference on Computer Communications*, pages 1775–1783, 2007.

[29] V. M. Zolotarev. *One-Dimensional Stable Distributions*, volume 65 of *Translation of Mathematical Monographs*. American Mathematical Soceity, Providence, RI, 1986. Translated from the Russian by H. H. McFaden, translation edited by Ben Silver.