# Applications of Smoothly Varying Functions and Tail Index Estimation

by

Chenhua Zhang

(Under the direction of William P. McCormick)

## Abstract

This dissertation focuses on heavy tail analysis. The most important classes of heavy tail distributions are the class of regularly varying distribution functions and the class of subexponential distributions. The class of smoothly varying functions, i.e. functions with continuous derivatives being regularly varying at infinity, is a subclass of regularly varying functions. Under an assumption of smooth variation on the step size distribution we obtain higher order expansions for the tail distribution of the global maximum of random walk with negative drift, and under mild regularity conditions that of randomly stopped sums of subexponential random variables with smoothly varying hazard function.

Tail index is the key parameter for distributions with regularly varying tail. We study the asymptotic properties of Hill's estimator, one of the commonly used tail index estimators, for shot noise sequences and first-order bifurcating autoregressive processes.

Index words: Heavy tail analysis, Regularly varying, Subexponential distribution, Tail expansion, Random walk, Wiener-Hopf, Hill's estimator, Shot noise processes, Bifurcating autoregressive processes

# Applications of Smoothly Varying Functions and Tail Index Estimation

by

Chenhua Zhang

B.S. Shanxi University, 1997

M.S., Institute of Applied Mathematics, Chinese Academy of Sciences, 2001

A Dissertation Submitted to the Graduate Faculty
of The University of Georgia in Partial Fulfillment
of the
Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2007

# Applications of Smoothly Varying Functions and Tail Index Estimation

by

Chenhua Zhang

Approved:

Major Professor: William P. McCormick

Committee: Lynne Seymour
Tharuvai N. Sriram
XiangRong Yin
Qing Zhang

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2007

Dedication

to my family

## Acknowledgments

First of all I would like to thank my major professor Dr. William P. McCormick. Without his help I can never ever finish my dissertation.

Thanks to Dr. Lynne Seymour, Dr. Tharuvai N. Sriram, Dr. Qing Zhang and Dr. Xiangrong Yin for serving as my graduate committee members.

Thanks to Dr. Ishwar V. Basawa for introducing to me the class of bifurcating autoregressive processes.

Special thanks to Dr. Philippe Barbe for his advise and endless help.

I would like to thank my parents, my wife and my son for their endless love and support.

## Table of Contents

# Chapter 1

## Introduction and Literature Review

We focuses on heavy tail analysis in this dissertation. Different authors use different definitions for heavy-tailed distributions. We choose the following definition.

**Definition.(Heavy-tailed Distribution)** *A distribution function $F$ is a (right) heavy-tailed distribution if its moment generating function doesn't exist on the positive real line, i.e.*

$$\int e^{tx}\,\mathrm{d}F(x) = \infty \quad \text{for any } t > 0.$$

Let $\overline{F} = 1 - F$ denote the tail distribution of $F$. It is easy to show that $F^{+} = \max(F, 0)$ has a moment generating function in some right neighborhood of 0 if and only if (iff) there exist positive real numbers $M$ and $t$ such that

$$\overline{F}(x) \leq Me^{-tx} \quad \text{for all } x > 0.$$

Thus we can say a distribution is heavy-tailed iff it has heavier tail than any exponential distribution. Clearly there are many distributions, such as Student's t, F, Cauchy, Pareto, log-normal, log-gamma, Weibull($\tau(<1)$,c) distributions, which are heavy-tailed. The most commonly used heavy-tailed distributions are subexponential distributions, which were introduced by Chistyakov (1964).

**Definition.(Subexponential Distribution)** *A distribution function $F$ with support $(0, \infty)$ is subexponential if for all $n \geq 2$*

$$\lim_{x\to\infty} \frac{\overline{F^{\star n}}(x)}{\overline{F}(x)} = n, \tag{1.1}$$

*where $\star$ denotes the convolution.*

The intuitive characterization of subexponentiality is for a sequence of independent random variables (r.v.s), $\{X_n\}$, with common distribution $F$

$$P\{S_n > x\} \sim P\{\max_{1\leq i\leq n} X_i > x\} \sim nP\{X_1 > x\}$$

as $x$ tends to infinity for any $n \geq 2$. This means for subexponential r.v.s their sum is large iff their maxima is large. It can be shown (Embrechts, Klüppelberg and Mikosch (1997)) that the above equality holds for all $n \geq 2$ provided it is true for $n = 2$, or, equivalently $\limsup_{x\to\infty} \frac{\overline{F^{\star 2}}(x)}{\overline{F}(x)} \leq 2$. We extend the definition of subexponential distribution to distribution with support $(-\infty, \infty)$, by saying $F$ is subexponential if $F^+ = \max(F, 0)$ is subexponential.

The most important subclass of subexponential distributions is the class of regularly varying distributions. We refer to Bingham, Goldie and Teugels (1989) for details about regularly varying function and their properties.

**Definition.(Regularly Varying)** *A real measurable function $f$ is regularly varying at infinity with index $-\alpha$ if for all $t > 0$*

$$\lim_{x\to\infty} \frac{f(tx)}{f(x)} = t^{-\alpha}. \tag{1.2}$$

A distribution $F$ has regularly varying tail if $\overline{F}$ is regularly varying at infinity with nonpositive index i.e.

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

where $L$ satisfies $\lim_{x\to\infty} \frac{L(tx)}{L(x)} = 1$ for any positive t. We call $L$ slowing varying at infinity, i.e. regularly varying at infinity with index 0.

Let $t^{-\infty} = \lim_{\alpha\to\infty} t^{-\alpha}$. We can define regularly varying at infinity with infinite index accordingly. This kind of function is called rapidly varying function.

For positive regularly varying function we have the Karamata representation, which is also called Karamata's theorem.

**Theorem 1.(Karamata's theorem)** *$f$ is a positive regularly varying function at infinity with index $-\alpha$ iff*

$$f(x) = c(x) \exp\left\{ \int_{x_0}^{x} \frac{-\alpha + \epsilon(t)}{t} \,\mathrm{d}t \right\}, \qquad x \geq x_0, \tag{1.3}$$

*where $c(x) = c + o(1)$ for some $c > 0$ and $\epsilon(t) = o(1)$.*

When $c(x) \equiv c$ we call it normalized regularly varying. Karamata's theorem yields that a absolutely continuous distribution function $F$ has a normalized regularly varying tail iff it satisfies Von Mises Condition

$$\lim_{x\to\infty} \frac{xF'(x)}{1 - F(x)} = \alpha, \tag{1.4}$$

where $\alpha$ is a positive real number.

All distributions with regularly varying tails used in applications are ultimately infinitely differentiable and their derivatives are regularly varying. Examples include the Pareto, Cauchy, Student, Burr and log-gamma distributions. Therefore it is quite natural to study smoothly varying functions, which is introduced by Barbe and McCormick (2004).

**Definition.(smoothly varying)** *A real measurable function $f$ is smoothly varying with index $-\alpha$ and order $m$ if it is ultimately $m$-times continuously differentiable and the $m$-th derivative $f^{(m)}$ is regular varying with index $-\alpha-m$. We denote the set of all such functions by $SR_{-\alpha,m}$.*

Any function smoothly varying in the sense of Bingham, Goldie and Teugels (1989, § 1.8.1) is smoothly varying of any fixed order. The class $SR_{-\alpha,m}$ may be extended to noninteger orders. This is useful to present sharp results. To define $SR_{-\alpha,\omega}$ where $\omega$ is a positive real number, we introduce the following notation. For any function $h$, let

$$\Delta^r_{t,x}(h) = \mathrm{sign}(x) \frac{h\big(t(1-x)\big) - h(t)}{|x|^r \, h(t)} \,.$$

**Definition.** *Let $\omega$ be a positive real number. Write $\omega = m + r$ where $m$ is the integer part of $\omega$ and $r$ is in $[0,1)$. A function $h$ is smoothly varying of index $-\alpha$ and order $\omega$ if it belongs to $SR_{-\alpha,m}$ and*

$$\lim_{\delta \to 0} \limsup_{t \to \infty} \sup_{0<|x|\leq \delta} \Delta^r_{t,x}(h) = 0 \, .$$

*We write $SR_{-\alpha,\omega}$ for the class of all such functions.*

We remark that the spaces $SR_{-\alpha,\omega}$ are nested because $SR_{-\alpha,r} \supset SR_{-\alpha,s}$ for $r < s$. In particular, if $\omega$ is positive with integer part $m$ and $r = \omega - m$, membership in $SR_{-\alpha,\omega}$ is guaranteed by that in $SR_{-\alpha,m+1}$, that is, by checking that the $m+1$-derivative is regularly varying of index $-\alpha - m - 1$. For further properties of smoothly varying functions of finite order, we refer to Barbe and McCormick (2004).

We present some applications of smoothly varying function in Chapters 2 and 3. In Chapter 2 we obtain higher order tail area asymptotics for the distribution of the maximum of the random walk generated by a heavy-tailed distribution with negative mean. The results can be applied to queueing processes and ruin probability calculation directly. Random walk is a fundamental stochastic model. For a sequence of independent and identically distributed (i.i.d.) r.v.s $(X_n)_{n\geq 1}$ with negative mean the associated random walk is defined by $S_0 = 0$ and for any integer $n$ positive, $S_n = X_1 + \cdots + X_n$. The distribution of its maximum, $M = \max_{n\geq 0} S_n$, can be represented as a compound-geometric distribution. By the Wiener-Hopf factorization Veraverbeke (1977) showed that, as $x$ tends to infinity,

$$\overline{W}(x) = P(M > x) \sim \frac{1}{1-p} \overline{F}_+(x) \sim \frac{-1}{\mu} \int_x^\infty \overline{F}(t) \, \mathrm{d}t \, ,$$

where $F_+$ is the distribution of first ascending ladder height and $\mu$ the first moment of $F$, which denotes the step size distribution. An asymptotic local limit result for $W$, the distribution of $M$, has been obtained in Asmussen et al. (2002). We also mention who Zachary (2004) presented a direct probabilistic proof of this result which circumvents use of the ascending ladder height distribution. In refining the estimate for compound sums, Omey and

Willekens (1986,1987) obtained second order results for $\overline{W}$ in terms of $\overline{F}_+$. Building upon the harmonic renewal theoretic method of Greenwood, Omey and Teugels (1982), Grübel (1988) obtained second order results using the so-called Banach algebra technique when the increments have a discrete distribution. Borovkov (2002) also presented a second-order result for $\overline{W}$. In order to obtain a second order result in terms of the original tail $\overline{F}$ one needs to derive an expansion for the tail of the right Wiener-Hopf factor in terms of $\overline{F}$. One of the contributions of Chapter 2 is to show that in some sense, the asymptotic regularity of $\overline{F}$ is inherited by $\overline{F}_+$, and to derive a higher order expansion for that Wiener-Hopf factor. We derive higher order results for $\overline{W}$ when $F$ has smoothly varying tail.

In Chapter 3, we construct asymptotic expansions for the tail area of a compound sum, when the distribution of the summands belongs to a class of rapidly varying subexponential distributions. To be more precise, let $X_i$, $i \geq 1$, be a sequence of independent random variables, all having the same distribution $F$. For any positive integer $n$ the distribution of the partial sums $S_n = X_1 + \cdots + X_n$ is the $n$-fold convolution $F^{\star n}$. We set $S_0 = 0$ and therefore $F^{\star 0}$ is defined as the distribution of the point mass at the origin. Let $N$ be a nonnegative integer-valued random variable, independent of the $X_i$'s. We consider the distribution $G$ of the compound sum $S_N$, that is $EF^{\star N}$, and we are seeking an asymptotic expansion for its tail area $\overline{G} = 1 - G$.

First order asymptotic results for $\overline{G}$ have been obtained by Embrechts, Goldie and Veraverbeke (1979), Cline (1987), and Embrechts (1985). A second order formula may be found in Grübel (1987) and Omey and Willekens (1987).

Compound sums or subordinated distributions arise as distributions of interest in several stochastic models. In insurance risk theory, it models the total claim amount. For a discussion of issues related to random sums and insurance risk, we refer to Embrechts, Klüppelberg and Mikosch (1997), Asmussen (1997), and Goldie and Klüppelberg (1998). Compound sums also appear in queueing theory, in connection with the stationary distribution of waiting times in the GI/G/1 queue. The connection here is not as direct as in the insurance risk model in that it is derived from an analysis of ladder heights for transient random walks; see, for example, Asmussen (1987, p.80), Feller (1971, p.396) and Pakes (1975). Another common way in which this model occurs is through the solution of a transient renewal equation. An example of this occurs in branching processes, where we obtain a geometric-compound sum in the analysis of the mean number of particles alive at a given time in an age-dependent subcritical process; see Athreya and Ney (1972, p.151). We refer to Feller (1971, chapter XI) for a discussion of transient renewal theory. For further applications of subexponentiality in transient renewal theory, we refer to Teugels (1975) and Embrechts and Goldie (1982).

Tail index $-\alpha$ is a key parameter for distributions with regularly varying tail. A well studied estimator of the reciprocal of the tail index, Hill's estimator (Hill (1975), is defined by

$$\frac{1}{k}\sum_{i=1}^{k}\log\frac{X_{(i)}}{X_{(k+1)}},$$

where $X_{(i)}$ denotes the $i$th largest order statistics, e.g. $X_{(1)}$ denotes the maximum sample value. de Haan and Resnick (1998) established asymptotically normality for iid data. Hsing (1991), Rootzen et al. (1990) studied the asymptotic properties of Hill's estimator under weak dependence assumptions. Resnick and Stărică (1995,1998) proved the weak consistency of Hill's estimator for certain classes of stationary sequence. Resnick and Stărică (1997), Ling and Peng (2004) studied the asymptotic normality of Hill's estimator for AR and ARMA model.

In Chapter 4 we study shot noise sequences

$$X_j = \sum_{i\le j} A_i h(\tau_j - \tau_i), \tag{1.5}$$

where $\tau_j$'s are renewal points, and $\{A_i\}$ is a sequence of iid positive r.v.s, independent of the renewal process, with heavy-tailed distribution. $h$ is real valued function on $[0,\infty)$ and is called impulse response function. Under a weak integrability condition on $h$ we established asymptotic normality of Hill's estimator for the tail distribution of the $A_i$. When $\tau_i$'s are degenerate r.v.s the shot noise sequence is an infinite order moving average process. Our condition on the coefficients is weaker than that in Resnick and Stărică (1997).

In Chapter 5 we consider first-order bifurcating autoregressive processes

$$X_t = \phi X_{\lfloor t/2 \rfloor} + \epsilon_t, \tag{1.6}$$

where $0 \leq \phi < 1$, $\lfloor x \rfloor$ is the largest integer which does not exceed $x$. Cowan (1984) and Cowan and Staudte (1986) proposed the bifurcating autoregressive (BAR) model to analyze cell lineage data. Huggins (1995) derived asymptotic properties of both robust and maximum likelihood estimators. Zhou and Basawa(2005) studied asymptotic properties of parameter estitmator $\hat{\phi}$ for BAR(1) processes with bivariate exponential errors. We consider first-order bifurcating autoregressive processes with Weibull type errors. The coefficient $\phi$ and tail index $\alpha$ are two parameters we are interested in. Using point process methods, we obtain the joint limit distribution of $(\hat{\phi}, \hat{\alpha})$. They are asymptotically independent.

# Chapter 2

# Tail Expansions for the Distribution of the Maxima of Random Walks with Negative Drift and Regularly Varying Increments[1]

[1]Barbe, Ph., McCormick, W.P., Zhang, C. Submitted to *Stochastic Processes and their Applications,* 4/2006.

**Abstract**

Let $F$ be a distribution function with negative mean and regularly varying right tail. Under a mild smoothness condition we derive higher order asymptotic expansions for the tail distribution of the maxima of the random walk generated by $F$. The expansion is based on an expansion for the right Wiener-Hopf factor which we derive first. An application to ruin probabilities is developed.


**Keywords:** tail expansion, random walk, regularly varying, Wiener-Hopf factor, ruin probability

## 2.1 Introduction

There is hardly a more basic stochastic model than a random walk, and for random walks with negative drift, a basic issue of study is the distribution of its global maximum. One reason for interest in this quantity is its connection to queueing processes (Asmussen, 1987, § III.7) and to ruin computations in insurance (Embrechts, Klüppelberg and Mikosch 1997, § 1.1).

In this chapter, we are interested in proving higher order tail area asymptotics for the distribution of the maximum of the random walk when the increments have a heavy-tailed distribution with negative mean. To explain the contribution of this chapter and the technique it uses, we first need to succinctly explain how first and second order results have been obtained. To this aim, let $(X_n)_{n\geq 1}$ be a sequence of independent and identically distributed random variables having negative mean. The associated random walk is defined by $S_0 = 0$ and for any integer $n$ positive, $S_n = X_1 + \cdots + X_n$. The distribution of its maximum, $M = \max_{n\geq 0} S_n$, can be represented as a compound-geometric distribution as follows. We first agree that the minimum of the empty set is $+\infty$. Then, let $\tau$ denote the hitting time for the positive half-line

$$\tau = \min\{\, n : S_1 \leq 0, \ldots, S_{n-1} \leq 0\,,\ S_n > 0\,\}\,.$$

This hitting time may be infinite, but it is finite with probability

$$p = P\{\,\tau < \infty\,\} = 1 - P\{\,S_1 \le 0, S_2 \le 0, \dots\,\}\,.$$

Recall that the first strict ascending ladder height distribution is defined by

$$F_+(x) = P\{\,S_\tau \le x\,,\,\tau < \infty\,\}\,.$$

Since the random walk has a negative drift, $F_+$ is a defective distribution with defect $1-p$. It follows that $M$ has a compound-geometric distribution subordinate to the distribution $H = p^{-1}F_+$ and with subordinator a geometric distribution with parameter $p$. More explicitly and following Feller (1971, § XII.5), writing $H^{\star n}$ for the $n$-fold convolution of $H$, the distribution $W$ of $M$ is

$$W = (1-p)\sum_{n\ge 0} p^n H^{\star n}\,. \tag{2.1}$$

This step has replaced the original question of analyzing the distribution of global maximum of a random walk with the more elementary question of analyzing that of a compound sum, at the price, however, of introducing a derived distribution, namely, the ascending ladder height distribution, which requires its own analysis. In the case of a heavy-tailed step size distribution as prescribed by a subexponentiality assumption, Veraverbeke (1977) supplies an answer to this question through use of the distributional form of the Wiener-Hopf factorization. His result establishes inheritability of the subexponential property of the right Wiener-Hopf factor, i.e. that factor having mass concentrated in positive half line, from that of the underlying distribution. To state that result, we agree that for any possibly defective distribution function $G$, we write $\overline{G}$ for its tail, that is, the function whose value at $x$ is

$$\overline{G}(x) = \lim_{t\to\infty} G(t) - G(x)\,.$$

Consider the Wiener-Hopf factorization $F = F_+ + F_- - F_+ \star F_-$, where $F_-$ and $F_+$ are concentrated on $(-\infty, 0\,]$ and $(0,\infty)$ respectively (see Feller, 1971, § XII.3). Let $\mu$ be the mean of $F$, which we assume to be negative. Veraverbeke (1977) showed that, as $x$ tends to infinity,

$$\overline{F}_+(x) \sim \frac{1-p}{-\mu} \int_x^\infty \overline{F}(t)\,\mathrm{d}t\,.$$

We remark that the Wiener-Hopf factors are given by the strict ascending ladder height distribution and the weak descending ladder height distribution.

With these two steps in place, a first-order analysis of the distribution of $M$ may be completed by using a result on tail-area asymptotics for subordinated probability distributions — in this case for the subordinator given by a geometric distribution. For example, the result in Athreya and Ney (1972, § IV.4) for first-order asymptotics of compound subexponential distributions with geometric subordinator gives the expected result that, under the assumption of $F$ subexponential with negative mean $\mu$,

$$\overline{W}(x) \sim \frac{1}{1-p}\overline{F}_+(x) \sim \frac{-1}{\mu} \int_x^\infty \overline{F}(t)\,\mathrm{d}t\,,$$

as $x$ tends to infinity.

In refining the estimate for compound sums, Omey and Willekens (1987) obtained second order results for $\overline{W}$ in terms of $\overline{F}_+$. In order to obtain a second order result in terms of the original tail $\overline{F}$ one needs to derive an expansion for the tail of the right Wiener-Hopf factor in terms of $\overline{F}$. One of the contributions of this chapter is to show that in some sense, the asymptotic regularity of $\overline{F}$ is inherited by $\overline{F}_+$, and to derive a higher order expansion for that Wiener-Hopf factor.

Before discussing further the technique used in this chapter, we mention that when $\overline{F}$ is regularly varying with index in the range from $-1$ to $-2$ and the mean is finite and negative, Omey and Willekens (1986) establish a second-order result for the tail $\overline{W}$; see also Geluk (1992, 1996).

Our approach is based on the algebraic formalism developed in Barbe and McCormick (2004). Beyond the specific expansions, a contribution of this chapter is to extend this formalism, suggesting a simple way of inverting some operators which are not invertible in the framework of Barbe and McCormick (2004).

Building upon the harmonic renewal theoretic method of Greenwood, Omey and Teugels (1982), Grübel (1988) obtained second order results using the so-called Banach algebra technique when the increments have a discrete distribution. As pointed out in Grübel (1987), the Banach algebra technique has the potential to lead to higher order expansions. While this technique has been used very successfully by Grübel (1987) to derive an expansion for generalized renewal measures with non-summable weights (a case which our technique cannot handle so far), it has not yet produced as high-order expansions as our technique when the weights decay fast enough. The Banach algebra technique has not been used successfully to obtain higher-order expansion for tail of weighted convolutions, which was the main motivation of Barbe and McCormick (2004). In our view, the domain of application of the two techniques clearly have a nonempty intersection, but the two approaches appear to complement each other quite nicely, with one becoming useful when the other one appears to run into some difficulties (a good illustration is in implicit renewal theory, by comparing the assumptions of Goldie (1991), with that of Barbe and McCormick, 2004). On a different note, our algebraic formalism appears to lead to compact expressions, suitable to use with a computer algebra package, while the Banach algebra method did not yield such a clean formalism. However, it should be pointed that regardless of the method used, an asymptotic expansion is unique once the asymptotic scale is chosen, and therefore, in that respect, which method to use is sometimes a pure matter of taste.

To describe our approach further and to show its conceptual simplicity, we need to introduce the algebraic formalism to express the results. To that end, let $\mathrm{D}$ be the derivation operator, mapping a differentiable function to its derivative. We define $\mathrm{D}^0$ to be the identity operator, and by induction, we set $\mathrm{D}^k = \mathrm{D}\mathrm{D}^{k-1}$ for any positive integer $k$. We also write $\mathrm{Id}$ for the identity operator, mapping a function $f$ to itself; that is, if $f$ is a function, $\mathrm{Id}f$ is $f$. The symbol $\mathrm{Id}$ also denotes the identity function on the real line, in which case, $\mathrm{Id}f$ is the function mapping $x$ to $xf(x)$. Which use is intended will be clear from the context.

Let $\mathbb{R}_m[\,\mathrm{D}\,]$ denote the ring of real polynomials in D modulo the ideal generated by $\mathrm{D}^{m+1}$. In other words, any polynomial in D divisible by $\mathrm{D}^{m+1}$ is set equal to 0.

For a possibly defective distribution function $G$ with at least $k$ moments finite, we write $\mu_{G,k}$ its $k-$th moment. Note in particular that $\mu_{G,0}$ is the total mass of $G$, equal to 1 if and only if $G$ is not defective.

**Definition.** *The Laplace character of order $m$ of a possibly defective distribution $G$ having a finite $m$-th moment is the element of $\mathbb{R}_m[\,\mathrm{D}\,]$ given by*

$$L_{G,m} = \sum_{0\leq k\leq m} \frac{(-1)^k}{k!} \mu_{G,k} \mathrm{D}^k \,.$$

It is easy to see that the Laplace characters realize an operator-valued representation of a convolution semi-group, in the sense that $L_{F\star G,m} = L_{F,m} L_{G,m}$, where the multiplication in the right hand side is in $\mathbb{R}_m[\,\mathrm{D}\,]$. The usefulness of Laplace characters is suggested in the following heuristic. Consider the Wiener-Hopf factorization, written on tails and on the nonnegative half-line as

$$\overline{F}(t) = \overline{F}_+(t) - \int_{-\infty}^0 \overline{F}_+(t-x)\,\mathrm{d}F_-(x)\,.$$

In the integral term, suppose that we can apply Taylor's formula to obtain the approximation

$$\begin{aligned} \overline{F}(t) &\approx \overline{F}_+(t) - \sum_{0\leq k\leq m} \frac{(-1)^k}{k!} \int_{-\infty}^0 x^k \,\mathrm{d}F_-(x) \overline{F}_+^{(k)}(t) \\ &= (\mathrm{Id} - L_{F_-,m}) \overline{F}_+(t)\,; \end{aligned} \tag{2.2}$$

it is then tempting to invert the operator $\mathrm{Id} - L_{F_-,m}$ and write

$$\overline{F}_+ \approx (\mathrm{Id} - L_{F_-,m})^{-1} \overline{F}\,,$$

obtaining an expansion of the Wiener-Hopf factor by a simple formal series expansion of $(\mathrm{Id} - L_{F_-,m})^{-1}$ as function of D. This is not possible because the constant term of $\mathrm{Id} - L_{F_-,m}$ vanishes, or in other words, because this operator is in the ideal generated by D. A contribution of this chapter mentioned earlier is to note that one can factor D in $\mathrm{Id} - L_{F_-,m}$, define a suitable inverse of D (a simple integration), and justify rigorously our heuristic. Then, we would like to push forward this expansion into one on $\overline{W}$ by using the higher order expansion for compound sums derived in Barbe and McCormick (2004). For those to be applicable, one needs to show that higher derivatives of the right Wiener-Hopf factor exist and are regularly varying. One cannot hope this to be true without regularity assumptions on $\overline{F}$, and therefore, another contribution of this chapter is to show that the order of differentiability of the tail $\overline{F}$ is ultimately inherited by $\overline{F}_+$ and that regular variation of the derivatives of the tail $\overline{F}$ translates into regular variation of the derivatives of the Wiener-Hopf factor.

## 2.2 Expansions

Since the Laplace characters are differential operators, formula (2.2) suggests that some smoothness requirement is needed. Moreover, that same formula suggests that one would like to have the derivatives of $\overline{F}_+$ to be asymptotically of smaller and smaller order. The following rather natural smoothness condition will be needed for our results.

**Definition.** *A real measurable function $f$ is smoothly varying with index $-\alpha$ and order $m$ if it is ultimately $m$-times continuously differentiable and the $m$-th derivative $f^{(m)}$ is regular varying with index $-\alpha - m$. We denote the set of all such functions by $SR_{-\alpha,m}$.*

All distributions with regularly varying tails used in applications are smoothly varying of arbitrary order. Examples include the Pareto, Cauchy, Student, Burr and log-gamma distributions. Any function smoothly varying in the sense of Bingham, Goldie and Teugels (1984, § 1.8.1) is smoothly varying of any fixed order.

The class $SR_{-\alpha,m}$ may be extended to noninteger orders. This is useful to present sharp results. To define $SR_{-\alpha,\omega}$ where $\omega$ is a positive real number, we introduce the following notation. For any function $h$, let

$$\Delta^r_{t,x}(h) = \mathrm{sign}(x) \frac{h\bigl(t(1-x)\bigr) - h(t)}{|x|^r \, h(t)} \, .$$

**Definition.** *Let $\omega$ be a positive real number. Write $\omega = m + r$ where $m$ is the integer part of $\omega$ and $r$ is in $[0,1)$. A function $h$ is smoothly varying of index $-\alpha$ and order $\omega$ if it belongs to $SR_{-\alpha,m}$ and*

$$\lim_{\delta\to 0}\limsup_{t\to\infty}\sup_{0<|x|\le\delta}\Delta^r_{t,x}(h) = 0\,.$$

*We write $SR_{-\alpha,\omega}$ for the class of all such functions.*

We remark that the spaces $SR_{-\alpha,\omega}$ are nested, for $SR_{-\alpha,r} \supset SR_{-\alpha,s}$ for $r < s$. In particular, if $\omega$ is positive with integer part $m$ and $r = \omega - m$, membership in $SR_{-\alpha,\omega}$ is guaranteed by that in $SR_{-\alpha,m+1}$, that is, by checking that the $m+1$-derivative is regularly varying of index $-\alpha - m - 1$. For further properties of smoothly varying functions of finite order, we refer to Barbe and McCormick (2004).

The Laplace characters have been introduced in the introductory section. As mentioned in the last paragraph of that section, we would like to factor D in $\mathrm{Id} - \mathsf{L}_{F_-,m}$. To do this properly, we define the backward signed shift $\mathsf{S}$ on polynomials in D by extending linearly the equalities $\mathsf{S}\mathrm{D}^0 = 0$ and for any $j$ positive integer, $\mathsf{S}\mathrm{D}^j = -\mathrm{D}^{j-1}$. It maps $\mathbb{R}_m[\,\mathrm{D}\,]$ to $\mathbb{R}_{m-1}[\,\mathrm{D}\,]$. In particular, one sees that for any proper distribution $G$,

$$\mathrm{Id} - \mathsf{L}_{G,m} = \mathrm{D}(\mathsf{S}\mathsf{L}_{G,m}) = (\mathsf{S}\mathsf{L}_{G,m})\mathrm{D}$$

and that $\mathsf{S}\mathsf{L}_{G,m}$ is invertible in $\mathbb{R}_m[\,\mathrm{D}\,]$ if and only if $\mu_{F,1}$ does not vanish.

We define the inverse of the differentiation on some functions as follows. If $f$ is a function regularly varying of index less than $-1$, we set

$$\mathrm{D}^{-1}f(t) = -\int_t^\infty f(x)\,\mathrm{d}x\,.$$

Clearly, $\mathrm{D}\mathrm{D}^{-1}$ is the identity on functions which are regularly varying of index less than $-1$, while $\mathrm{D}^{-1}\mathrm{D}$ is the identity on the smoothly varying functions of negative index and order at least 1.

We now present our results. Recall $F$ denotes the step size distribution about which we assume its first moment is negative and that its right tail is regularly varying of index $-\alpha$.

The strict ascending ladder height distribution is $F_+$ and $W$ is the distribution of $M$. Let

$$\kappa = \sup\Bigl\{\, r \geq 0 \, : \, \int_{-\infty}^0 |x|^r \, \mathrm{d}F(x) < \infty \,\Bigr\}\,.$$

This may be less than $\alpha$ if the lower tail of $F$ is heavier than the upper one. In the following theorems, it is supposed that $\kappa$ is greater than 1.

Our first theorem gives the regularity of the positive Wiener-Hopf factor and its asymptotic expansion in terms of $\overline{F}$ and its derivatives.

**Theorem 2.1.** *Suppose that $\overline{F}$ is smoothly varying of index $-\alpha$ and order $\omega$. Then*
*(i) $\overline{F}_+$ is smoothly varying of index $-\alpha+1$ and same order $\omega$ as $\overline{F}$;*
*(ii) for any integer $m$ at least 1 and less than $\omega\wedge\alpha\wedge\kappa$, the moments $\mu_{F_-,m}$ and $\mu_{F_+,m-1}$ are finite and*

$$\overline{F}_+ = (\mathsf{SL}_{F_-,m})^{-1}\mathrm{D}^{-1}\overline{F} + o(\mathrm{Id}^{-m+2}\overline{F})\,.$$

Our next result is then an expansion for the tail of the maximum of the random walk.

**Theorem 2.2.** *Suppose that $\overline{F}$ is smoothly varying of index $-\alpha$ and order $\omega$. Then, for any integer $m$ at least 1 and less than $\omega\wedge\alpha\wedge\kappa$*

$$\overline{W} = (1-p)(\mathrm{Id} - \mathsf{L}_{F_+,m-1})^{-2}(\mathsf{SL}_{F_-,m})^{-1}(\mathrm{D}^{-1}\overline{F}) + o(\mathrm{Id}^{-m+2}\overline{F})\,.$$

**Remark.** As previously mentioned, Laplace characters of order $m-1$ are elements of the ring $\mathbb{R}_{m-1}[\,\mathrm{D}\,]$. The inverses $(\mathrm{Id} - L_{F_+,m-1})^{-2}$ and $(\mathsf{SL}_{F_-,m})^{-1}$ are taken in that ring, multiplied together in that ring, and applied to $\mathrm{D}^{-1}\overline{F}$.

**Remark.** The result may seem a little mysterious and not so explicit at a first glance. However, the computations related to Laplace characters are usually easy to carry out because they amount to expanding in a Taylor series in D whatever formal expansion is being considered. In particular, if needed, those computations can be implemented with a computer algebra package. For instance, the following very short `Maple` code calculates the expansion given in the Theorem. In that code, `Fp` and `Fm` stand for $F_+$ and $F_-$.

```
restart; m:=4: mu[Fp,0]:=1-q:
LFp:=sum('(-1)^j*mu[Fp,j]*x^j/j!','j'=0..m-1):
SLFm:=sum('(-1)^j*mu[Fm,j+1]*x^j/(j+1)!','j'=0..m-1):
a:=taylor((1-LFp)^(-2),x=0,m-1):
b:=taylor(SLFm^(-1),x=0,m-1):
expand(convert(q*taylor(a*b,x=0,m-1),polynom)/x);
```

One simply replaces `x` and `q` in the output by D and $1-p$, with the convention that `1/x` should be replaced by $\mathrm{D}^{-1}$. For instance, using that $\mu_{F,1} = (1-p)\mu_{F_-,1}$, taking $m$ to be 4, we deduce the 3-terms expansion

$$\begin{aligned}\overline{W} \;=\; & \frac{1}{\mu_{F,1}}\mathrm{D}^{-1}\overline{F} + \frac{1}{2\mu_{F,1}^2}\big((1-p)\mu_{F_-,2} - 4\mu_{F_+,1}\mu_{F_-,1}\big)\overline{F} \\ & + \frac{1}{12\mu_{F,1}^3}\Big(3(1-p)^2\mu_{F_-,2}^2 + 12\mu_{F,1}(\mu_{F_-,1}\mu_{F_+,2} - \mu_{F_+,1}\mu_{F_-,2}) \\ & \qquad\qquad + 36\mu_{F_-,1}^2\mu_{F_+,1}^2 - 2(1-p)\mu_{F,1}\mu_{F_-,3}\Big)\overline{F}\,' \\ & + o(\mathrm{Id}^{-1}\overline{F})\,.\end{aligned}$$

**Remark.** An important point to mention with regard to the main result is that the expansion it provides for the tail distribution $\overline{W}$ is based on the underlying distribution $F$ of the random walk, its derivatives and its integrated tail. This is notable, because the starting point to obtain this result is that of a tail area expansion for a subordinated distribution based on underlying distribution given by $F_+$. Since $F_+$ is generally unattainable in an explicit form, the formulation of our main result is more attractive than that which results from a direct application of a tail area result for subordinated distributions. It is a comment on the usefulness of this algebraic approach that such an improvement is so easily and transparently attained compared to the effort to accomplish the same goal analytically. We conclude this remark by noting that our proof shows that a penultimate expansion based on $F_+$ is given by

$$\overline{W} = (1-p)(\mathrm{Id} - \mathsf{L}_{F_+,m})^{-2}\overline{F}_+ + o(\mathrm{Id}^{-m+1}\overline{F}) \tag{2.3}$$

provided $m$ is less than $\omega \wedge (\alpha - 1)$. When $m$ is 1 in the above, we obtain a second-order result in agreement with Theorem 2.2 in Omey and Willekens (1987). Comparing this formula with that given in our theorem, we see that the latter is slightly less accurate. The reason is that, under the assumption of the theorem, replacement of $\overline{F}_+$ with an approximation based on $\overline{F}$, its derivatives and its integral comes with a one-order lower error bound, viz.

$$\overline{F}_+ = (\mathsf{S}\mathsf{L}_{F_-,m})^{-1}\mathrm{D}^{-1}\overline{F} + o(\mathrm{Id}^{-m+2}\overline{F})\,.$$

But note that in (2.3), $m$ is required to be below $\alpha - 1$ whereas in Theorem 2.2, $m$ is required to be less than $\alpha$. Therefore, provided that the descending ladder height distribution has adequately high enough moments, the two expansions of $\overline{W}$ in (2.3) and Theorem 2.2 when taken to their fullest length provide the same order of magnitude for the error bound.

## 2.3 Application

Finally, we present an application to insurance risk. To that end, we introduce some notation. Let $R_0 = x$ be the initial capital of an insurance company. We assume that the claim amounts, $(A_n)_{n\geq 1}$, are independent, with common distribution function $L$ having a smoothly varying tail of index $-\alpha$ and order $m+1$. We also assume that the interclaim times $(T_n)_{n\geq 1}$ are independent, with common distribution $K$, and independent of the claim amounts. Finally, we assume the intensity of the gross risk premium is some positive $c$. The net loss to the company in period $n$ is $X_n = A_n - c\,T_n$. The sequence $(X_n)_{n\geq 1}$ is a sequence of independent random variables with distribution $F(x) = \int_0^\infty L(x + c\,t)\,\mathrm{d}K(t)$. Under the assumptions on $L$, it follows that $F$ is ultimately $m$-times differentiable and

$$\frac{F^{(m)}(x)}{L^{(m)}(x)} = \int_0^\infty \frac{L^{(m)}(x+ct)}{L^{(m)}(x)}\,\mathrm{d}K(t)\,.$$

This implies that $\overline{F}$ is smoothly varying of index $-\alpha$ and order $m$. Let $\psi(R_0)$ be the probability of eventual ruin given $R_0$. Writing as before $S_n$ for the random walk with increment $X_i$ and $M$ for its maximum, $\psi(x) = P\{\, M > x \,\}$. We follow our established notation and set $F_+$ and $F_-$ for the strict ascending and weak descending ladder height distributions for the random walk $S_n$. We assume that $X_1$ has negative expectation. Then, we have the following expansion of the ruin probability, which obviously follows from the theorem.

**Corollary.** *Assume that $\overline{L}$ is smoothly varying of index $-\alpha$ and order $m+1$. Assume also that $\mu = EA_1 - cET_1$ is finite and negative and that $m$ less than $\alpha$. Then,*

$$\psi = (1-p)(\mathrm{Id} - \mathsf{L}_{F_+,m-1})^{-2}(\mathsf{SL}_{F_-,m})^{-1}\mathrm{D}^{-1}\overline{F} + o\big(\mathrm{Id}^{-m+2}\overline{F}\big)\,.$$

## 2.4 Proof of the Theorems

We first show that $\mu_{F_-,m}$ is indeed finite. The distributional form of the Wiener-Hopf factorization implies that on the negative half-line,

$$F_- = F + F_+ \star F_-\,.$$

Since $F_+$ has defect $1-p$, this yields $F_- \leq F + pF_-$, from which we deduce $F_- \leq (1-p)^{-1}F$. This proves the finiteness of $\mu_{F_-,m}$. That of $\mu_{F_+,m-1}$ follows and this proves part of statement (ii) of Theorem 2.1.

We begin the proof of the main part of our theorem by establishing a preparatory lemma.

**Lemma 1.** *Let $(Y_i)_{i\geq 1}$ be a sequence of nonnegative random variables, independent and identically distributed with finite and positive mean. Let $(Z_n)_{n\geq 0}$ be their corresponding random walk. Furthermore, let $f$ be a regularly varying function of index $-\beta$ less than $-1$. Then,*

$$\lim_{x\to\infty} \frac{1}{xf(x)} \sum_{n\geq 0} Ef(x+Z_n) = \frac{1}{(\beta-1)EY_1}\,.$$

**Proof.** By Potter's theorem (Bingham, Goldie and Teugels, § 1.5.4) it is sufficient to prove it with $f(x) = x^{-\gamma}$ for any positive $\gamma$. Let $\theta$ be positive and less than the mean of the $Y_i$'s. We have the trivial bound

$$\sum_{n\geq 0} Ef(x+Z_n) \leq \sum_{n\geq 0} f(x+\theta n) + f(x)\sum_{n\geq 0} P\{\, Z_n \leq \theta n\,\}\,.$$

Since $f$ is a decreasing funciton,

$$\sum_{n\geq 0} f(x+\theta n) = \sum_{n\geq 0}(x+\theta n)^{-\gamma} \sim \int_0^\infty (x+\theta s)^{-\gamma}\,\mathrm{d}s = \frac{xf(x)}{\theta(\gamma-1)}\,.$$

as $x$ tends to infinity. Let $M$ be such that $E(Y_1\wedge M)$ is greater than $\theta$. Such $M$ exists by monotone convergence of $Y_1 \wedge M$ to $Y_1$. Let $(Z_n^M)_{n\geq 0}$ be the random walk associated to the sequence $(Y_i\wedge M)_{i\geq 1}$. By the Hsu and Robbins theorem (see Chow and Teicher, 1988, § 10.4), the series $\sum_{n\geq 0} P\{\, Z_n^M \leq n\theta\,\}$ is finite. Since $Z_n^M$ is at most $Z_n$, this series is at least $\sum_{n\geq 0} P\{\, Z_n \leq n\theta\,\}$, and the latter is finite as well. Therefore, since $\theta$ is any positive number less than $EY_1$,

$$\limsup_{x\to\infty} \frac{1}{xf(x)} \sum_{n\geq 0} Ef(x+Z_n) \leq \frac{1}{(\gamma-1)EY_1}\,.$$

To obtain a matching lower bound, note that $f$ is a convex function. By Hölder's inequality

$$\sum_{n\geq 0} Ef(x+Z_n) \geq \sum_{n\geq 0} f(x+nEY_1) \sim \frac{xf(x)}{(\gamma-1)EY_1}$$

as $x$ tends to infinity. ■

Note that with unimportant and additional assumptions an alternate proof of Lemma 1 based on the renewal theorem may be given, slightly shorter, but not as direct. To sketch it, write $G$ the distribution function of $Y_i$ and consider the renewal function $U = \sum_{n\geq 0} G^{\star n}$. We see that $\sum_{n\geq 0} Ef(x+Z_n) = \int f\,\mathrm{d}U$. When $f$ is smooth, an integration by parts and a change of variable bring this integral to the form $f(x)U(0) - s\int_1^\infty f'(xs)U\big(x(s-1)\big)\,\mathrm{d}s$. The renewal theorem (Feller, 1971, § XI.3) yields $U\big(x(s-1)\big) \sim x(s-1)/EY_1$ as $x$ tends to infinity, uniformly in $s$ at least 1. The result then follows by standard arguments involving regular variation. This shows that Lemma 1 is implied by the renewal theorem. Theorem 1.7.4 in Bingham, Goldie and Teugels (1989) showed that conversely, the renewal theorem is implied by Lemma 1.

The next lemma restate part (i) of Theorem 2.1 and restate a known first order equivalent for the right Wiener-Hopf factor.

**Lemma 2.** *The strict ascending ladder height distribution $F_+$ is smoothly varying of index $-\alpha+1$ and same order $\omega$ as $F$. Moreover,*

$$\overline{F}_+ \sim \frac{-1}{(\alpha-1)\mu_{F_-,1}} \,\mathrm{Id}\,\overline{F}\,. \tag{2.4}$$

**Proof.** The proof has four steps.

*Step 1. A representation for $\overline{F}_+$.* By the distributional form of Wiener-Hopf factorization, we have

$$F = F_+ + F_- - F_+ \star F_-\,. \tag{2.5}$$

It is convenient to introduce the following integral operator,

$$U_{F_-} g(t) = \int_{-\infty}^{0} g(t-u)\,\mathrm{d}F_-(u)\,.$$

As usual, powers of operators are defined inductively. In particular, $U^0_{F_-}$ is the identity and $U^n_{F_-} = U_{F_-} \circ U^{n-1}_{F_-}$ for any integer $n$ positive. On $(0,\infty)$, we can write (2.5) as $\overline{F}_+ = \overline{F} + \overline{F_+ \star F_-}$, which leads to

$$\overline{F}_+ = \overline{F} + U_{F_-}\overline{F}_+\,. \tag{2.6}$$

By recursion this yields

$$\overline{F}_+ = \sum_{0\leq i\leq n} U^i_{F_-}\overline{F} + U^{n+1}_{F_-}\overline{F}_+\,.$$

Note that $F_-$ cannot be the distribution degenerate at 0 since $F$ is assumed to have a negative mean. Let $(Y_i)_{i\geq 1}$ be a sequence of independent random variables, all with the same distribution $F_-$, and let $(Z_n)_{n\geq 0}$ be their random walk (note that the signs are changed compared to the previous lemma). Observe that $\sum_{0\leq i\leq n} U^i_{F_-}\overline{F}$ is nondecreasing in $n$ and that, by dominated convergence, $U^{n+1}_{F_-}\overline{F}_+(x) = E\overline{F}_+(x - Z_{n+1})$ tends to 0 as $n$ goes to infinity. Consequently, we obtain the representation

$$\overline{F}_+ = \sum_{i\geq 0} U^i_{F_-}\overline{F}\,.$$

Note that combined with Lemma 1, this representation yields Veraverbeke's (1977) theorem asserting that (2.4) holds.

*Step 2. A representation for* $\overline{F}^{(k)}_+$. Let $k$ be a positive integer at most $\omega\wedge(\alpha-1)$. Using the mean value theorem, there exists a sequence of real numbers, $(\theta_n)_{n\geq 0}$, nonnegative and at most 1, such that

$$\begin{aligned}
&\sum_{n\geq 0}\Big|\frac{1}{\epsilon}\Big(U^n_{F_-}\overline{F}^{(k-1)}(x+\epsilon) - U^n_{F_-}\overline{F}^{(k-1)}(x)\Big) - U^n_{F_-}\overline{F}^{(k)}(x)\Big|\\
=\;& \sum_{n\geq 0}\big|U^n_{F_-}\overline{F}^{(k)}(x+\theta_n\epsilon) - U^n_{F_-}\overline{F}^{(k)}(x)\big|\\
\leq\;& \sum_{n\geq 0} E\big|\overline{F}^{(k)}(x+\theta_n\epsilon - Z_n) - \overline{F}^{(k)}(x-Z_n)\big|\,.
\end{aligned}$$

Since the absolute value of a difference is at most the sum of the absolute values, Lemma 1 shows that the above series is bounded as a function of $x$ and uniformly in $\epsilon$ in some interval $(0,\eta)$. Moreover, every summand tends to 0 as $\epsilon$ tends to 0. Therefore, the series tends to 0 as $\epsilon$ tends to infinity. This proves that

$$\overline{F}^{(k)}_+ = \sum_{i\geq 0} U^i_{F_-}\overline{F}^{(k)}. \tag{2.7}$$

By recursion this yields on some neighborhood of infinity.

*Step 3.* $\overline{F}^{(k)}_+$ *is regularly varying.* The asymptotic equivalence in (3.0) implies that $\overline{F}_+$ is regularly varying with index $-\alpha+1$. Recall that $k$ is at most $\omega\wedge(\alpha-1)$. By assumption $\overline{F}^{(k)}$ is regularly varying. By representation (2.7) and Lemma 1, $\overline{F}^{(k)}_+$ is regularly varying of index $-\alpha-k+1$. Taking $k$ to be $m$, that is, $\lfloor\omega\rfloor$, this proves that $\overline{F}_+$ is smoothly varying of index $-\alpha+1$ and order $m$.

*Step 4. Concluding the proof of the lemma.* Following Barbe and McCormick (2004), for a function $h$ define

$$\overline{\Delta}^r_{\tau,\delta}(h) = \sup_{t\geq\tau} \sup_{0<|x|\leq\delta} |\Delta^r_{t,x} h| \, .$$

This quantity is nonincreasing in $\tau$ and nondecreasing in $\delta$. Using representation (2.7), we see that

$$\overline{F}^{(m)}_+\big(t(1-x)\big) - \overline{F}^{(m)}_+(t) = \sum_{n\geq 0} E\Big(F^{(m)}\big(t(1-x)-Z_n\big) - F^{(m)}(t-Z_n)\Big) \, .$$

Consider $x$ in the range $[\,-\delta,\delta\,]\setminus\{\,0\,\}$. Factoring $t-Z_n$ in $t(1-x)-Z_n$, the $n$-th summand in the series above is at most

$$\Big(\frac{t}{t-Z_n}|x|\Big)^r |\overline{F}^{(m)}(t-Z_n)| \; \overline{\Delta}^r_{t-Z_n, t\delta/(t-Z_n)} \overline{F}^{(m)} \, .$$

Consequently, for $|x|$ positive and at most $\delta$,

$$|\Delta^r_{t,x}\overline{F}^{(m)}_+| \leq E\sum_{n\geq 0}\Big|\frac{\overline{F}^{(m)}(t-Z_n)}{\overline{F}^{(m)}_+(t)}\Big| \; \overline{\Delta}^r_{t,\delta}\overline{F}^{(m)} \, .$$

It follows from step 3, $\overline{F}^{(m)}_+ \asymp \mathrm{Id}\overline{F}^{(m)}$. It then follows from Lemma 1 and our assumption on $F$ that

$$\lim_{\delta\to 0}\lim_{t\to\infty}\sup_{0<|x|<\delta} |\Delta^r_{t,x}\overline{F}^{(m)}_+| = 0 \, ,$$

proving the smooth variation of order $\omega$ of $\overline{F}_+$. ■

Finally, we present a technical lemma of some independent interest, particularly in the light of Marić's (2000) work. It is needed for the proof of our main result. We remark that the result is not proved under optimal conditions.

**Lemma 3.** *Let $(a_i)_{0\leq i\leq m}$ be a sequence of real numbers with $a_0$ different from 0. For any nonnegative integer $k$ at most $m$, define the differential operators $P_k(\mathrm{D}) = \sum_{0\leq i\leq k} a_i \mathrm{D}^i$. Let $\psi$ be a function. Let $f$ and $g$ be two functions smoothly varying with index $-\alpha$ and order at least $m$ satisfying the differential equations*

$$P_{m-k}(\mathrm{D})\mathrm{D}^k f = \mathrm{D}^k g + o(\psi) \, , \qquad k = 0,1,\ldots,m.$$

*Then, viewing $P_m(\mathrm{D})$ in $\mathbb{R}_m[\,\mathrm{D}\,]$,*

$$f = P_m(\mathrm{D})^{-1} g + o(\psi)\,.$$

The lemma may be interpreted as saying that if the functions $\mathrm{D}^k g$ have a generalized asymptotic expansion in the the asymptotic scale $\mathrm{D}^k f$, then $f$ has a generalized asymptotic expansion in the asymptotic scale $\mathrm{D}^k g$.

**Proof.** Write $b_k$ the $k$-th coefficient of $P_m(\mathrm{D})^{-1}$. Then

$$P_m(\mathrm{D})^{-1} g = \sum_{0\leq k\leq m} b_k \mathrm{D}^k g\,. \tag{2.8}$$

In this sum, by assumption, we can replace $\mathrm{D}^k g$ by $P_{m-k}(\mathrm{D})\mathrm{D}^k f + o(\psi)$. Since $P_{m-k}(\mathrm{D})\mathrm{D}^k = P_m(\mathrm{D})\mathrm{D}^k$ in $\mathbb{R}_m[\,\mathrm{D}\,]$, the definition of the $b_k$ and (2.8) yield $P_m(\mathrm{D})^{-1} g = f + o(\psi)$, which is the result. ■

We now conclude the proof of our two theorems. Using (2.6) and applying Lemma 2 and a variant of Theorem 2.3.1 in Barbe and McCormick (2004), we obtain for any nonnegative $k$ at most $m$,

$$\overline{F}^{(k)} = \overline{F}_+^{(k)} - \mathsf{L}_{F_-,m-k} \overline{F}_+^{(k)} + o(\mathrm{Id}^{-m}\overline{F}_+) = \mathsf{SL}_{F_-,m-k}\mathrm{D}^{k+1}\overline{F}_+ + o(\mathrm{Id}^{-m}\overline{F}_+)\,.$$

By Veraverbeke's (1977) theorem or (3.0), this implies

$$\mathsf{SL}_{F_-,m-k}\mathrm{D}^k\mathrm{D}\overline{F}_+ = \mathrm{D}^k\overline{F} + o(\mathrm{Id}^{-m+1}\overline{F})\,.$$

Applying Lemma 3, we obtain

$$\mathrm{D}\overline{F}_+ = (\mathsf{SL}_{F_-,m})^{-1}\overline{F} + o(\mathrm{Id}^{-m+1}\overline{F})\,.$$

Hence, integrating, we obtain

$$\overline{F}_+ = (\mathsf{SL}_{F_-,m})^{-1}\mathrm{D}^{-1}\overline{F} + o(\mathrm{Id}^{-m+2}\overline{F})\,. \tag{2.9}$$

This is the equivalent stated in (ii) of Theorem 2.1 and concludes the proof of that theorem.

■

**Proof of Theorem 2.2.** Note that if $N$ is a random variable with geometric distribution with parameter $p$, then $ENL_{H,m}^{N-1}$ is $p(1-p)(\mathrm{Id}-pL_{H,m})^{-2}$ in $\mathbb{R}_m[\,\mathrm{D}\,]$. Representation (2.1) and Theorem 4.4.1 in Barbe and McCormick (2004) yield formula (2.1), that is,

$$\overline{W} = (1-p)(\mathrm{Id} - L_{F_+,m})^{-2}\overline{F}_+ + o(\mathrm{Id}^{-m}\overline{F}_+)\,.$$

To obtain the statement of the Theorem, we again use representation (1.1) and apply Theorem 4.4.1 in Barbe and McCormick (2004) to obtain that if $m$ is less than $\alpha \wedge \omega$,

$$\overline{W} = (1-p)(\mathrm{Id} - L_{F_+,m-1})^{-2}\overline{F}_+ + o(\mathrm{Id}^{-m+1}\overline{F}_+)\,.$$

Then, we use Lemma 2 and (3.6) to conclude. ■

## 2.5 References


[1] Athreya, K.B. and Ney, P. (1972). *Branching Processes*, Springer.

[2] Asmussen, S. (1987). *Applied Probability and Queues*, Wiley.

[3] Barbe,Ph. and McCormick, W.P. (2004). Asymptotic expansions for infinite weighted convolutions of heavy tail distributions and applications, *Mem. Amer. Math. Soc.*, to appear, http://www.arxiv. org/abs/math.PR/0412537.

[4] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1989). *Regular Variation*, 2nd ed. Cambridge University Press.

[5] Chow, Y.S. and Teicher, H. (1978). *Probability Theory, Independence, Interchangeability, Martingales*, Springer.

[6] Doney, R.A. (1980). Moments of Ladder Heights in Random Walks, *J. Appl. Prob.*, 17, 248–252.

[7] Embrechts, P., Goldie, C.M. and Veraverbeke, N. (1979). Subexponentiality and infinite divisibility, *Z. Wahrsch. verw. Geb.*, 49, 335–347.

[8] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events*, Springer.

[9] Embrechts, P. and Veraverbeke, N. (1982). Estimates for the probability of ruin with special emphasis on the possibility of large claims, *Insurance: Mathematics and Economics*, 1, 55–72.

[10] Geluk, J.L. (1992). Second order tail behaviour of a subordinated probability distribution. *Stoch. Proc. Appl.*, 40, 325–337.

[11] Geluk, J.L. (1996). Tails of subordinated laws: The regularly varying case. *Stoch. Proc. Appl.*, 61, 147–161.

[12] Feller, W. (1971). *An Introduction to Probability Theory and its Applications*, 2nd ed., Wiley.

[13] Goldie, C.M. (1991). Implicit renewal theory and tails of solutions of random equations, *Ann. Appl. Probab.*, 1, 126–166.

[14] Greenwood, P., Omey, E. and Teugels, J.L. (1982). Harmonic renewal measures, *Z. Wahrssch. Verw. Gebiete*, 59, 391–409.

[15] Grübel, R. (1987). Harmonic renewal sequences and the first positive sum, *J. London Math. Soc. (2)*, 38, 179–192.

[16] Grübel, R. 1987). On subordinated distributions and generalized renewal measures, *Ann. Probab.*, 15, 394–415.

[17] Marić, V. (2000). *Regular Variation and Differential Equations*, *Lecture Notes in Mathematics*, 1726, Springer.

[18] Omey, E. and Willekens E. (1986). Second order behaviour of the tail of a subordinated probability distribution. *Stoch. Proc. Appl.*, 2, 339–353.

[19] Omey, E. and Willekens E. (1987). Second-order behaviour of distributions subordinate to a distribution with finite mean. *Comm. Statist. Stoch. Models*, 3, 311–342.

[20] Veraverbeke, N. (1977). Asymptotic behavior of Wiener-Hopf factors of a random walk. *Stoch. Proc. Appl.*, 5, 27–37.

# Chapter 3

## Asymptotic Expansions for Distributions of Compound Sums of Random Variables with Rapidly Varying Subexponential Distribution[1]

[1]Barbe, Ph., McCormick, W.P., Zhang, C. Submitted to *Journal of Applied Probability,* 10/2006.

**Abstract**

We derive an asymptotic expansion for the distribution of a compound sum of independent random variables, all having the same rapidly varying subexponential distribution. The examples of a Poisson and geometric number of summands serve as an illustration of the main result. Complete calculations are done for a Weibull distribution, with which we derive, as examples and without any difficulties, 7 terms expansions.

**Keywords:** asymptotic expansion, convolution, tail area approximation, regular variation, subexponential distributions.

## 3.1 Introduction

In this chapter, we construct asymptotic expansions for the tail area of a compound sum, when the distribution of the summands belongs to a class of rapidly varying subexponential distributions. To be more precise, let $X_i$, $i \geq 1$, be a sequence of independent random variables, all having the same distribution $F$. For any positive integer $n$ the distribution of the partial sums $S_n = X_1 + \cdots + X_n$ is the $n$-fold convolution $F^{\star n}$. We set $S_0 = 0$ and therefore $F^{\star 0}$ is defined as the distribution of the point mass at the origin. Let $N$ be a nonnegative integer-valued random variable, independent of the $X_i$'s. We consider the distribution $G$ of the compound sum $S_N$, that is, $EF^{\star N}$, and we are seeking an asymptotic expansion for its tail area $\overline{G} = 1 - G$.

First order asymptotic results for $\overline{G}$ have been obtained by Embrechts, Goldie and Veraverbeke (1979), Cline (1987), and Embrechts (1985). A second order formula may be found in Grübel (1987) and Omey and Willekens (1987).

Compound sums or subordinated distributions arise as distributions of interest in several stochastic models. In insurance risk theory, it models the total claim amount. For a discussion of issues related to random sums and insurance risk, we refer to Embrechts, Klüppelberg and Mikosch (1997), Asmussen (1997), and Goldie and Klüppelberg (1998). Compound sums also appear in queueing theory, in connection with the stationary distribution of waiting times in the GI/G/1 queue. The connection here is not as direct as in the insurance risk model in that it is derived from an analysis of ladder heights for transient random walks; see, for example, Asmussen (1987, p.80), Feller (1971, p.396) and Pakes (1975). Another common way in which this model occurs is through the solution of a transient renewal equation. An example of this occurs in branching processes, where we obtain a geometric-compound sum in the analysis of the mean number of particles alive at a given time in an age-dependent subcritical process; see Athreya and Ney (1972, p.151). We refer to Feller (1971, chapter XI) for a discussion of transient renewal theory. For further applications of subexponentiality in transient renewal theory, we refer to Teugels (1975) and Embrechts and Goldie (1982).

We conclude this introduction by discussing two technical points. While the technique developed in this chapter is applied in a more restrictive setting than that promoted by Grübel (1987), it appears to lead to results in a form more suitable for computation. Moreover, while Grübel (1987), in his improvement of the so-called Banach algebra method pointed out that, 'in principle, this should lead to arbitrarily fine expansions', to our best knowledge the Banach algebra method has yet to be used to produce such explicit fine results comparable to the one we present here for the problem at hand.

A second technical aspect of the current chapter is that for regularly varying tails, Barbe and McCormick (2004) obtained expansions for compound sums based on an expansion for weighted convolutions with an estimate of the remainder term. In contrast to this, for rapidly varying subexponential tail, Barbe and McCormick (2005) does not provide explicit bounds, thereby precluding use of the method from Barbe and McCormick (2004) to obtain applied probability applications to be carried over to the present setting. Thus, one goal of the current chapter is to obtain compound sum tail expansions in the rapidly varying subexponential setting with less information than was available in the regularly varying setting in Barbe and McCormick (2004).

Throughout the chapter, we assume that the $X_i$'s are nonnegative.

## 3.2 Main results

If it exists, the hazard rate $h = F'/\overline{F}$ yields the representation of the distribution function $F$ as

$$\overline{F}(t) = \overline{F}(t_0) \exp\Bigl(-\int_{t_0}^{t} h(u)\,\mathrm{d}u\Bigr)\,.$$

We write Id the identity function on $\mathbb{R}$; for any positive real number $r$, the function $\mathrm{Id}^r$ maps $t$ to $t^r$. From the representation of $\overline{F}$ in terms of its hazard rate, we see that if $h \sim \alpha/\mathrm{Id}$ at infinity, then $\overline{F}$ is regularly varying with index $-\alpha$. If $\lim_{t\to\infty} h(t) = \alpha$, then $\overline{F}(t) = e^{-\alpha t(1+o(1))}$ has a tail behavior close to that of an exponential distribution. Since we are interested in rapidly varying subexponential tails, it is natural to consider hazard rates such that

$$\begin{aligned} &h \text{ is regularly varying,} \\ &\lim_{t\to\infty} th(t) = +\infty \qquad \text{and} \qquad \lim_{t\to\infty} h(t) = 0\,. \end{aligned} \tag{3.1}$$

In order to be not too close to the Pareto type distributions, we will strengthen this assumption by requiring that

$$\liminf_{t\to\infty} th(t)/\log t > 0\,. \tag{3.2}$$

This excludes distributions with tail $e^{-(\log t)^a}$ with $a < 2$, but includes those for which $a \geq 2$. It also includes the subexponential Weibull distributions, or more generally, those with tail of the form $t^{\beta} e^{-t^{\alpha}}$ with $\alpha$ positive and less than 1.

As observed in Barbe and McCormick (2004, 2005), smoothness is a key requirement to obtain asymptotic expansions. For our purposes, a good class of regularly varying functions are the smoothly varying ones of a given order, whose definition we now recall.

**Definition.** *A function $h$ is smoothly varying of index $\alpha$ and order $m$ if it is ultimately $m$-times continuously differentiable and its $m$-th derivative is regularly varying of index $\alpha - m$.*

Clearly, if the hazard rate is $m$ times differentiable, the tail function $\overline{F}$ can be differentiated $m+1$ times.

The next notation we need to introduce pertains to the Laplace characters. We write D the derivation operator; that is, if $g$ is differentiable, D$g$ is its derivative. As is customary, we define $\mathrm{D}^0$ to be the identity, and for any positive integer $i$ we define $\mathrm{D}^i$ by induction as $\mathrm{D}\mathrm{D}^{i-1}$.

We write $\mu_{F,i}$ the $i$-th moment of $F$.

**Definition.** *(Barbe and McCormick, 2004). Let $F$ be a distribution function having at least $m$ moments. Its Laplace character of order $m$ is the differential operator*

$$\mathsf{L}_{F,m} = \sum_{0\leq i\leq m} \frac{(-1)^i}{i!}\mu_{F,i}\mathrm{D}^i\,.$$

Laplace characters have useful algebraic properties which are described in Barbe and McCormick (2004). In particular, consider the ring $\mathbb{R}_m[\,D\,]$ defined as the quotient ring of polynomials in D with real coefficients modulo the ideal generated by $\mathrm{D}^{m+1}$. Laplace characters are elements of this ring, and can be multiplied. It may be helpful to think of a Laplace character as a formal Laplace transform $Ee^{-X\mathrm{D}}$ where $X$ has distribution $F$, expressed as a formal Taylor series in D, dropping all terms in $\mathrm{D}^{m+1}, \mathrm{D}^{m+2}, \ldots$. Then, the multiplication in the ring $\mathbb{R}_m[\,\mathrm{D}\,]$ amounts to the usual multiplication of Taylor series, dropping any term in $\mathrm{D}^{m+1}, \mathrm{D}^{m+2}, \ldots$ In particular, in $\mathbb{R}_m[\,\mathrm{D}\,]$, we have $\mathsf{L}_{H\star K,m} = \mathsf{L}_{H,m}\mathsf{L}_{K,m}$. In what follows, we always consider Laplace characters of order $m$ as members of $\mathbb{R}_m[\,D\,]$, and all the operations on Laplace characters are in that quotient ring.

The following theorem provides an asymptotic expansion for the tail of $G$. A less elegant but more explicit formulation is given as a corollary.

**Theorem 2.1.** *Let $F$ be a distribution function whose hazard rate is smoothly varying with negative index at least $-1$ and positive order $m$. Assume further that (3.2) holds and that the moment generating function of $N$ is finite in a neighborhood of the origin. Then for any nonnegative integer $k$ at most $m$*

$$\overline{G} = EN\mathsf{L}_{F^{\star(N-1)},k}\overline{F} + o(h^k\overline{F})\,.$$

**Remark.** It is shown in Barbe and McCormick (2005, Lemma 4.1.1) that under the assumptions of Theorem 3.2 the asymptotic equivalence $\overline{F}^{(k)} \sim (-1)^k h^k \overline{F}$ holds. Therefore, the remainder term in the above formula could be written as $o(\overline{F}^{(k)})$.

We will see in the next section that the formulation of Theorem 2.1 is quite adequate for practical calculations. A more explicit formulation may be interesting both for understanding the meaning of Theorem 2.1 and for further theoretical developments. To write down such a formulation, we introduce the factorial moments of $N$; writing $(N)_p = N(N-1)\ldots(N-p+1)$ for the falling factorial of order $p$ of $N$, the $p$-th factorial moment of $N$ is $E(N)_p$. Given a multiindex $(k_1, \ldots, k_n)$ of nonnegative integers and extending the classical notation for the multinomial coefficients, we write $\binom{n}{k}$ or $\binom{n}{k_1 \ldots k_n}$ for $n!/k_1! \ldots k_n!$, and the norm $|k| = k_1 + \cdots + k_n$. We define the following coefficients, involving only the moments of $F$ and the factorial moments of $N$: when $i = 0$, we set $g_0 = EN$ and for any $i$ positive,

$$\begin{aligned} g_i &= \frac{(-1)^i}{i!} \sum_{\substack{k_1+2k_2+\cdots+ik_i=i \\ k_1,\ldots,k_i\geq 0}} \binom{i}{k} E(N)_{|k|+1} \Big(\frac{\mu_{F,1}}{1!}\Big)^{k_1} \cdots \Big(\frac{\mu_{F,i}}{i!}\Big)^{k_i} \\ &= \frac{(-1)^i}{i!} \sum_{1\leq p\leq i} \frac{E(N)_{p+1}}{p!} \sum_{\substack{j_1+\cdots+j_p=i \\ j_1,\ldots,j_p>0}} \binom{i}{j_1 \ldots j_p} \mu_{F,j_1} \ldots \mu_{F,j_p} \, . \end{aligned}$$

The following corollary asserts that $\overline{G}$ has an asymptotic expansion in the natural asymptotic scale $(\overline{F}^{(i)})_{0\leq i\leq m}$, whose coefficients are given by the $g_i$.

**Corollary.** *Under the assumptions of Theorem 2.1,*

$$\overline{G} = \sum_{0\leq i\leq m} g_i \overline{F}^{(i)} + o(h^m \overline{F}) \, .$$

**Proof of the Corollary.** Consider the probability generating function $\Lambda(t) = Et^N$ of $N$ and the Laplace transform $\mathcal{L}(t) = Ee^{-tX}$ of $X$. Considered as an element in $\mathbb{R}_m[\,D\,]$, the operator $\mathcal{L}(\mathrm{D})$ coincides with the Laplace character $L_{F,m}$. Consequently, introducing the derivative $\Lambda'$ of $\Lambda$, in $\mathbb{R}_m[\,\mathrm{D}\,]$

$$ENL_{F^{\star(N-1)},m} = ENL_{F,m}^{N-1} = \Lambda'\big(\mathcal{L}(\mathrm{D})\big) \, .$$

Given Theorem 2.1, it suffices to evaluate the right hand side of this equality. For this purpose, consider the function defined in a neighborhood of the origin of the real line $g(t) = \Lambda'\big(\mathcal{L}(t)\big)$. This function has a Taylor expansion $\sum_{i\geq 0} t^i g^{(i)}(0)/i!$. Applying Faà di Bruno formula (see e.g. Roman, 1980), we see that $g^{(i)}(0) = i!g_i$. The result follows. ■

## 3.3 Examples

We illustrate the use of Theorem 2.1, considering the cases where $N$ has a Poisson and a geometric distribution.

*Example 1.* Assume that $N$ has a Poisson distribution with parameter $a$. Sums with a Poisson number of summands are commonly used in insurance mathematics, modelling total claim size (Beirlant et al., 1996, Embrechts et al., 1997, Willmot and Lin, 2000). The following expansion is easily derived.

**Proposition 3.1.** *Let $F$ be a distribution function satisfying the assumptions of Theorem 2.1. If $N$ has a Poisson distribution with parameter $a$, then $\overline{G} = aL_{G,m}\overline{F} + o(h^m\overline{F})$. Moreover, $L_{G,m} = e^{a(L_{F,m}-\mathrm{Id})}$.*

**Proof.** Combine Theorem 2.1 and the proof of Corollary 4.4.2 in Barbe and McCormick (2004) to obtain the expansion $aL_{G,m}\overline{F}$. To obtain the expression for $L_{G,m}$, write, in the quotient ring,

$$ENL_{F,m}^{N-1} = e^{-a}\sum_{n\geq 1} n\frac{a^n}{n!}L_{F,m}^{n-1} = ae^{a(L_{F,m}-\mathrm{Id})}\,. \qquad ■$$

The above formula is easily implemented with a computer algebra system. For example, the following `Maple` code calculates $ae^{a(L_{F,m}-\mathrm{Id})}$.

```
mu[0]:=1:
LF:=sum('(-1)^j*mu[j]*D^j/j!','j'=0..m+1):
taylor(a*exp(a*(LF-1)),D=0,m+1);
```

Setting $m=3$ in the previous code yields the first four terms,

$$\begin{aligned} E(N\mathsf{L}_{F,4}^{N-1}) &= a\mathrm{Id} - a^2\mu_{F,1}\mathrm{D} + \frac{a^2}{2}(a\mu_{F,1}^2+\mu_{F,2})\mathrm{D}^2 \\ &\qquad - \frac{a^2}{6}(a^2\mu_{F,1}^3 + 3a\mu_{F,1}\mu_{F,2} + \mu_{F,3})\mathrm{D}^3\,. \end{aligned}$$

To give a very concrete example, assume that $\overline{F}$ is the Weibull distribution with parameter $1/3$, so that $\overline{F}(t) = e^{-t^{1/3}}$. Define $e_r(t) = t^r e^{-t^{1/3}}$. We obtain, after evaluation of $E(N\mathsf{L}_{F,4}^{N-1})$, and using a computer algebra package,

$$\begin{aligned} \overline{G} &= ae_0 + 2a^2e_{-2/3} + 2a^2(20+a)e_{-4/3} + 4a^2(20+a)e_{-5/3} \\ &\quad + \frac{4a^2(1680+60a+a^2)}{3}e_{-2} + 8a^2\bigl(1680+60a+a^2\bigr)e_{-7/3} \\ &\quad + \frac{2a^2(403200+9120a+140a^2+a^3)}{3}e_{-8/3} + o\bigl(e_{-8/3}\bigr)\,. \end{aligned}$$

Perhaps the only remarkable feature of such 7 terms expansion is that it can be done.

*Example 2.* Motivated by applications to queueing theory (see e.g., Cohen, 1972, or Bingham, Goldie and Teugels, 1987, p.387), consider the case where $N$ has a geometric distribution with parameter $a$, that is, $N$ is a nonnegative integer $n$ with probability $(1-a)a^n$. Again, Theorem 2.1 provides a compact expression of the asymptotic expansion of $\overline{G}$, and the issue is how to actually compute it.

Any polynomial in D with nonvanishing constant term is invertible in the quotient ring $\mathbb{R}_m[\,\mathrm{D}\,]$. Therefore, since $a$ is positive and less than 1,

$$EN\mathsf{L}_{F^{\star(N-1)},m} = (1-a)\sum_{n\geq 1} a^n n \mathsf{L}_{F,m}^{n-1} = a(1-a)(\mathrm{Id} - a\mathsf{L}_{F,m})^{-2}\,.$$

Consequently, the following result holds.

**Proposition 3.2** *Let $F$ be a distribution function satistfying the assumptions of Theorem 2.1. If $N$ has a geometric distribution with parameter $a$, then $\overline{G} = a(1-a)(\mathrm{Id} - a\mathsf{L}_{F,m})^{-2}\overline{F} + o(h^m\overline{F})$.*

Setting $m=3$, we obtain, as in the previous example, with the help of a computer algebra package, with $b = a/(1-a)$,

$$EN\mathsf{L}_{F,3}^{N-1} = b\,\mathrm{Id} - 2b^2\mu_{F,1}\mathrm{D} + b^2(\mu_{F,2} + 3b\mu_{F,1}^2)\mathrm{D}^2$$

$$-\frac{b^2}{3}(12b^2\mu_{F,1}^3 + 9b\mu_{F,1}\mu_{F,2} + \mu_{F,3})\mathrm{D}^3 \,.$$

For instance, when $F$ is the Weibull distribution with parameter $1/3$, the calculation of $ENL_{F,4}^{N-1}$ yields the following 7 terms expansion, where $e_r(t) = t^r e^{-t^{1/3}}$ — expressed solely with $a$, the formula contains alternating signs; expressing it with $b = a/(1-a)$ makes it slightly more stable numerically —

$$\begin{aligned}
\overline{G} \;=\; & b\, e_0 + 4b^2\, e_{-2/3} + 4b^2(20+3b)\, e_{-4/3} + 8b^2(20+3b)\, e_{-5/3} \\
& + 32b^2(140+15b+b^2)\, e_{-2} + 192b^2(140+15b+b^2)\, e_{-7/3} \\
& + 80b^2(6720+456b+28b^2+b^3)\, e_{-8/3} + o(e_{-8/3}) \,.
\end{aligned}$$

## 3.4 Proof of the Theorem.

When $m$ vanishes, Theorem 2.1 is due to Embrechts, Goldie and Veraverbeke (1979, p.342). Therefore, we will prove it when $m$ is at least 1.

It is convenient to introduce a pseudo-semi-norm on tails. If $K$ is a distribution function, we write

$$|\overline{K}|_F = \sup_{t\geq 0} (\overline{K}/\overline{F})(t) \,,$$

with the convention $0/0 = 0$. This generates balls $B(F,r)$ containing all tails $\overline{K}$ which are less than $r\overline{F}$. We write $B(F)$ the union of all these balls for all positive $r$.

We write $G_n$ for the $n$-fold convolution $F^{\star n}$.

We start by recalling Kesten's global bound on tail function of self-convolutions of subexponential distributions; see Athreya and Ney (1972, § IV.4, Lemma 7). It asserts that for any positive $\epsilon$ there exists a positive $A$ such that for all positive integers $n$,

$$|\overline{G}_n|_F \leq A(1+\epsilon)^n \,. \tag{3.3}$$

We also need a precise estimate of the order of magnitude of derivatives of $\overline{F}$. As noted in the Remark following Theorem 2.1, Lemma 4.1.1 in Barbe and McCormick (2005) showed that for any nonnegative $k$ at most $m$,

$$\overline{F}^{(k)} \sim (-1)^k h^k \overline{F}\,. \tag{3.4}$$

Finally, we also need a basic representation of convolution in terms of operators. For any distribution function $K$ with support in the nonnegative half-line and any $\eta$ positive and less than 1, define the operator

$$\mathsf{T}_{K,\eta} f(t) = \int_0^{\eta t} f(t-x)\,\mathrm{d}K(x)\,.$$

For any positive $c$ we also define the multiplication operator $\mathsf{M}_c$ acting on functions by

$$\mathsf{M}_c f(t) = f(t/c)\,.$$

These two operators allow us to write a convolution in a way suitable for our analysis. Define the powers $\mathsf{T}_{K,\eta}^n$ by $\mathsf{T}_{K,\eta}^0 = \mathrm{Id}$ and $T_{K,\eta}^{n+1} = T_{K,\eta} T_{K,\eta}^n$. Using Proposition 5.1.1 in Barbe and McCormick (2004) inductively, we obtain a representation for the distribution function, valid on the nonnegative half line,

$$\overline{G}_n = \sum_{1\leq i\leq n} \mathsf{T}_{F,\eta}^{i-1} \mathsf{T}_{G_{n-i},1-\eta} \overline{F} + \sum_{1\leq i\leq n} \mathsf{T}_{F,\eta}^{i-1} \big( \mathsf{M}_{1/\eta} \overline{F} \mathsf{M}_{1/(1-\eta)} \overline{G}_{n-i} \big)\,. \tag{3.5}$$

Our first lemma is a simple moment bound.

**Lemma 4.1** *Let $i$ be a nonnegative integer, and let $\epsilon$ be a positive real number. There exists $t_1$ such that for any $t$ at least $t_1$ and any distribution function $K$ in $B(F)$,*

$$\int_t^\infty x^i\,\mathrm{d}K(x) \leq (1+\epsilon)|\overline{K}|_F t^i \overline{F}(t)\,.$$

**Proof.** For any nonnegative integer $i$, an integration by parts yields

$$\int_t^\infty x^i\,\mathrm{d}K(x) = t^i\overline{K}(t) + i\int_t^\infty x^{i-1}\overline{K}(x)\,\mathrm{d}x\,. \tag{3.6}$$

The right hand side of this equality is less than $|\overline{K}|_F$ times the same expression with $K$ replaced by $F$. Consequently, it suffices to prove the result when $K$ is $F$. In that case, let $M$

be a positive real number so that $\epsilon(M-i)\geq i$. Since $\mathrm{Id}\, h$ tends to infinity at infinity, $h$ is more than $M/\mathrm{Id}$ ultimately. For any $t$ large enough and any $x$ at least $t$,

$$\frac{\overline{F}(x)}{\overline{F}(t)}=\exp\Bigl(-\int_t^x h(u)\,\mathrm{d}u\Bigr)\leq\Bigl(\frac{t}{x}\Bigr)^M .$$

This implies that the integral in the right hand side of (3.6), when $F$ is substituted for $K$, is at most $\epsilon t^i\overline{F}(t)$. ■

Our next lemma contains the main argument of the proof, namely that a $\mathsf{T}_{K,\eta}$ operator is in some sense very close to a Laplace character as far as tail behavior is concerned when applied to $\overline{F}$ and its derivatives.

**Lemma 4.2.** *For any fixed integer $p$ at most $m$,*

$$\lim_{t\to\infty}\sup_{K\in B(F)}\frac{|(\mathsf{T}_{K,\eta}-\mathsf{L}_{K,m-p})\overline{F}^{(p)}|}{|\overline{K}|_F h^m\overline{F}}(t)=0 .$$

**Proof.** The proof of Lemma 4.2.3 in Barbe and McCormick (2005) showed that for any $\delta$ positive,

$$\Bigl|\int_0^{\delta/h(t)}\overline{F}^{(p)}(t-x)\,\mathrm{d}K(x)-\mathsf{L}_{K,m-p}\overline{F}^{(p)}\Bigr| \tag{3.7}$$

is at most

$$\sum_{0\leq j\leq m-p}|\overline{F}^{(p+j)}(t)|\int_{\delta/h(t)}^\infty x^j\,\mathrm{d}K(x) \tag{3.8}$$

$$+\int_0^{\delta/h(t)}\int_0^x\frac{y^{m-p-1}}{(m-p-1)!}|\overline{F}^{(m)}(t-x+y)-\overline{F}^{(m)}(t)|\,\mathrm{d}y\,\mathrm{d}K(x) . \tag{3.9}$$

Let $\epsilon$ be a positive number. Using Lemma 3.4 and (3.4), we see that for large $t$, the term (3.8) is less than

$$\overline{F}(t)2|\overline{K}|_F\Bigl(\frac{\delta}{h(t)}\Bigr)^j\overline{F}\Bigl(\frac{\delta}{h(t)}\Bigr) .$$

Since $\overline{F}$ is rapidly varying, this is ultimately less than $\epsilon|\overline{K}|_F h^m\overline{F}$.

The proof of Lemma 4.2.3 in Barbe and McCormick (2005) showed that for $\delta$ small enough, for any $t$ large enough and for any $K$ in $B(F)$, the double integral (3.9) is at most $\epsilon|\overline{K}|_F\mu_{F,m-p}h^m\overline{F}$. Hence, we have shown that (3.7) is at most $\epsilon|\overline{K}|_F(\mu_{F,m-p}+1)h^m\overline{F}$ ultimately uniformly over $B(F)$.

The proof of Lemma 4.2.4 in Barbe and McCormick (2005) showed that for any positive $\delta$ and $\eta$, ultimately uniformly over $B(F)$,

$$\int_{\delta/h(t)}^{\eta t} |\overline{F}^{(p)}(t-x)|\,\mathrm{d}K(x) \leq \epsilon|\overline{K}|_F h^m\overline{F}(t)\,.$$

This proves the Lemma 4.2. ■

Lemma 4.2. yields the following estimate on an operator $\mathsf{T}$ composed with a Laplace character applied to a derivative of $\overline{F}$.

**Lemma 4.3.** *The following uniform limit holds:*

$$\lim_{t\to\infty}\sup_{K\in B(F)} \frac{|\mathsf{T}_{K,\eta}\mathsf{L}_{H,m-p}\overline{F}^{(p)} - \mathsf{L}_{K\star H,m-p}\overline{F}^{(p)}|}{|\overline{K}|_F\,h^m\overline{F}\,\sum_{0\leq j\leq m-p}\frac{\mu_{H,j}}{j!}} = 0\,.$$

**Proof.** Since $\mathsf{T}_{K,\eta}$ is linear and

$$\mathsf{L}_{H,m-p}\overline{F}^{(p)} = \sum_{0\leq j\leq m-p}\frac{(-1)^j}{j!}\mu_{H,j}\overline{F}^{(p+j)}\,,$$

the result follows from Lemma 4.1 and Lemma 2.1.4 in Barbe and McCormick (2004). ■

The next two lemmas will take care of some remainder terms. The first one asserts that terms of order $o(h^m\overline{F})$ remain so through the action of some $\mathsf{T}$ operators.

**Lemma 4.4.** *Let $q$ be a nonnegative integer and $\epsilon$ be a positive real number. There exist $t_2$, some positive $A$ and $\eta$, such that for any positive integer $i$,*

$$\mathsf{T}^i_{F,\eta}(h^q\overline{F}) \leq A(1+\epsilon)^i h^q\overline{F}$$

*on $[\,t_2,\infty)$.*

**Proof.** Let $\epsilon$ be a positive real number. Since $h$ is regularly varying with negative index, provided $\eta$ is small enough, $h(t-x) \leq (1+\epsilon)h(t)$ for any $t$ large enough and any $x$ nonnegative and at most $\eta t$.

Therefore, for $t$ at least $t'_2$,

$$\begin{aligned} \mathsf{T}_{F,\eta}(h^q\overline{F})(t) &= \int_0^{\eta t} h^q\overline{F}(t-x)\,\mathrm{d}F(x) \\ &\leq (1+\epsilon)h^q(t)\int_0^{\eta t}\overline{F}(t-x)\,\mathrm{d}F(x) \\ &\leq (1+\epsilon)h^q(t)\overline{F}^{\star 2}(t)\,. \end{aligned}$$

By induction, it follows that

$$\mathsf{T}^i_{F,\eta}(h^q\overline{F})(t) \leq (1+\epsilon)^i\big(h^q\overline{F}^{\star(i+1)}\big)(t)\,.$$

Using Kesten's bound, (3.3) above, this yields that $\mathsf{T}^i_{F,\eta}(h^q\overline{F})$ is ultimately at most $A(1+\epsilon)^{2i}h^q\overline{F}$, finishing the proof since $\epsilon$ is arbitrary. ■

Our penultimate lemma will be used to handle the terms involving the multiplication operators in (3.5).

**Lemma 4.5.** *Let $\epsilon$ be a positive real number. There exists $t_3$ such that for any positive integers $i$ and $m$,*

$$|\mathsf{M}_{1/\eta}\overline{F}\mathsf{M}_{1/(1-\eta)}\overline{G}_i| \leq (1+\epsilon)^i h^{m+1}\overline{F}$$

*on $[\,t_3,\infty)$.*

**Proof.** Kesten's bound in (3.3) shows that

$$\big|\overline{F}(t\eta)\overline{G}_i\big(t(1-\eta)\big)\big| \leq \overline{F}(t\eta)A(1+\epsilon)^i\overline{F}\big(t(1-\eta)\big)\,.$$

Arguing as in Lemma 4.2.1 in Barbe and McCormick (2005), $\overline{F}(t\eta)\overline{F}\big(t(1-\eta)\big)$ is $o\big(h^q\overline{F}(t)\big)$ for any positive $q$. This implies the result. ■

Our last lemma is stated merely to avoid digression in the argument later on.

**Lemma 4.6.** *Let $\epsilon$ be a positive number. There exists $A$ such that for any positive integer $n$*

$$\sum_{0\leq j\leq m} \frac{\mu_{G_n,j}}{j!} \leq A(1+\epsilon)^n .$$

**Proof.** The lemma follows from Marcinkiewicz-Zygmund's inequality (see Chow and Teicher, 1988, §10.3, Theorem 3), which implies that $\mu_{G_n,j} \leq An^j$ for some constant $A$. ■

We can now conclude the proof of Theorem 2.1. Combining Lemmas 4.4 and 4.5, there exists an interval $[\,t_3,\infty)$ on which for any $j$ and $k$ with $0\leq j\leq k\leq n$, any positive $i$ and $n$ with $i\leq n$,

$$|\,\mathsf{T}_{F,\eta}^{i-1}(\mathsf{M}_{1/\eta}\overline{F}\mathsf{M}_{1/(1-\eta)}\overline{G}_{n-i})| \leq A(1+\epsilon)^n h^{m+1}\overline{F} .$$

Representations (3.5) yield, on $[\,t_3,\infty)$,

$$|\overline{G}_n - \sum_{1\leq i\leq n} \mathsf{T}_{F,\eta}^{i-1}\mathsf{T}_{G_{n-i},1-\eta}\overline{F}| \leq An(1+\epsilon)^n h^{m+1}\overline{F} . \tag{3.10}$$

Let $\epsilon$ be a positive real number, small enough so that $E(1+2\epsilon)^N$ is finite. Let $\delta$ be a positive real number. Combining Lemmas 4.1, 4.3, 4.4 and 4.6, using also Kesten's bound, ultimately, uniformly in $n$ and $i$ at most $n$,

$$\begin{aligned}
&|\,\mathsf{T}_{F,\eta}^{i-1}\mathsf{T}_{G_{n-i},1-\eta}\overline{F} - \mathsf{T}_{F,\eta}^{i-2}\mathsf{L}_{G_{n-i+1},m}\overline{F}| \\
\leq\ & \mathsf{T}_{F,\eta}^{i-1}|(\mathsf{T}_{G_{n-i},1-\eta} - \mathsf{L}_{G_{n-i},m})\overline{F}| + \mathsf{T}_{F,\eta}^{i-2}|(\mathsf{T}_{F,\eta}\mathsf{L}_{G_{n-i},m} - \mathsf{L}_{G_{n-i+1},m})\overline{F}| \\
\leq\ & 2\delta(A^2+A)(1+\epsilon)^n h^m\overline{F} . 
\end{aligned} \tag{3.11}$$

Using the same combination of lemmas, we also have, ultimately, uniformly in $n$ and $j$ at most $n-1$,

$$|\,\mathsf{T}_{F,\eta}^{j}\mathsf{L}_{G_{n-j-1},m}\overline{F} - \mathsf{T}_{F,\eta}^{j-1}\mathsf{L}_{G_{n-j},m}\overline{F}| \leq A^2(1+\epsilon)^n\delta h^m\overline{F} . \tag{3.12}$$

We take $A$ to be at least 1, simply to ensure that $A^2$ is more than $A$. Summing (3.12) for $j$ positive and less than $i$ and adding (3.11), we obtain

$$|\,\mathsf{T}_{F,\eta}^{i-1}\mathsf{T}_{G_{n-i},1-\eta}\overline{F} - \mathsf{L}_{G_{n-1},m}\overline{F}| \leq 4A^2\delta h^m\overline{F}i(1+\epsilon)^n$$

on some interval $[\,t_4, \infty)$. Summing these inequalities for $i$ positive and at most $n$ and combining with (3.10) yield

$$|\overline{G}_n - n\mathsf{L}_{G_{n-1},m}\overline{F}| \leq 10A^2\delta n(n+1)(1+\epsilon)^n h^m \overline{F}\,.$$

Since the moment generating function of $N$ is finite at $\log(1+2\epsilon)$ and $\delta$ is arbitrary, Theorem 2.1 follows. ■

## 3.5 References


[1] Asmussen, S. (1987). *Applied Probability*, Wiley.

[2] Asmussen, S. (1997). *Ruin Probabilities*, World Scientific.

[3] Athreya, K.B. and Ney, P.E. (1972). *Branching Processes.* Springer.

[4] Barbe, Ph. and McCormick, W.P. (2004). Asymptotic expansions for infinite weighted convolutions of heavy tail distributions and applications, *Mem. Amer. Math. Soc.*(in press), http://www.arxiv. org/abs/math.PR/0412537.

[5] Barbe, Ph. and McCormick, W.P. (2005). Asymptotic expansions for convolutions of light tailed subexponential distributions, *Prob. Theory Rel. Fields*, to appear, http://www. arxiv.org/abs/math.PR/0512141.

[6] Beirlant, J., Teugels, J.L. and Vynckier, P. (1996). *Practical Analysis of Extreme Values*, Leuven University Press, Leuven, Belgium.

[7] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1989). *Regular Variation*, 2nd ed., Cambridge.

[8] Cline, D.B.H. (1987). Convolutions of distributions with exponential and subexponential tails, *J. Austr. Math. Soc. (A)*, 43, 347–365.

[9] Chow, Y.S. and Teicher, H. (1988). *Probability Theory, Independence, Interchangeability, Martingales*, 2nd ed., Springer.

[10] Cohen, J.W. (1972). On the tail of the stationary waiting-time distribution and limit theorem for M/G/1 queue, *Ann. Inst. H. Poincaré,B*, 8, 255–263.

[11] Embrechts, P. (1985). Subexponential distribution functions and their applications: a review, in *Proceedings of the Seventh Conference on Probability Theory, (Braşov, 1982)*, 125–136, VNU Sci. Press, Utrecht, 1985.

[12] Embrechts, P. and Goldie, C.M. (1982). On convolution tails, *Stoch. Proc. Appl.*, 13, 263–278.

[13] Embrechts, P., Goldie, C.M. and Veraverbeke, N. (1979). Subexponentiality and infinite divisibility, *Z. Wahrsch. verw. Geb.*, 49, 335–347.

[14] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events*, Springer.

[15] Feller, W. (1971). *An Introduction to Probability and its Applications*, 2nd ed., Wiley.

[16] Goldie, C. and Klüppelberg, C. (1998). Subexponential distributions, in *A Practical Guide to Heavy Tails, Statistical Techniques and Applications*, R. Adler, R. Feldman and M. Taqqu eds., Birkhäuser, 435–460.

[17] Grübel, R. (1987). On subordinated distributions and generalized renewal measures, *Ann. Probab.*, 15, 394–415.

[18] Omey, E. and Willekens E. (1987). Second-order behaviour of distributions subordinate to a distribution with finite mean, *Comm. Statist. Stoch. Models*, 3, 311-342.

[19] Pakes, A.G. (1975). On the tails of waiting-time distributions, *J. Appl. Probab.*, 12, 555–564.

[20] Roman, S. (1980). The formula of Faà di Bruno, *Amer. Math. Monthly*, 87, 805–809.

[21] Teugels, J.L. (1975). The class of subexponential distributions, *Ann. Probab.*, 3, 1000–1011.

[22] Willmot, G.E. and Lin, X.S. (2000). *Lundberg Approximations for Compound Distributions with Insurance Applications*, *Lecture Notes in Statistics*, 156, Springer.

# Chapter 4

## Asymptotic Properties of Hill's Estimator for Shot Noise Sequence [1]

[1]Zhang, C., McCormick, W.P., Submitted to *Journal of Statistical Planning and Inference,* 5/2007.

**Abstract**

We study a shot noise sequence of the form $X_j = \sum_{i\le j} h(\tau_j - \tau_i)A_i$, where the distribution of $A_i$ has regularly varying tail with index $-\alpha$. Under a mild integration condition on $h$ the distribution of $X_i$ also has regularly varying tail with index $-\alpha$ and the Hill's estimator is a weakly consistent estimator of $1/\alpha$. If the tail distribution of $X_i$ is normalized and second order regularly varying the Hill's estimator is asymptotically normal for a large class of functions $h$.

**Keywords:** shot noise, heavy tail, regular variation, tail index, Hill's estimator

## 4.1 Introduction

Shot noise processes are flexible models capable of capturing stochastic features for a variety of natural phenomena. Applications of such models have been made to model traffic flow (Bartlett (1963)), computer failure times (Lewis (1964)), earthquake after shocks (Vere-Jones (1970)), river flow (Lawrance and Kottegoda (1977)), storage models (Lund (1996)) as well as many more diverse applications. Parzen (1962) provides further examples. More recently, shot noise models have been applied in modeling financial time series and in insurance risk. Being a natural generalization of a compound Poisson process, nonstationary shot noise processes are particularly well suited as a model in insurance risk theory. See, for example, Klüppelberg et al. ((2003). Heavy tail behavior, clustering, long-range dependence are properties financial data exhibit and Samorodnitsky (1998) has demonstrated that these features can be captured by shot noise processes.

A general shot noise process can be represented as a superposition of independently and identically distributed (i.i.d.) stochastic processes with time index shifted by the locations of an independent process. A useful specialization of the above model takes the form

$$X(t) = \sum_{\tau_i \le t} A_i h(t - \tau_i),$$

where $\tau_j$'s are renewal points, $\{A_i\}$ is a sequence of i.i.d. positive random variables (r.v.'s), independent of the renewal process. $h$ is a real valued function on $[0, \infty)$ and is called the impulse response function.

We study the following shot noise sequence in this paper

$$X_j = \sum_{i \leq j} A_i h(\tau_j - \tau_i). \tag{4.1}$$

In this paper we assume that the right tail of the distribution of $A_i$ is regularly varying with negative index $-\alpha$. Under a mild condition on $h$, we know that the tail distribution of $X_i$ is also regularly varying with the same index $-\alpha$. A well studied estimator of the reciprocal of tail index, Hill's estimator (Hill (1975)), is defined as

$$\frac{1}{k} \sum_{i=1}^{k} \log \frac{X_{(i)}}{X_{(k+1)}},$$

where $X_{(i)}$'s are upper order statistics of $X_1, \cdots, X_n$. For example, $X_{(1)}$ is the largest value of $X_1, \cdots, X_n$. de Haan and Resnick (1998) established asymptotic normality for iid data. For dependent data Hsing (1991), Rootzen et al. (1990) studied the asymptotic properties of Hill's estimator under weak dependence assumptions. Resnick and Stărică (1995,1998) proved the weak consistency of Hill's estimator for certain classes of stationary sequence while Resnick and Stărică (1997), and Ling and Peng (2004) studied the asymptotic normality of Hill's estimator for AR and ARMA models.

## 4.2 Results

The following conditions will be imposed throughout our paper. We refer to these collectively as the basic assumptions.

1. $\{A_j, -\infty < j < \infty\}$ denotes a sequence of positive i.i.d. random variables, which have regularly varying tail at infinity with negative tail index, $-\alpha$.

2. $\{\tau_j, -\infty < j < \infty\}$ is a renewal process with a fixed renewal at $\tau_0 = 0$, and is independent of $\{A_j, -\infty < j < \infty\}$.

3. the impulse response function $h$ is nonnegative and nonincreasing on $[0, \infty)$.

Without loss of generality we assume $h(0) = 1$. We write $C_{ji} = h(\tau_j - \tau_i)$, $j \geq i$, for typographical ease.

If $\int_0^\infty h^{(\alpha\wedge 1)-\delta}(s)\,\mathrm{d}s < \infty$ for some $\delta > 0$, it can be shown that $X_j$'s are almost surely finite and

$$\lim_{x\to\infty} \frac{P\{X_j > x\}}{P\{A_1 > x\}} = \sum_{i\leq j} EC_{ji}^{\alpha} < \infty.$$

Then the distribution of $X_j$ has regularly varying tail with index $-\alpha$.

For completeness and since its proof can be established following the procedure in Resnick and Stărică [59], we mention without proof the following weak consistency result.

**Theorem 2.1.** *Under the basic assumptions, if* $\int_0^\infty h^{(\alpha\wedge 1)-\delta}(s)\,\mathrm{d}s < \infty$ *for some* $\delta > 0$*, we have the weak consistency of Hill's estimator, i.e. for any sequence of positive integers* $\{k := k_n\}$ *such that* $k \wedge \frac{n}{k} \to \infty$

$$\frac{1}{k}\sum_{i=1}^{k} \log \frac{X_{(i)}}{X_{(k+1)}} \xrightarrow{P} \frac{1}{\alpha}.$$

In order to get asymptotic normality of Hill's estimator we further assume that **(i)** the marginal distribution of the shot noise sequence has a density $F'_{X_1}$ satisfying the Von Mises condition $\lim\limits_{t\to\infty} \frac{tF'_{X_1}(t)}{1-F_{X_1}(t)} = \alpha$, or equivalently, $\overline{F}_{X_1} = 1 - F_{X_1}$ is differentiable and normalized regularly varying, **(ii)** $\overline{F}_{X_1}$ is second order regularly varying, which is to say that there exists an $\alpha < 0$, $\rho \leq 0$ and $L$ a slowly varying function such that $\overline{F}_{X_1}(x) = x^{-\alpha}L(x)$ and $\lim_{t\to\infty}\left(\frac{L(tx)}{L(t)} - 1\right)/g(t) = K\frac{x^\rho - 1}{\rho}$ for all $x > 0$, where $g$ is a regularly varying function with index $\rho$ and $\lim_{t\to\infty} g(t) = 0$. In the case that $\rho = 0$, $\frac{x^\rho-1}{\rho}$ is to be interpreted as $\log x$. We choose $k := k_n$ such that $\lim_{n\to\infty}\sqrt{k}g\left(b\left(\frac{n}{k}\right)\right) = 0$, where $b(t) = F^{\leftarrow}_{X_1}(1 - \frac{1}{t})$ and $F^{\leftarrow}$ denotes the left continuous inverse function given by $F^{\leftarrow}(t) = \inf\{s : F(s) \geq t\}$.

**Theorem 2.2.** *For shot noise sequence $\{X_j\}$ as specified above and $k \in (n^{\delta'}, n^{1-\delta'})$ for some $\delta' > 0$, if*

$$\int_0^\infty s^{\gamma+(\alpha+2\delta)I_{\{\alpha\geq 1\}}} h^{(\alpha\wedge 1)-\delta}(s)\,\mathrm{d}s < \infty \tag{4.2}$$

*for some $\gamma > 2$, $\delta > 0$ and $\left(k^{\frac{1}{2}} \wedge \frac{n}{k}\right)^{\gamma}/k \to \infty$ we have the asymptotic normality of Hill's estimator*

$$\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k}\log\frac{X_{(i)}}{X_{(k+1)}} - \frac{1}{\alpha}\right) \xrightarrow{d} N(0,\lambda),$$

*where $\lambda = \frac{1}{\alpha^2}\left(1 + \frac{\sum_{i\geq 1} 2iEh^{\alpha}(\tau_i)}{\sum_{i\geq 0} Eh^{\alpha}(\tau_i)}\right)$.*

## 4.3 Proof

Let M denote a generic constant which may change value from appearance to appearance.

**Proposition 3.1.** *For a renewal sequence $\tau_n$ with $\tau_0 = 0$ and positive interarrival mean, if $\int_0^\infty s^p h^{\beta/\gamma'}\,\mathrm{d}s < \infty$ for some $\gamma' \geq 1, \beta > 0$ and $p \geq 0$ then*

$$\sum_{n\geq 0}(n+1)^p\big(Eh^{\beta}(\tau_n)\big)^{\frac{1}{\gamma'}} < \infty.$$

**Proof: Proof:** Since $h$ is a nonincreasing function we may assume that $0 \leq \tau_i - \tau_{i-1} \leq 1$. Let $\mu = E\tau_1$. By Markov's inequality we get

$$\begin{aligned}
&\sum_{n\geq 0}(n+1)^p\big(Eh^{\beta}(\tau_n)\big)^{\frac{1}{\gamma'}} \\
\leq\ & h^{\beta/\gamma'}(0) + 2^p\sum_{n\geq 1} n^p\big(Eh^{\beta}(\tau_n)I_{\{|\tau_n-\mu n|<\frac{\mu n}{2}\}}\big)^{\frac{1}{\gamma'}} + 2^p h^{\beta/\gamma'}(0)\sum_{n\geq 1} n^p\big(EI_{\{|\tau_n-\mu n|\geq\frac{\mu n}{2}\}}\big)^{\frac{1}{\gamma'}} \\
\leq\ & h^{\beta/\gamma'}(0) + 2^p\sum_{n\geq 1} n^p h^{\beta/\gamma'}\left(\frac{\mu n}{2}\right) + M\sum_{n\geq 1} n^p\Big(\frac{E|\tau_n-\mu n|^m}{n^m}\Big)^{\frac{1}{\gamma'}}.
\end{aligned}$$

Note that $n^p h^{\beta/\gamma'}\left(\frac{\mu n}{2}\right) \leq 2^p s^p h^{\beta/\gamma'}\left(\frac{\mu s}{2}\right)$ for any $n-1 \leq s \leq n$ and $n \geq 2$. Thus

$$\sum_{n\geq 1} n^p h^{\beta/\gamma'}\left(\frac{\mu n}{2}\right) \ \leq\ h^{\beta/\gamma'}\left(\frac{\mu}{2}\right) + 2^p\int_1^\infty s^p h^{\beta/\gamma'}\left(\frac{\mu s}{2}\right)\mathrm{d}s,$$

which is finite by assumption. The Marcinkiewicz-Zygmund's inequality provides that $E|\tau_n - \mu n|^m = O\big(n^{\frac{m}{2}}\big)$ (see Corollary 10.3.2, Chow and Teicher [4]). Let $m > 2\gamma'(p+1)$. We obtain

$$\sum_{n\geq 1} n^p\Big(\frac{E|\tau_n-\mu n|^m}{n^m}\Big)^{\frac{1}{\gamma'}} \leq M\sum_{n\geq 1} n^{-(\frac{m}{2\gamma'}-p)} < \infty.$$

This ends the proof. ■

**Remark 1.** A sufficient condition for $\int_0^\infty s^p h^{\beta/\gamma'}\,\mathrm{d}s < \infty$ with $\gamma' > 1$ is $\int_0^\infty s^{(p+1)\gamma'+\delta''} h^\beta\,\mathrm{d}s < \infty$ for some $\delta'' > 0$.

Next we get asymptotic behavior for the joint distribution of $(X_1, X_j)$.

**Lemma 3.1.** *If $\int_0^\infty h^{(\alpha\wedge 1)-\delta}(s)\,\mathrm{d}s < \infty$ for some $\delta > 0$, then for any $y > 0$, $j \geq 1$*

$$\lim_{x\to\infty} \frac{P\{X_1 > x, X_j > xy\}}{P\{A_1 > x\}} = \sum_{i\leq 1} E(C_{1i} \wedge C_{ji}y^{-1})^\alpha.$$

**Proof:** Recall $C_{ji} = h(\tau_j - \tau_i)$. Conditional on $\{C_{ji}\}$, we obtain (see the proofs of Lemma 4.24 in Resnick [58] and Theorem 2.1 in Resnick and Willekens [63]) for any $x > x_0(\delta)$

$$\begin{aligned}\frac{P\{X_1 > x, X_j > xy|\{C_{ji}\}\}}{P\{A_1 > x\}} &\leq \frac{P\{X_1 > x|\{C_{ji}\}\}}{P\{A_1 > x\}} \\ &\leq M\sum_{i\leq 1} C_{1i}^{\alpha-\delta} + MI_{\{\alpha\geq 1\}}\Big(\sum_{i\leq 1} C_{1i}^{\frac{\alpha-\delta^2}{\alpha+\delta}}\Big)^{\alpha+\delta}. \qquad (4.3)\end{aligned}$$

Using Minkowski's inequality we have

$$E\Big(\sum_{i\leq 1} C_{1i}^{\frac{\alpha-\delta^2}{\alpha+\delta}}\Big)^{\alpha+\delta} \leq \Big(\sum_{i\leq 1}\big(EC_{1i}^{\alpha-\delta^2}\big)^{\frac{1}{\alpha+\delta}}\Big)^{\alpha+\delta}.$$

Note that $\frac{\alpha-\delta^2}{\alpha+\delta} \geq 1-\delta$ if $\alpha \geq 1$. Applying Proposition 4.3 with $(p, \beta, \gamma') = (0, \alpha-\delta, 1)$ and $(p, \beta, \gamma') = (0, \alpha-\delta^2, \alpha+\delta)$, we have that the expression on the right hand side of (4.3) is finite. Hence the desired result follows from Lemma 5.1 in Datta and McCormick (1998) and dominated convergence theorem. ■

**Remark 2.** By taking $j = 1$, $y = 1$, Lemma 4.3 implies that $\overline{F}_{X_1}$ is regularly varying with index $-\alpha$.

Recall that $b(t) = F_{X_1}^{\leftarrow}(1 - \frac{1}{t})$ and usual lattice notation $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}$. Next we will deduce an upper bound for $\frac{n}{k}P\left\{X_1 > b\left(\frac{n}{k}\right)x, X_j > b\left(\frac{n}{k}\right)y\right\}$.

**Proposition 3.2.** *If $\int_0^\infty s^{2+\delta}h^{(\alpha\wedge 1)-\delta}(s)\,\mathrm{d}s < \infty$ for some $\delta > 0$, then for large $n$, any $j > 1$, $x, y \geq \delta_0 > 0$, and $k = o(n)$*

$$\begin{aligned}&\frac{n}{k}P\Big\{X_1 > b\left(\frac{n}{k}\right)x, X_j > b\left(\frac{n}{k}\right)y\Big\} \\ \leq\ & \frac{k}{n}\frac{M}{(xy)^{\alpha-\delta}} + M\sum_{i\leq 1} E\Big(\frac{C_{1i}}{x} \wedge \frac{C_{ji}}{y}\Big)^{\alpha-\delta} + \frac{M}{(xy)^{\frac{\alpha-\delta}{2}}} j^{-\left(1+\frac{\delta}{3}\right)}.\end{aligned}$$

**Proof:** For any $j > 1$

$$\begin{aligned}
&\frac{n}{k}P\Big\{X_1 > b\left(\frac{n}{k}\right)x, X_j > b\left(\frac{n}{k}\right)y\Big\}\\
\leq\ & \frac{n}{k}P\Big\{X_1 > b\left(\frac{n}{k}\right)x\Big\}P\Big\{\sum_{1<i\leq j} A_iC_{ji} > b\left(\frac{n}{k}\right)y/2\Big\}\\
&+\frac{n}{k}P\Big\{X_1 > b\left(\frac{n}{k}\right)x, \sum_{i\leq 1} A_iC_{ji} > b\left(\frac{n}{k}\right)y/2\Big\}\\
\leq\ & \frac{n}{k}P\Big\{X_1 > b\left(\frac{n}{k}\right)x\Big\}P\Big\{X_j > b\left(\frac{n}{k}\right)y/2\Big\}\\
&+\frac{n}{k}P\Big\{\bigvee_{i\leq 1} C_{1i}A_i > b\left(\frac{n}{k}\right)x \text{ or } \bigvee_{i\leq 1} C_{ji}A_i > b\left(\frac{n}{k}\right)y/2,\\
&\qquad X_1 > b\left(\frac{n}{k}\right)x, \sum_{i\leq 1} A_iC_{ji} > b\left(\frac{n}{k}\right)y/2\Big\}\\
&+\frac{n}{k}P\Big\{\sum_{i\leq 1} C_{1i}A_i\wedge b\left(\frac{n}{k}\right)x > b\left(\frac{n}{k}\right)x, \sum_{i\leq 1} C_{ji}A_i\wedge (b\left(\frac{n}{k}\right)y/2) > b\left(\frac{n}{k}\right)y/2\Big\}\\
=\ & I_{j,1}+I_{j,2}+I_{j,3}.
\end{aligned}$$

By Potter's bound (Bingham et al. (19890 Theorem 1.5.6), for any $0<\delta<\alpha$ and large $n$, we can bound $I_{j,1}$ by $\frac{M}{(xy)^{\alpha-\delta}}\frac{k}{n}$. It is easy to show that $I_{j,2}$ is at most It is easy to show that $I_{j,2}$ is at most

$$\frac{2n}{k}\sum_{i\leq 1} P\Big\{C_{1i}A_i > b\left(\frac{n}{k}\right)x/2, C_{ji}A_i > b\left(\frac{n}{k}\right)y/4\Big\} \tag{4.4}$$

$$+\frac{n}{k}\sum_{i\leq 1} P\Big\{C_{1i}A_i > b\left(\frac{n}{k}\right)x, \sum_{m\leq 1, m\neq i} A_mC_{1m} > b\left(\frac{n}{k}\right)y/4\Big\} \tag{4.5}$$

$$+\frac{n}{k}\sum_{i\leq 1} P\Big\{\sum_{m\leq 1, m\neq i} C_{1m}A_m > b\left(\frac{n}{k}\right)x/2, C_{1i}A_i > b\left(\frac{n}{k}\right)y/2\Big\}. \tag{4.6}$$

Applying Potter's bound again, the term (4.4) is less than $M\sum_{i\leq 1} E\Big(\frac{C_{1i}}{x}\wedge\frac{C_{ji}}{y}\Big)^{\alpha-\delta}$. Conditional on $\{C_{ji}\}$, using Potter's bound and inequality (4.3), we know that the term (4.5) is at most

$$\frac{M}{(xy)^{\alpha-\delta}}\frac{k}{n}E\sum_{i\leq 1} C_{1i}^{\alpha-\delta}\left(\sum_{i\leq 1} C_{1i}^{\alpha-\delta} + MI_{\{\alpha\geq 1\}}\Big(\sum_{i\leq 1} C_{1i}^{\frac{\alpha-\delta^2}{\alpha+\delta}}\Big)^{\alpha+\delta}\right).$$

Minkowski's inequality yields $E(\sum_{i\leq 1} C_{1i}^{\alpha-\delta})^2 \leq \Big(\sum_{i\leq 1}\big(E{C_{1i}}^{2(\alpha-\delta)}\big)^{\frac{1}{2}}\Big)^2$, which is finite by Proposition 1. If $\alpha\geq 1$, by Proposition 1, Hölder's and Minkowski's inequalities yield for any conjugate exponents $p$ and $q$ greater than 1

$$
\begin{aligned}
E\sum_{i\le 1} C_{1i}^{\alpha-\delta}\Big(\sum_{i\le 1} C_{1i}^{\frac{\alpha-\delta_1}{\alpha+\delta_2}}\Big)^{\alpha+\delta_2} &\le \Big(E\big(\sum_{i\le 1} C_{1i}^{\alpha-\delta}\big)^p\Big)^{\frac{1}{p}}\Big(E\Big(\sum_{i\le 1} C_{1i}^{\frac{\alpha-\delta}{\alpha+\delta}}\Big)^{q(\alpha+\delta)}\Big)^{\frac{1}{q}} \\
&\le \Big(\sum_{i\le 1}\big(EC_{1i}{}^{p(\alpha-\delta)}\big)^{\frac{1}{p}}\Big)\Big(\sum_{i\le 1}\big(EC_{1i}^{q(\alpha-\delta)}\big)^{\frac{1}{q(\alpha+\delta)}}\Big)^{\alpha+\delta} \\
&< \infty.
\end{aligned}
$$

Therefore

$$
I_{j2} \;\le\; M\sum_{i\le 1} E\Big(\frac{C_{1i}}{x}\wedge\frac{C_{ji}}{y}\Big)^{\alpha-\delta} + \frac{k}{n}\frac{M}{(xy)^{\alpha-\delta}}.
$$

$$
\begin{aligned}
E\sum_{i\le 1} C_{1i}^{\alpha-\delta}\Big(\sum_{i\le 1} C_{1i}^{\frac{\alpha-\delta^2}{\alpha+\delta}}\Big)^{\alpha+\delta} &\le \Big(E\big(\sum_{i\le 1} C_{1i}^{\alpha-\delta}\big)^p\Big)^{\frac{1}{p}}\Big(E\Big(\sum_{i\le 1} C_{1i}^{\frac{\alpha-\delta^2}{\alpha+\delta}}\Big)^{q(\alpha+\delta)}\Big)^{\frac{1}{q}} \\
&\le \Big(\sum_{i\le 1}\big(EC_{1i}{}^{p(\alpha-\delta)}\big)^{\frac{1}{p}}\Big)\Big(\sum_{i\le 1}\big(EC_{1i}^{q(\alpha-\delta^2)}\big)^{\frac{1}{q(\alpha+\delta)}}\Big)^{\alpha+\delta} \\
&< \infty.
\end{aligned}
$$

Therefore

$$
I_{j,2} \;\le\; M\sum_{i\le 1} E\Big(\frac{C_{1i}}{x}\wedge\frac{C_{ji}}{y}\Big)^{\alpha-\delta} + \frac{k}{n}\frac{M}{(xy)^{\alpha-\delta}}.
$$

From Markov's and Hölder's inequalities we have for any $a, b > 1$

$$
P\{|X|>a, |Y|>b\} \le E\left(\frac{|XY|}{ab}\right)^{\frac{\alpha+\delta}{2}} \le \left(E\left(\frac{|X|}{a}\right)^{\alpha+\delta}\right)^{\frac{1}{2}}\left(E\left(\frac{|Y|}{b}\right)^{\alpha+\delta}\right)^{\frac{1}{2}}.
$$

Thus the last term $I_{j,3}$ is bounded by

$$
\begin{aligned}
&M\frac{n}{k}\Big(\left(b\left(\frac{n}{k}\right)x\right)^{-(\alpha+\delta)} E\Big(\sum_{i\le 1} C_{1i}A_i\wedge b\left(\frac{n}{k}\right)x\Big)^{\alpha+\delta}\Big)^{\frac{1}{2}} \\
&\qquad\times\Big(\left(b\left(\frac{n}{k}\right)y\right)^{-(\alpha+\delta)} E\Big(\sum_{i\le 1} C_{ji}A_i\wedge b\left(\frac{n}{k}\right)y/2\Big)^{\alpha+\delta}\Big)^{\frac{1}{2}}.
\end{aligned}
$$

From the proof of Theorem 2.1 in Resnick and Willekens (1991), we know for large $n$ and $t \ge \frac{\delta_0}{2}$

$$
\begin{aligned}
&\frac{n}{k}\left(b\left(\frac{n}{k}\right)t\right)^{-(\alpha+\delta)} E\Big(\sum_{i\le 1} C_{ji}A_i\wedge b\left(\frac{n}{k}\right)t\Big)^{\alpha+\delta} \\
\le\;& MI_{\{\alpha<1\}}t^{-(\alpha-\delta)}\sum_{i\le 1} EC_{ji}{}^{\alpha-\delta} + MI_{\{\alpha\ge 1\}}t^{-(\alpha-\delta)}\Big(\sum_{i\le 1}\big(EC_{ji}^{\alpha-\delta^2}\big)^{\frac{1}{\alpha+\delta}}\Big)^{\alpha+\delta}.
\end{aligned}
$$

The assumption $\int_0^\infty s^{2+\delta} h^{(\alpha\wedge 1)-\delta}(s)\,\mathrm{d}s < \infty$ for some $\delta > 0$ yields

$$\sum_{i\le 1} E{C_{ji}}^{\alpha-\delta} \le j^{-2\left(1+\frac{\delta}{3}\right)} \sum_{i\le 1} (j-i+1)^{2+\delta} E{C_{ji}}^{\alpha-\delta} \le M j^{-2\left(1+\frac{\delta}{3}\right)}$$

and, if $\alpha \ge 1$,

$$\begin{aligned} \sum_{i\le 1} \left(EC_{ji}^{\alpha-\delta^2}\right)^{\frac{1}{\alpha+\delta}} &\le j^{-\frac{2\left(1+\frac{\delta}{3}\right)}{\alpha+\delta}} \sum_{i\le 1} (j-i+1)^{\frac{2\left(1+\frac{\delta}{3}\right)}{\alpha+\delta}} \left(E{C_{ji}}^{\alpha-\delta^2}\right)^{\frac{1}{\alpha+\delta}} \\ &\le M j^{-\frac{2\left(1+\frac{\delta}{3}\right)}{\alpha+\delta}}. \end{aligned}$$

Therefore

$$I_{j,3} \le \frac{M}{(xy)^{\frac{\alpha-\delta}{2}}} j^{-\left(1+\frac{\delta}{3}\right)}.$$

This completes the proof of the proposition. ■

Define $c_0 = \sum_{i\ge 0} Eh^\alpha(\tau_i)$, $\lambda^{(1)} = \frac{2}{\alpha^2}\Big(1 + c_0^{-1}\sum_{j\ge 2}\sum_{i\le 1} EC_{ji}^\alpha\big(2+\alpha\log\frac{C_{1i}}{C_{ji}}\big)\Big)$, $\lambda^{(2)} = 1 + 2c_0^{-1}\sum_{j\ge 2}\sum_{i\le 1} EC_{ji}^\alpha$, and $\lambda = \frac{1}{\alpha^2}\Big(1 + 2c_0^{-1}\sum_{j\ge 2}\sum_{i\le 1} EC_{ji}^\alpha\Big)$. The proof of following lemma should be compared with that of Remark 2.1 in Resnick and Stărică (1997).

**Lemma 3.2.** *Let $a^+ = a \vee 0$ and $r_n$ be a sequence of positive integers such that $r_n = o(\frac{n}{k})$. If $\int_0^\infty s^{2+\delta} h^{(\alpha\wedge 1)-\delta}(s)\,\mathrm{d}s < \infty$ for some $\delta > 0$, then as $n \to \infty$,*

$$\begin{aligned} &\frac{n}{kr_n} Var\Big(\sum_{1\le j\le r_n} \Big(\log\frac{X_j}{b\left(\frac{n}{k}\right)}\Big)^+\Big) \to \lambda^{(1)}, \\ &\frac{n}{kr_n} Var\Big(\sum_{1\le j\le r_n} I_{\{X_j/b\left(\frac{n}{k}\right) > x\}}\Big) \to \lambda^{(2)} x^{-\alpha}, \\ &\frac{n}{kr_n} Var\Big(\sum_{1\le j\le r_n} \Big(\log\frac{X_j}{b\left(\frac{n}{k}\right)}\Big)^+ - \frac{1}{\alpha} I_{\{X_j/b\left(\frac{n}{k}\right) > 1\}}\Big) \to \lambda. \end{aligned}$$

**Proof:** It is easy to show that (see for example Hsing (1991) for $m \ge 1$

$$\frac{n}{k} E\Big(\Big(\log\frac{X_1}{b\left(\frac{n}{k}\right)}\Big)^+\Big)^m \sim \frac{m!}{\alpha^m}.$$

For any $\beta > 0$ straightforward calculation leads to

$$E\int_1^\infty\int_1^\infty \left(\frac{C_{1i}}{x} \wedge \frac{C_{ji}}{y}\right)^\beta \frac{dx}{x}\frac{dy}{y} = \frac{1}{\beta^2} EC_{ji}^\beta \left(2 + \beta\log\frac{C_{1i}}{C_{ji}}\right).$$

Since $\log x \leq \frac{x^{\alpha-\delta}}{\alpha-\delta}$ for $x \geq 1$, we have $C_{ji}^{\alpha-\delta} \log \frac{C_{1i}}{C_{ji}} \leq \frac{1}{\alpha-\delta} C_{1i}^{\alpha-\delta}$. Therefore we conclude that

$\sum_{j\geq 1} \sum_{i\leq 1} E \int_1^\infty \int_1^\infty \left(\frac{C_{1i}}{x} \wedge \frac{C_{ji}}{y}\right)^{\alpha-\delta} \frac{dx}{x}\frac{dy}{y}$ is bounded by $M \sum_{n\geq 1}(n+1)Eh^{\alpha-\delta}(\tau_n)$, which is finite by Proposition 4.3 provided $\int_0^\infty sh^{\alpha-\delta}(s)\,\mathrm{d}s < \infty$. Now the rest of the proof follows by the arguments given in Resnick and Stărică (1997) (see Remark 2.1 in (1997) ). ■

Proof of the asymptotic normality of the Hill's estimator for stationary dependent sequences is greatly eased by the work of Rootzen et al. (1990), wherein sufficient conditions are presented for such an asymptotic result. This was the approach taken in Resnick and Stărică (1997) and is used here as well. The methodology utilized in Rootzen et al. (1990) is based on a blocking technique to handle dependency. At this time, it will be convenient to recount the blocking and other notation from Rootzen et al. (1990), which we need for our proof. Firstly, we let $\{k_n^*, n \geq 1\}$ denote a sequence of positive integers diverging to infinity. $k_n^*$ will represent the number of big blocks. Recall that we have used $k = k_n$ to indicate the number of upper order statistics needed to define Hill's estimator. Our choice in this notation stems from our desire to maintain a close notational affinity with the papers, Rootzen et al. [66] and Resnick and Stărică (1997). The choice of $k$ for the Hill's estimator parameter is that taken in Resnick and Stărică (1997), whereas the same quantity in Rootzen et al. (1990) is denoted by $c_n$ while $k$ is used to denote number of big blocks. To make comparisons to these two papers and ours easier, we choose to reserve the quantity $k$ as the Hill's parameter as in Resnick and Stărică (1997) and to use $k^*$ as the big block multiplicity. We also choose to use $c_n$ interchangeably with $k$ in order to maintain a close notational connection with Rootzen et al. (1990). More precisely, we set $c_n = k + 1$.

Let $r_n = [n/k_n^*]$, the integer part of $n/k_n^*$, denote big block size and let $l_n = o(r_n)$ denote the gap size between blocks. Further let $J_i = \{(i-1)r_n + 1, \cdots, ir_n\}$, and $J_i^*$ denote the first $r_n - l_n$ integers in $J_i$, i.e. the $i$th big block, $1 \leq i \leq k_n^*$. Define a function $\psi(x) =$

$xI_{\{x\geq 0\}} - \frac{1}{\alpha} I_{\{x>0\}}$ and

$$\lambda_n = \tfrac{k_n^*}{c_n} \mathrm{Var}\Big( \textstyle\sum_{j=1}^{r_n} \psi\big((X_j^* - u_n)^+\big)\Big),$$
$$U_i = U_{n,i} = (\lambda_n c_n)^{-\frac{1}{2}} \textstyle\sum_{j \in J_i^*} \psi\big((X_j^* - u_n)^+\big),$$
$$V_i = V_{n,i} = (\lambda_n c_n)^{-\frac{1}{2}} \textstyle\sum_{j \in J_i \setminus J_i^*} \psi\big((X_j^* - u_n)^+\big),$$

where $X_j^* = \log(X_j)$, $u_n = \log(F_{X_1}^{\leftarrow}(1 - \frac{k}{n}))$.

For our proof of asymptotic normality, we require some growth conditions on the Hill's estimator parameter $k$, namely, $r_n \sim \frac{k^{\frac{1}{2}} \wedge \frac{n}{k}}{\log^2 k}$, $l_n \sim \frac{r_n}{\log k}$. By Lemma 4.3, $\lambda_n$ converges to $\lambda$ where $\lambda$ is defined just prior to the statement of Lemma 2.

**Proof of Theorem 2.2.** We begin by recounting the conditions under which asymptotic normality of Hill's estimator for a stationary sequence holds according to Theorem 4.3 in Rootzen et al. (1990). According to that result in order that for a stationary sequence $\{X_n, n \geq 1\}$ with marginal distribution $F$ having tail $\overline{F} = 1 - F$ regularly varying with index $-\alpha$ to be such that

$$\sqrt{c_n - 1}\left( \frac{1}{c_n - 1} \sum_{i=1}^{c_n} \log \frac{X_{(i)}}{X_{(c_n)}} - \frac{n}{c_n} E\left(\log X_1 - u_n\right)_+ \right) \Rightarrow N(0, \lambda)$$

for an appropriate sequence of positive integers $c_n$ and positive constant $\lambda$, the following conditions suffice:

(i) $\{X_n\}$ is a stationary strong mixing sequence with strong mixing coefficient

$$\alpha_{n,l} = \sup\{|P(A \cap B) - P(A)P(B)| : A \in \mathcal{B}_{1,k}, B \in \mathcal{B}_{k+l,n}, 1 \leq k \leq n - l\},$$

where $\mathcal{B}_{i,j} = \sigma(X_k : i \leq k \leq j)$ and $\alpha_{n,l}$ is such that for some sequence $l_n = o(n)$, $\alpha_{n,l_n} = o(1)$; further, integers $k_n^*$ diverging to infinity are chosen such that $k_n^*(\alpha_{n,l_n} + l_n/n) = o(1)$ and set $r_n = [n/k_n^*]$;

(ii) integers $c_n$ tending to infinity and reals $u_n$ are chosen so that $\overline{F}(e^{u_n}) \sim c_n/n$ and, moreover, $c_n = o(k_n^*)$.

Defining functions $\psi_1(x) = xI_{\{x\geq 0\}}$; $\psi_2(x) = I_{\{x>0\}}$; $\psi(x) = \psi_1(x) - \frac{1}{\alpha}\psi_2(x)$ and quantities

$$\lambda_n^{(i)} = \frac{k_n^*}{c_n}\mathrm{Var}\left(\sum_{j=1}^{r_n}\psi_i\left((\log X_j - u_n)_+\right)\right),\ i=1,2 \quad \text{and} \quad \lambda_n = \frac{k_n^*}{c_n}\mathrm{Var}\left(\sum_{j=1}^{r_n}\psi\left((\log X_j - u_n)_+\right)\right),$$

the list of sufficient conditions continues with:

(iii) $\lambda_n^{(i)} = O(1)$, $i = 1, 2$ and $\lambda_n \to \lambda > 0$;

(iv) the marginal distribution $F$ of the stationary sequence satisfies the Von Mises condition $tF'/\overline{F}(t) \to \alpha$ and finally

(v) for any sequence of reals $z_n$ with $\sqrt{c_n/\lambda_n}(z_n - u_n) = O(1)$, we have that

$$\frac{k_n^*}{\lambda_n c_n}\mathrm{Var}\left(\sum_{i=1}^{r_n} I_{\{u_n\wedge z_n \leq \log X_i \leq u_n \vee z_n\}}\right) = o(1).$$

We can now explain our strategy to establish asymptotic normality. Each of the conditions (ii)-(v) above are easily verified. With regard to condition (i), strong mixing, we take a different approach. A close examination of the proof of asymptotic normality in Rootzen et al. (1990) indicates where exactly the strong mixing condition is used and offers the opportunity to establish that step without that condition. This is the program which we will embark upon, thereby obtaining the conclusion of Theorem 4.3 in Rootzen et al. (1990) without the assumption (i) of strong mixing.

Before entering into the details of our proof, we settle one issue regarding centering. By Hsing (1991), we have that under the assumption of second order regular variation on the marginal distribution tail $\overline{F}$ with auxiliary function $g$ and $b(t)$ denoting $F^{\leftarrow}(1-1/t)$

$$\sqrt{k}\left(\frac{n}{k}E(\log X_1 - u_n)_+ - \frac{1}{\alpha}\right) = \sqrt{k}O\left(g(b(\frac{n}{k}))\right).$$

By this relation, under an assumption of second order regular variation on the marginal tail distribution, if the number of upper order statistics used in the definition of the Hill's estimator $k$ is controlled so that $\sqrt{k}g(b(\frac{n}{k})) = o(1)$, then the Hill's estimator may be centered at the reciprocal of the tail index as we have written in our Theorem 2.

Condition (ii) holds by definition of $u_n$ and since by assumption $rk \le 2n/\log^2 k$ for large $n$, whereas $rk^* \sim n$ where recall $c_n = k+1$. (iii) holds by virtue of Lemma 2. (iv) is true by assumption and (v) may be derived using Lemmas 1 and 2. Thus, by our earlier remark concerning centering by $1/\alpha$, Theorem 2 follows from Theorem 4.3 in Rootzen et al. (1990) once we address the mixing question to which problem a careful examination of their paper reveals that it suffices to show

$$\left| E\exp(it\sum_{j=1}^{k_n^*}(U_j - EU_j)) - \prod_{j=1}^{k_n^*} E\exp(it(U_j - EU_j))\right| \to 0 \tag{4.7}$$

and

$$\left| E\exp(it\sum_{j=1}^{k_n^*}(V_j - EV_j)) - \prod_{j=1}^{k_n^*} E\exp(it(V_j - EV_j))\right| \to 0. \tag{4.8}$$

Define $X_j^{(m)} = \sum_{j-m<i\le j} h(\tau_j - \tau_i)A_i$, for $j \ge 1$ and $m \ge 1$. Clearly the sequence $\{X_j^{(m)}, j \ge 1\}$ is stationary and $m$-dependent for $m \ge 1$. Let

$$U_i^* = (\lambda_n c_n)^{-\frac{1}{2}} \sum_{j \in J_i^*} \psi\left(\left(\log\left(X_j^{(j-(i-1)r_n+l_n)}/b\left(\frac{n}{k}\right)\right)\right)^+\right),$$

$$V_i^* = (\lambda_n c_n)^{-\frac{1}{2}} \sum_{j \in J_i \setminus J_i^*} \psi\left(\left(\log\left(\log(X_j^{(j-(i-1)r_n)}/b\left(\frac{n}{k}\right)\right)\right)^+\right),\ 1 \le i \le k_n.$$

It can be checked that $\{V_i^*\}$, $\{U_i^*\}$ are independent sequences of i.i.d. r.v.'s. Since $|\prod a_i - \prod b_i| \le \sum |a_i - b_i|$ for any complex numbers $a_i, b_i$s satisfying $|a_i| \le 1, |b_i| \le 1$, and $|e^{ix} - e^{iy}| \le$

$\left|\sin(\frac{x-y}{2})\right| \leq \frac{|x-y|}{2}$, we obtain that

$$
\begin{aligned}
& \Big| E\exp(it\sum_{j=1}^{k_n^*}(U_j - EU_j)) - \prod_{j=1}^{k_n^*} E\exp(it(U_j - EU_j))\Big| \\
\leq \; & \Big| E\exp(it\sum_{j=1}^{k_n^*} U_j) - E\exp(it\sum_{j=1}^{k_n^*} U_j^*)\Big| + \Big|\prod_{j=1}^{k_n^*} E\exp(itU_j^*) - \prod_{j=1}^{k_n^*} E\exp(itU_j)\Big| \\
\leq \; & 2|t|k_n^* E|U_1 - U_1^*| \\
\leq \; & Mk^{-\frac{1}{2}}k_n^* \sum_{j\in J_1^*} E\left(\log\left(X_j/b\left(\frac{n}{k}\right)\right)\right)^+ I_{\{X_j^{(j+l_n)} \leq b\left(\frac{n}{k}\right)/2\}} \\
& + Mk^{-\frac{1}{2}}k_n^* \sum_{j\in J_1^*} E\log\left(X_j/X_j^{(j+l_n)}\right) I_{\{X_j^{(j+l_n)} > b\left(\frac{n}{k}\right)/2\}} \\
& + Mk^{-\frac{1}{2}}k_n^* \sum_{j\in J_1^*} E\Big(I_{\{X_j > b\left(\frac{n}{k}\right)\}} - I_{\{X_j^{(j+l_n)} > b\left(\frac{n}{k}\right)\}}\Big) \\
= \; & I_a + I_b + I_c.
\end{aligned}
$$

We will examine each term in turn. Note that for any $j \in J_1^*$

$$
\begin{aligned}
& E\left(\log\left(X_j/b\left(\frac{n}{k}\right)\right)\right)^+ I_{\{X_j^{(j+l_n)} \leq b\left(\frac{n}{k}\right)/2\}} \\
= \; & \int_0^\infty P\Big\{\left(\log\left(X_j/b\left(\frac{n}{k}\right)\right)\right)^+ I_{\{X_j^{(j+l_n)} \leq b\left(\frac{n}{k}\right)/2\}} > u\Big\}\,\mathrm{d}u \\
\leq \; & \int_0^\infty P\Big\{X_j - X_j^{(j+l_n)} > (e^u - 1/2)b\left(\frac{n}{k}\right)\Big\}\,\mathrm{d}u \\
\leq \; & \int_0^\infty P\Big\{\sum_{i\geq l_n} h(\tau_i)A_i > (e^u - 1/2)b\left(\frac{n}{k}\right)\Big\}\,\mathrm{d}u \\
\leq \; & MP\Big\{\sum_{i\geq l_n} h(\tau_i)A_i > b\left(\frac{n}{k}\right)\Big\},
\end{aligned}
$$

where we used the inequality (4.3) in the last step. From inequality (4.3) and Minkowski's inequality we have that $I_a$ is bounded by

$$
Mk^{\frac{1}{2}}l_n^{-\gamma}\Big(\sum_{i\geq l_n} i^\gamma Eh^{\alpha-\delta}(\tau_i) + I_{\{\alpha\geq 1\}}\Big(\sum_{i\geq l_n} i^{\frac{\gamma}{\alpha+\delta}}\left(Eh^{\alpha-\delta^2}(\tau_i)\right)^{\frac{1}{\alpha+\delta}}\Big)^{\alpha+\delta}\Big).
$$

Since $k^{\frac{1}{2}}l_n^{-\gamma} = o(1)$ and $\int_0^\infty s^\gamma h^{(\alpha\wedge 1)-\delta}(s)\,\mathrm{d}s < \infty$, using Proposition 4.3, we have

$$
I_a \to 0.
$$

For term $I_b$, because on the set $\{X_j^{(j+l_n)} > b\left(\frac{n}{k}\right)/2\}$

$$
\log(X_j/X_j^{(j+l_n)}) \leq \log\Big(1 + (X_j - X_j^{(j+l_n)})/\big(b\left(\frac{n}{k}\right)/2\big)\Big)
$$

we get

$$
\begin{aligned}
I_b &\leq Mk^{-\frac{1}{2}}k_n^* \sum_{j\in J_1^*} \int_0^\infty P\Big\{X_j - X_j^{(j+l_n)} > (e^u-1)b\left(\frac{n}{k}\right)/2, X_j^{(j+l_n)} > b\left(\frac{n}{k}\right)/2\Big\}du \\
&\leq Mk^{-\frac{1}{2}}k_n^* \sum_{j\in J_1^*} \Big\{ \int_0^\infty P\Big\{u \leq \log(1+2h(-\tau_{-l_n})), X_j^{(j+l_n)} > b\left(\frac{n}{k}\right)/2\Big\}\,\mathrm{d}u \\
&\quad + \int_0^\infty P\Big\{X_j - X_j^{(j+l_n)} > (e^u-1)b\left(\frac{n}{k}\right)/2, (e^u-1)/2 > h(-\tau_{-l_n})\Big\}\,\mathrm{d}u\Big\} \\
&\leq Mk^{-\frac{1}{2}}n \int_0^\infty P\Big\{u \leq 2h(-\tau_{-l_n}), X_{r_n-l_n}^{(r_n)} > 2^{-1}b\left(\frac{n}{k}\right)\Big\}\,\mathrm{d}u \\
&\quad + Mk^{-\frac{1}{2}}n \int_0^\infty P\Big\{X_1 - X_1^{(1+l_n)} > (e^u-1)b\left(\frac{n}{k}\right)/2, (e^u-1)/2 > h(-\tau_{-l_n})\Big\}\,\mathrm{d}u \\
&= I_{b,1} + I_{b,2}.
\end{aligned}
$$

By conditioning on $\{C_{ji}\}$ and using inequality (4.3) we obtain

$$
\begin{aligned}
I_{b,1} &\leq Mk^{\frac{1}{2}}\Big\{Eh(-\tau_{-l_n})\sum_{i=-l_n+1}^{r_n-l_n} h^{\alpha-\delta}(\tau_{r_n-l_n}-\tau_i) \\
&\qquad + I_{\{\alpha\geq 1\}}Eh(-\tau_{-l_n})\Big(\sum_{i=-l_n+1}^{r_n-l_n} h^{\frac{\alpha-\delta^2}{\alpha+\delta}}(\tau_{r_n-l_n}-\tau_i)\Big)^{\alpha+\delta}\Big\}.
\end{aligned}
$$

By Hölder's inequality and Minkowski's inequality this is at most

$$
\begin{aligned}
&Mk^{\frac{1}{2}}\Big\{\big(Eh^p(\tau_{l_n})\big)^{\frac{1}{p}}\sum_{i\geq 0}\big(Eh^{q(\alpha-\delta)}(\tau_i)\big)^{\frac{1}{q}} \\
&\qquad + I_{\{\alpha\geq 1\}}\big(Eh^p(\tau_{l_n})\big)^{\frac{1}{p}}\Big(\sum_{i\geq 0}\big(Eh^{q(\alpha-\delta^2)}(\tau_i)\big)^{\frac{1}{q(\alpha+\delta)}}\Big)^{\alpha+\delta}\Big\},
\end{aligned}
$$

where $p, q > 1$, and $\frac{1}{p}+\frac{1}{q}=1$. By Proposition 4.3 we have that

$$
I_{b,1} \leq Mk^{\frac{1}{2}}\big(Eh^p(\tau_{l_n})\big)^{\frac{1}{p}}.
$$

Without loss of generality we may assume $0 \leq \tau_{i+1}-\tau_i \leq 1$ for any integer $i$, and $\mu = E\tau_1 > 0$. Clearly

$$
k^{\frac{1}{2}}\big(Eh^p(\tau_{l_n})\big)^{\frac{1}{p}} \leq h(0)k^{\frac{1}{2}}\Big(P\Big\{|\tau_{l_n}-\mu l_n| > \frac{\mu l_n}{2}\Big\}\Big)^{\frac{1}{p}} + k^{\frac{1}{2}}h\Big(\frac{\mu l_n}{2}\Big).
$$

By the Marcinkiewicz-Zygmund's inequality we have that the first term in right hand side of above inequality converges to 0 if $k^{\frac{1}{2}} = O(l_n^\beta)$ for some $\beta > 0$. Since $h$ is a nonincreasing function, assumption (4.2) yields $h(s) = o(s^{-\gamma})$. Since $k^{\frac{1}{2}} l_n^{-\gamma} = o(1)$, we conclude

$$I_{b,1} \to 0.$$

On the set $\{(e^u - 1)/2 > h(-\tau_{-l_n})\}$, $\frac{h(\tau_1-\tau_i)}{(e^u-1)/2} < 1$ for any $i \le -l_n$. Inequality (4.3) yields

$$I_{b,2} \le Mk^{\frac{1}{2}} E \int_0^\infty \Big\{ \sum_{i \le -l_n} \Big(\frac{h(\tau_1 - \tau_i)}{(e^u - 1)/2}\Big)^{\alpha-\delta} + I_{\{\alpha \ge 1\}} \Big( \sum_{i \le -l_n} \Big(\frac{h(\tau_1 - \tau_i)}{(e^u - 1)/2}\Big)^{\frac{1-\delta}{\alpha+\delta}} \Big)^{\alpha+\delta} \Big\} \,\mathrm{d}u.$$

Because for any $\delta < \beta \le 1$

$$\int_0^\infty \left(\frac{e^u - 1}{2}\right)^{-(\beta-\delta)} \mathrm{d}u = \int_0^\infty \frac{2}{1+2t} t^{-(\beta-\delta)} \,\mathrm{d}t < \infty, \ u = \log(1 + 2t),$$

we have that

$$\begin{aligned} I_{b,2} &\le Mk^{\frac{1}{2}} \Big( E \sum_{i \ge l_n} h^{\alpha-\delta}(\tau_i) + I_{\{\alpha \ge 1\}} \Big( \sum_{i \ge l_n} \big(Eh^{1-\delta}(\tau_i)\big)^{\frac{1}{\alpha+\delta}} \Big)^{\alpha+\delta} \Big) \\ &\le Mk^{\frac{1}{2}} l_n^{-\gamma} \Big( \sum_{i \ge l_n} i^\gamma E h^{\alpha-\delta}(\tau_i) + I_{\{\alpha \ge 1\}} \Big( \sum_{i \ge l_n} i^{\frac{\gamma}{\alpha+\delta}} \big(Eh^{1-\delta}(\tau_i)\big)^{\frac{1}{\alpha+\delta}} \Big)^{\alpha+\delta} \Big) \\ &\to 0. \end{aligned}$$

For last item

$$\begin{aligned} I_c &= Mk^{-\frac{1}{2}} k_n^* \sum_{j \in J_1^*} P\Big\{ X_j > b\left(\frac{n}{k}\right), X_j^{(j+l_n)} \le b\left(\frac{n}{k}\right) \Big\} \\ &\le Mk^{-\frac{1}{2}} k_n^* \sum_{j \in J_1^*} \Big( P\Big\{ b\left(\frac{n}{k}\right) < X_j \le (1+\epsilon_n) b\left(\frac{n}{k}\right) \Big\} \\ &\qquad\qquad + P\Big\{ X_j > (1+\epsilon_n) b\left(\frac{n}{k}\right), X_j^{(j+l_n)} \le b\left(\frac{n}{k}\right) \Big\} \Big) \\ &\le Mk^{\frac{1}{2}} \left( 1 - \frac{P\{X_1 > (1+\epsilon_n) b\left(\frac{n}{k}\right)\}}{P\{X_1 > b\left(\frac{n}{k}\right)\}} \right) \\ &\quad + M\, k^{-\frac{1}{2}} n P\Big\{ X_1 - X_1^{(1+l_n)} > \epsilon_n b\left(\frac{n}{k}\right) \Big\} \\ &= I_{c,1} + I_{c,2}. \end{aligned}$$

Since $\overline{F}_{X_1}$ is normalized regularly varying, $I_{c,1} \sim M\alpha k^{\frac{1}{2}} \epsilon_n = o(1)$ for $\epsilon_n = k^{-\frac{1}{2}} / \log k$.

$$\begin{aligned} I_{c,2} &\le M\, k^{-\frac{1}{2}} n P\{ |\tau_{-l_n} - \mu(-l_n)| \ge \mu l_n/2 \} \\ &\quad + M\, k^{-\frac{1}{2}} n P\Big\{ \sum_{i \le -l_n} h(\tau_1 - \tau_i)/\epsilon_n A_i > b\left(\frac{n}{k}\right), |\tau_{-l_n} - \mu(-l_n)| < \mu l_n/2 \Big\}. \end{aligned}$$

In order to prove $I_{c,2} \to 0$ we assume $0 \le \tau_{i+1} - \tau_i \le 1$ for any integer $i$ again. By Marcinkiewicz-Zygmund's inequality and $\log n = O(\log l_n)$, we have that the first term in right hand side of above inequality converges to 0. Since $h(s) = o(s^{-\gamma})$, on the set $\{|\tau_{-l_n} - \mu(-l_n)| < \mu l_n/2\}$, we obtain for any $i \le -l_n$,

$$\frac{h(\tau_1 - \tau_i)}{\epsilon_n} \le \frac{h(-\tau_{-l_n})}{\epsilon_n} \le \frac{h(\mu l_n/2)}{\epsilon_n} = o(1).$$

Using inequality (4.3) we thus obtain that

$$\begin{aligned} I_{c,2} &\le o(1) + Mk^{\frac{1}{2}} \sum_{i \le -l_n} E\Big(\frac{h(\tau_1 - \tau_i)}{\epsilon_n}\Big)^{(\alpha\wedge 1)-\delta} + Mk^{\frac{1}{2}} I_{\{\alpha \ge 1\}} E\bigg(\sum_{i \le -l_n} \Big(\frac{h(\tau_1 - \tau_i)}{\epsilon_n}\Big)^{\frac{1-\delta}{\alpha+\delta}}\bigg)^{\alpha+\delta} \\ &\le o(1) + \ M\, k^{\frac{1}{2}} \epsilon_n^{-((\alpha\wedge 1)-\delta)} l_n^{-\gamma} \sum_{m \ge l_n} m^{\gamma} E h^{(\alpha\wedge 1)-\delta}(\tau_m) \\ &\quad + M\, I_{\{\alpha \ge 1\}} k^{\frac{1}{2}} \epsilon_n^{-(1-\delta)} l_n^{-\gamma} \Big(\sum_{m \ge l_n} m^{\frac{\gamma}{\alpha+\delta}} \big(E h^{1-\delta}(\tau_m)\big)^{\frac{1}{\alpha+\delta}}\Big)^{\alpha+\delta}. \end{aligned}$$

Therefore since $k^{1-\delta} l_n^{-\gamma} = o(1)$ for any $0 < \delta < \alpha \wedge 1$, by Proposition 4.3, we conclude $I_{c,2} \to 0$. Hence

$$\Big|E \exp(it \sum_{j=1}^{k_n^*} (U_j - EU_j)) - \prod_{j=1}^{k_n^*} E \exp(it(U_j - EU_j))\Big| \to 0.$$

Similarly we can show that

$$\Big|E \exp(it \sum_{j=1}^{k_n^*} (V_j - EV_j)) - \prod_{j=1}^{k_n^*} E \exp(it(V_j - EV_j))\Big| \to 0.$$

This ends the proof of theorem. ∎

## 4.4 References


[1] Barbe,Ph. and McCormick, W.P. (2004). Asymptotic expansions for infinite weighted convolutions of heavy tail distributions and applications, *Mem. Amer. Math. Soc.*, to appear, http://www.arxiv. org/abs/math.PR/0412537.

[2] Beirlant, J. and Teugels, J.L. (1992). Limit distributions for compounded sums of extreme order statistics. *J. Appl. Probab.*, **3,** 557–574.

[3] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1989). *Regular Variation*, 2nd Edition. Cambridge University Press. Cambridge.

[4] Chow, Y.S. and Teicher,H. (1988). *Probability theory: Independence, interchangeability, martingales*, 2nd edition. Springer-Verlag, New York.

[5] Datta, S. and McCormick, W.P. (1998). Inference for the tail parameters of a linear process with heavy tail innovations. *Ann. Inst. Statist. Math.*, **50,** 337–359.

[6] de Haan, L. and Resnick, S. (1998). On asymptotic normality of the Hill estimator. *Comm. Statist. Stochastic Models*, **4,** 849–866.

[7] Hill, B. (1975). A simple approach to inference about the tail of a distribution. *Ann. Statist.*, **3,** 743–756.

[8] Hsing, T. (1991). On tail index estimation using dependent data. *Ann. Statist.*, **19,** 1547–1569.

[9] Hsing, T. and Teugels, J. L. (1989). Extremal properties of shot noise processes. *Adv. Appl. Prob.*, **21,** 513-525.

[10] Lawrance, A.J. and Kottegoda, H.T. (1977). Stochastic modelling of riverflow time series. *J. R. Statist. Soc. A*, **140,** 1–14.

[11] Ling, S. and Peng., L. (2004). Hill's estimator for the tail index of an ARMA model. *J. Statist. Plann. Inference*, **123,** 279–293.

[12] Lund, R.B. (1996). The stability of storage models with shot noise input. *J. Appl. Prob.*, **33,** 830–839.

[13] McCormick, W.P. (1989). Extremal properties of shot noise processes. *Adv. Appl. Prob.*, **21,** 513-525.

[14] Pratt,J. (1960). On interchanging limits and integrals. *Ann. Math. Statist.*, **31,** 74-77.

[15] Resnick, S. (1987). *Extreme values, Regular Variation, and Point Processes*. Springer-Verlag, New York.

[16] Rootzen, H., Leadbetter, M.R. and de Haan, L.(1990). Tail and quantile estimation for strongly mixing stationary sequences. Technical Report 292, Center for Stochastic Processes, Department of Statistics, University of North Carolina, Chapel Hill, NC 27599–3260.

[17] Rootzen, H., Leadbetter, M.R. and de Haan, L. (1998). On the distribution tail array sums for strongly mixing stationary sequences. *Ann. Appl. Probab.*, **8,** 868–885.

[18] Resnick, S. and Stărică, C. (1995). Consistency of Hill's estimator for dependent data. *J. Appl. Prob.*, **32,** 139–167.

[19] Resnick, S. and Stărică, C. (1997). Asymptotic behavior of Hill's estimator for autoregressive data. *Comm. Statist. Stochastic Models*, **13,** 703–721.

[20] Resnick, S. and Stărică, C. (1998). Tail index estimation for dependent data. *Ann. Appl. Probab.*, **8,** 1156–1183.

[21] Resnick, S. and Willekens, E. (1991). Moving averages with random coefficients and random coefficient autoregressive models. *Comm. Statist. Stochastic Models*, **7,** 511–525.

# Chapter 5

## Parameter Estimation for First-order Bifurcating Autoregressive Processes with Weibull Type Innovations [1]

[1]Zhang, C. Submitted to *Journal of Time Series Analysis*, 5/2007.

**Abstract**

We study the first-order bifurcating autoregressive processes with Weibull type innovations, i.e. $X_t = \phi X_{\lfloor t/2 \rfloor} + \epsilon_t$, where $\lfloor x \rfloor$ is the largest integer which does not exceed $x$, and $\{(\epsilon_{2t}, \epsilon_{2t+1})\}$ is a sequence of independent and identically distributed (i.i.d.) random vectors with Weibull type distribution. The coefficient $\phi$ and tail index $\alpha$ are two parameters we are interested in. Using point process methods, we obtain the joint limit distribution of $(\hat{\phi}_n, \hat{\alpha})$. They are asymptotically independent.


**Keywords:** bifurcating autoregressive processes, Weibull distribution, regular variation, point process, tail index, Hill's estimator

## 5.1 Introduction

Consider the first-order bifurcating autoregressive process

$$X_t = \phi X_{\lfloor t/2 \rfloor} + \epsilon_t, \tag{5.1}$$

where $0 \leq \phi < 1$, $\lfloor x \rfloor$ is the largest integer which does not exceed $x$. We assume that $\{(\epsilon_{2t}, \epsilon_{2t+1})\}$ is a sequence of independent and identically distributed positive random vectors with bivariate Weibull distribution, or, more generally, the distribution function of $Y_t = \epsilon_{2t} \wedge \epsilon_{2t+1} = X_{2t} \wedge X_{2t+1} - \phi X_t$ is regularly varying at 0 with index $\alpha$. Roughly speaking, $P\{Y_t \leq x\} \sim x^\alpha$, as $x \to 0+$.

When $\alpha = 1$, the bivariate Weibull distribution is exactly bivariate exponential distribution. Zhou and Basawa (2005) have studied the maximum likelihood estimator, $\hat{\phi}_n = \min\{X_t / X_{\lfloor t/2 \rfloor} : 1 \leq t \leq n\}$, of $\phi$, and derived its exact and asymptotic distributions.

The first-order bifurcating autoregressive process, AR(1), is

$$X_t^* = \phi X_{t-1}^* + \epsilon_t^*,$$

where $\{\epsilon_t^*\}$ is a sequence of i.i.d. random variables with the same marginal distribution as $\epsilon_t$ and we assume $\epsilon_{2t}$ has the same distribution as $\epsilon_{2t+1}$.

Davis and McCormick (1989) have obtained the limit distribution of $\hat{\phi}_n = \min_t \frac{X_t}{X_{t-1}}$ for AR(1) processes. They have shown that when the innovation distribution is regularly varying at 0 for some constants $b_n$

$$P\{b_n(\hat{\phi}_n - \phi) > y\} \to \exp(-y^\alpha),$$

and $\hat{\phi}_n \xrightarrow{a.s.} \phi$.

The tail index, $\alpha$, is another parameter we are interested in. Hill's estimator is one of commonly used estimators for tail index. Resnick and Stărică (1997b) have obtained the asymptotic normality of $\hat{\alpha}$ based on $X_t^*$s and estimated residuals for AR(p) processes when the distribution of $\{\epsilon_t^*\}$ is regularly varying at infinity.

Point process method is used to derive that when $\alpha < 2$ for Weibull type innovation Hill's estimator based on estimated residuals, $\hat{Y}_t = X_{2t} \wedge X_{2t+1} - \hat{\phi}_n X_t$, has asymptotic normality under weak conditions. Meanwhile it is asymptotically independent of $\hat{\phi}_n$.

## 5.2 Main Results.

We assume that the $\{(\epsilon_{2t}, \epsilon_{2t+1})\}$ in model (5.1) is a sequence of independent random vectors such that the distribution of $Y_t = \epsilon_{2t} \wedge \epsilon_{2t+1} = X_{2t} \wedge X_{2t+1} - \phi X_t$, $F_Y$, is regularly varying at 0 with index $\alpha > 0$, i.e. for all $x > 0$

$$\lim_{t \to 0+} \frac{F_Y(tx)}{F_Y(t)} = x^\alpha.$$

Roughly speaking, $F_Y(x) \sim x^\alpha$, as $x \to 0+$.

**Example**: Bivariate Weibull distribution (Hougaard (1986)) for $0 < \delta \le 1$

$$\overline{F}(x_1, x_2) = P\{\epsilon_{2t} > x_1, \epsilon_{2t+1} > x_2\} = \exp\Big\{ - \beta\Big(\Big(\frac{x_1}{\lambda_1}\Big)^{\alpha_1/\delta} + \Big(\frac{x_2}{\lambda_2}\Big)^{\alpha_2/\delta}\Big)^\delta\Big\} \tag{5.2}$$

or (Marshall and Olkin (1967))

$$\overline{F}(x_1, x_2) = \exp\Big\{ - \lambda_1 x_1^{\alpha_1} - \lambda_2 x_2^{\alpha_2} - \lambda_{12}\big(x_1^{\alpha_1} \vee x_2^{\alpha_2}\big)\Big\} \tag{5.3}$$

Then $F_Y(x) = 1 - \exp\big\{ - \lambda x^\alpha\big\}$ is regularly varying with index $\alpha = \alpha_1 = \alpha_2$ at 0.

By Karamata theorem, Resnick (1987) or Bingham et al. (1989), $F_Y$ is regularly varying at 0 with index $\alpha$ if its density $f_Y$ is regularly varying at 0 with index $\alpha - 1$. Many commonly used distributions are regularly varying at 0 such as $\chi^2$, exponential, F, Weibull distributions.

Define finite order moving average processes $X^*_{t,q} = \sum_{j=0}^{q} \phi^j \epsilon^*_{t-j}$ and $X_{t,q} = \sum_{j=0}^{q} \phi^j \epsilon_{\lfloor t/2^j \rfloor}$ for $q \geq 1$. They have the same distribution. Let $b^{-1}(t) = \inf\{x : F_{Y_1}(x) \geq 1/t\}$ for $t > 1$.

First we state the following result on convergence in distribution of a sequence of point processes. Let $E$ be a locally compact second countable Hausdorff topological space, which for our purposes will be a subset of Euclidean space. For any $x \in E$, let $\varepsilon_{\{x\}}(A) = 1$ if $x \in A$ and 0 otherwise. Let $\mathscr{E}$ be the Borel $\sigma$-algebra of subsets of $E$, i.e. the $\sigma$-algebra generated by open sets (the topology), $M_+(E)$ ($M_p(E)$) be all non-negative (integer-valued) Radon measures on $(E, \mathscr{E})$ and $\mathscr{M}_+(E)$ ($\mathscr{M}_p(E)$) be the $\sigma$-algebra generated by vague topology. Because the vague topology on $M_+(E)$ and $M_p(E)$ is metrizable as a complete separable metric space we may use the terminology weak convergence, i.e. convergence in distribution which will be denoted by $\Longrightarrow$. For background information on point processes see Kallenberg (1983) or Resnick (1987).

**Lemma 2.1.** *For any $q \geq 1$, on $M_p\big([0,\infty) \times [0,\infty)\big)$,*

$$\xi_{n,q} := \sum_{t=2^q}^{n} \varepsilon_{(b(n)Y_t, X_{t,q})} \Longrightarrow \xi^*_q,$$

*where $\xi^*_q$ is Poisson point process with mean measure $\mu_q(\,\mathrm{d}x \times \,\mathrm{d}y) = \alpha x^{\alpha-1}\,\mathrm{d}x \times P\{X^*_{1,q} \in \mathrm{d}y\}$.*

The proof will be postponed to next section.

Resnick and Stărică (1997a) have shown that if $k$ is a sequence such that $k \wedge (n/k) \to \infty$

$$\frac{1}{\sqrt{k}} \sum_{t=1}^{n} \left(\varepsilon_{b(n/k)Y_t}([0, y^{1/\alpha}]) - P\{b(n/k)Y_t \leq y^{1/\alpha}\}\right) \Longrightarrow W(y)$$

on $D(0,\infty)$, where $W$ is a standard Wiener process, and, if $\sqrt{k}\left(\frac{n}{k} P\{b(n/k)Y_1 \leq y^{1/\alpha}\} - y\right) \to 0(*)$ for all $y > 0$,

$$\sqrt{k}\,(V_n(y) - y) := \sqrt{k}\left(\frac{1}{k}\sum_{t=1}^{n} \varepsilon_{b(n/k)Y_t}[0, y^{1/\alpha}] - y\right) \Longrightarrow W(y).$$

We remark that Condition (*) is actually a kind of necessary and reasonable condition for asymptotic normality of Hill's estimator.

**Example (cont.)**: For bivariate Weibull distribution, $b^{-1}(t) = \big(\frac{1}{c}\log(1-1/t)^{-1}\big)^{1/\alpha}$. If $k^{3/2} = O(n)$ then

$$\sqrt{k}\,\Big(\frac{n}{k}P\{b(n/k)Y_1 \le y^{1/\alpha}\} - y\Big) = \sqrt{k}\,\bigg(\frac{n}{k}\Big(1-\big(1-\frac{k}{n}\big)^y\Big) - y\bigg) = \frac{k^{3/2}}{n}o(1) = o(1)$$

Let $e$ denote the identity mapping. We conclude $\{(\xi_{n,q}, \sqrt{k}\,(V_n - e))\}$ is tight, or equivalently relatively compact in $M_p\big([0,\infty)\times[0,\infty)\big)\times D(0,\infty)$ because both $\xi_{n,q}$ and $\sqrt{k}\,(V_n - e)$ converge weakly.

**Lemma 2.2.** *If $\sqrt{k}\,\big(\frac{n}{k}P\{b(n/k)Y_1 \le y^{1/\alpha}\} - y\big) \to 0$ for all $y > 0$, finite-dimensional distributions of $\{(\xi_{n,q}, (V_n - e))\}$, or $\{(\xi_{n,q}, (V_n - EV_n))\}$ converge weakly.*

For complete separable metric space $M_p\big([0,\infty)\times[0,\infty)\big)\times D(0,\infty)$, by Prohorov's theorem, the tightness is equivalent to relative compactness. Then tightness and convergence of finite-dimensional distributions implies weak convergence. Consequently, for each $q \ge 1$,

$$(\xi_{n,q}, \sqrt{k}\,(V_n - e)) \Longrightarrow \xi_q^* \times W,$$

where $X \times Y$ means that $X$ is independent of $Y$.

**Theorem 2.1.** As $k \wedge (n/k) \to \infty$, in $M_p\big([0,\infty)\times[0,\infty)\big)\times D(0,\infty)$,

$$(\xi_n, \sqrt{k}\,(V_n - EV_n)) \Longrightarrow \xi^* \times W,$$

where $\xi_n := \sum_{t=1}^n \varepsilon_{(b(n)Y_t, X_t)}$, and $\xi^*$ is the Poisson process with mean measure $\mu(\,\mathrm{d}x \times \mathrm{d}y) = \alpha x^{\alpha-1}\,\mathrm{d}x \times P\{X_1^* \in \mathrm{d}y\}$. If $\sqrt{k}\,\big(\frac{n}{k}P\{b(n/k)Y_1 \le y^{1/\alpha}\} - y\big) \to 0$ for all $y > 0$,

$$(\xi_n, \sqrt{k}\,(V_n - e)) \Longrightarrow \xi^* \times W. \tag{5.4}$$

**Proof:** Clearly $\xi_q^* \times W \Longrightarrow \xi^* \times W$, as $q \to \infty$. The desired result follows by applying Theorem 3.2 in Billingsley (1996). For details see the proof of Lemma 2.2 in Davis and McCormick (1989). ∎

**Propostion 2.1.** *Let* $\hat\phi_n = \bigwedge_{t=2}^n \frac{X_t}{X_{\lfloor t/2\rfloor}}$. *For any* $y > 0$,

$$P\{\big(E(X_1^*)^\alpha\big)^{1/\alpha} b(n)(\hat\phi_n - \phi) > y\} \to \exp(-y^\alpha),$$

*and* $\hat\phi_n \xrightarrow{a.s.} \phi$.

**Proof:** Define a subset of $\mathbb{R}_+^2$ by $A_y = \{(x_1, x_2) : x_1/x_2 \le y, x_1, x_2 > 0\}$. Observe that

$$\{\xi_n(A_y) = 0\} = \Big\{\bigwedge_{t=2}^{2n+1} b(n)\Big(\frac{X_t}{X_{\lfloor t/2\rfloor}} - \phi\Big) > y\Big\}. \tag{5.5}$$

Thus by (5.4) and (5.5), we have that

$$\begin{aligned}
\lim_{n\to\infty} P\Big\{b(n)\Big(\bigwedge_{t=2}^{n} \frac{X_t}{X_{\lfloor t/2\rfloor}} - \phi\Big) > y\Big\} &= \lim_{n\to\infty} P\{\xi_n(A_y) = 0\} \\
&= P\{\xi^*(A_y) = 0\} \\
&= \exp\left(-\int_0^\infty \overline{F}_{X_1^*}(x/y)\alpha x^{\alpha-1}\,\mathrm{d}x\right) \\
&= \exp(-E(X_1^*)^\alpha y^\alpha).
\end{aligned}$$

Since $b(n) \to 0$ we must have $\hat\phi_n \xrightarrow{P} \phi$. But this implies $\hat\phi_n \xrightarrow{a.s.} \phi$ because $\hat\phi_n > \phi$ and is nonincreasing. ■

Via standard argument, see Resnick and Stărică (1997a), if $F_{Y_1}$ is regularly varying at 0 with index $\alpha$ and $\sqrt{k}\left(\frac{n}{k}F_{Y_1}(y^{1/\alpha}/b(n/k)) - y\right) \to 0$ for all $y > 0$, we obtain joint limit distribution of coefficient and tail index estimator, i.e. as $k \wedge (n/k) \to \infty$

$$\left(b(n)(\hat\phi_n - \phi), \sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k-1}\log\frac{Y_{(k)}}{Y_{(i)}} - \alpha\right)\right) \Longrightarrow \zeta \times N(0, \alpha^2),$$

where $P\{\zeta > x\} = e^{-E(X_1^*)^\alpha y^\alpha}$ and $Y_{(i)}$'s are the order statistics of $Y_t's, 1 \le t \le n = 2^{m-1} - 1$.

**Theorem 2.2.** *If* $F_{Y_1}$ *is regularly varying at 0 with index* $\alpha$ *and* $\sqrt{k}\left(\frac{n}{k}P\{b(n/k)Y_1 \le y^{1/\alpha}\} - y\right) \to 0$ *for all* $y > 0$, *when* $\alpha < 2$ *and* $Ee^{\lambda X_1} + Ee^{\lambda\epsilon_1} < \infty$ *for some* $\lambda > 0$, *we have joint limit distribution of coefficient and tail index estimators, i.e. as* $(k/\log n) \wedge (n/k) \to \infty$,

$$\left(b(n)(\hat\phi_n - \phi), \sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k-1}\log\frac{\hat Y_{(k)}}{\hat Y_{(i)}} - \alpha\right)\right) \Longrightarrow \zeta \times N(0, \alpha^2),$$

*where* $\hat Y_{(i)}s$ *are the order statistics of* $\hat Y_t = X_{2t} \wedge X_{2t+1} - \hat\phi_n X_t$.

**Proof:** The proof is similar to that of Proposition 3.2 in Resnick and Stărică (1997b). It is sufficient to show that for any interval $[c,d] \in (0,\infty)$ and $\delta > 0$

$$P\Big\{\frac{1}{\sqrt{k}} \sup_{x\in[c,d]} \sum_{t=1}^{n} \big(\epsilon_{b(n/k)\hat{Y}_t}[0,x] - \epsilon_{b(n/k)Y_t}[0,x]\big) > \delta\Big\} \longrightarrow 0.$$

Following the proof of Proposition 3.2 in Resnick and Stărică (1997b) we obtain

$$\begin{aligned}
& P\Big\{\frac{1}{\sqrt{k}} \sup_{x\in[c,d]} \sum_{t=1}^{n} \big(\epsilon_{b(n/k)\hat{Y}_t}[0,x] - \epsilon_{b(n/k)Y_t}[0,x]\big) > \delta\Big\} \\
\leq\ & P\Big\{\frac{1}{\sqrt{k}} \sup_{x\in[c,d]} \sum_{t=1}^{n} I_{\{x<b(n/k)Y_t\leq(1+\epsilon_n)x\}} > \delta/2\Big\} + P\Big\{\frac{1}{\sqrt{k}} \sum_{t=1}^{n} I_{\{(\hat{\phi}_n-\phi)b(n/k)X_t>\epsilon_n c\}} > \delta/2\Big\} \\
\leq\ & P\Big\{\frac{1}{\sqrt{k}} \sup_{x\in[c,d]} \sum_{t=1}^{n} I_{\{x<b(n/k)Y_t\leq(1+\epsilon_n)x\}} > \delta/2\Big\} \\
& \quad + P\Big\{b(n)(\hat{\phi}_n-\phi) \geq M\Big\} + \frac{2n}{\delta\sqrt{k}} e^{-\lambda c\epsilon_n b(n)/(Mb(n/k))} Ee^{\lambda X_n} \\
=\ & I_{a1} + I_{a2} + I_{a3},
\end{aligned}$$

and $I_{a1} \to 0$ if $\sqrt{k}\epsilon_n \to 0$. Choose $\ln n = o(k)$ and $\epsilon_n \sim k^{-\delta}$ with $1/2 < \delta < 1/\alpha$. Note that $Ee^{\lambda X_n} \leq Ee^{\lambda X_1} \prod_{j\geq 0} Ee^{\lambda\phi^j\epsilon_1}$ is bounded above iff $\sum_{j=0}^{\infty} E(e^{\lambda\phi^j\epsilon_1} - 1) < \infty$. It is easy to see that by Taylor's expansion $\sum_{j=k}^{\infty} E(e^{\lambda\phi^j\epsilon_1} - 1) \leq \sum_{j=k}^{\infty} E(\lambda\phi^j\epsilon_1 e^{\lambda\phi^j\epsilon_1}) \leq \sum_{j=k}^{\infty} \lambda\phi^j E(\epsilon_1^2 + e^{2\lambda\phi^j\epsilon_1}) < \infty$. Condition $Ee^{\lambda X_1} + Ee^{\lambda\epsilon_1} < \infty$ yields $I_{a3} \to 0$. Let $M \to \infty$. The desired result follows from proposition 5.2. ■

We remark that the larger $\alpha$ is, the less observations are in a neighborhood of 0. Then it is difficult to estimate large $\alpha$.

## 5.3 Proof of Lemmas.

**Proof of Lemma 2.1.** Define $DC - semiring$

$$\mathscr{T} = \Big\{[a,b)\times[a',b') : P\{X_{1,q}^* = a' \text{ or } b'\} = 0,\ a, a' \geq 0\Big\}.$$

By theorem 4.2 in Kallenberg (1983), it is equivalent to prove that for any disjoint sets, $R_i := [a_i,b_i)\times[a_i',b_i') \in \mathscr{T}$, $1 \leq i \leq m$,

$$\big(\xi_{n,q}(R_1), \xi_{n,q}(R_2), \cdots \xi_{n,q}(R_m)\big) \xrightarrow{d} \big(\xi_q^*(R_1), \xi_q^*(R_2), \cdots \xi_q^*(R_m)\big). \tag{5.6}$$

We will only prove (5.6) for the case $m = 2$ with $[a_1, b_1) \bigcap [a_2, b_2) = \emptyset$. The proof for the other cases is quite similar, and thus it is omitted.

Let $A_j := [a_j, b_j) \times [a'_j, b'_j)^c$, $j = 1, 2$ and $B_{k_1,k_2,l_1,l_2} := \{\xi_{n,q}(R_j) = k_j, \xi_{n,q}(A_j) = l_j, j = 1, 2\}$. For any nonnegative integers $k_1, k_2$ and $l_2 := m - l_1$,

$$P\{\xi_{n,q}(R_1) = k_1, \xi_{n,q}(R_2) = k_2\} = \sum_{m=0}^{\infty}\sum_{l_1=0}^{m} P\{B_{k_1,k_2,l_1,l_2}\} I_{\{m\le(n-2^q+1)-k_1-k_2\}}.$$

If $P\{B_{k_1,k_2,l_1,l_2}\} I_{\{m\le(n-2^q+1)-k_1-k_2\}}$ can be bounded by a sequence $s_n(m, l_1)$, which satisfies

$$\lim_{n\to\infty}\sum_{m=0}^{\infty}\sum_{l_1=0}^{m} s_n(m, l_1) = \sum_{m=0}^{\infty}\sum_{l_1=0}^{m}\lim_{n\to\infty} s_n(m, l_1), \tag{5.7}$$

then by a commonly used variant of dominated convergence theorem sometimes called Pratt's lemma (Pratt (1960)) we only need to show that

$$P\{\xi^*_{n,q}(R_1) = k_1, \xi^*_{n,q}(R_2) = k_2\} = \sum_{m=0}^{\infty}\sum_{l_1=0}^{m}\lim_{n\to\infty} P\{B_{k_1,k_2,l_1,l_2}\}.$$

Define for $m \le (n - 2^q + 1) - k_1 - k_2$

$$\begin{aligned} s_n(m, l_1) \;\; := \;\; & \frac{n^{k_1+k_2+m}}{k_1!\,k_2!\,m!}\binom{m}{l_1}\big(P\{b(n)Y_n \in [a_1, b_1)\}\big)^{k_1+l_1}\big(P\{b(n)Y_n \in [a_2, b_2)\}\big)^{k_2+(m-l_1)} \\ & \big(1 - P\{b(n)Y_n \in [a_1, b_1)\bigcup[a_2, b_2)\}\big)^{n-k_1-k_2-m-2^q}. \end{aligned}$$

Since $\#\{t : b(n)Y_t \in [a_j, b_j), 2^q \le t \le n\} = k_j + l_j$ for $j = 1, 2$ when $\xi_{n,q}(R_j) = k_j, \xi_{n,q}(A_j) = l_j$, $j = 1, 2$, it is easy to check that $s_n(m, l_1)$ is an upper bound of $P\{\xi_{n,q}(R_j) = k_j, \xi_{n,q}(A_j) = l_j, j = 1, 2\}$ and either side of equation (5.7) is $\frac{(b_1^\alpha-a_1^\alpha)^{k_1}}{k_1!}\frac{(b_2^\alpha-a_2^\alpha)^{k_2}}{k_2!}$.

Next we will prove

$$\begin{aligned} & P\{B_{k_1,k_2,l_1,l_2}\} \\ \to \;\; & \frac{\big((b_1^\alpha - a_1^\alpha)P\{X^*_{1,q} \in [a'_1, b'_1)\}\big)^{k_1}}{k_1!}\frac{\big((b_2^\alpha - a_2^\alpha)P\{X^*_{1,q} \in [a'_2, b'_2)\}\big)^{k_2}}{k_2!}e^{-(b_1^\alpha-a_1^\alpha+b_2^\alpha-a_2^\alpha)} \\ & \frac{1}{m!}\binom{m}{l_1}\big((b_1^\alpha - a_1^\alpha)P\{X^*_{1,q} \in [a'_1, b'_1)^c\}\big)^{l_1}\big((b_2^\alpha - a_2^\alpha)P\{X^*_{1,q} \in [a'_2, b'_2)^c\}\big)^{l_2} \\ & =: b(k_1, k_2, l_1, l_2). \end{aligned}$$

It is obvious that for each $t$, $\varepsilon_{(b(n)Y_t,X_{t,q})}$ is independent of all except finite, $m_q$(say), $\varepsilon_{(b(n)Y_s,X_{s,q})}$'s. Then there are at least $\frac{1}{k_1!\,k_2!\,l_1!\,l_2!}\prod_{i=1}^{k_1+k_2+m}((n - 2^q + 1) - m_q i) \sim \frac{n^{k_1+k_2+m}}{k_1!\,k_2!\,m!}\binom{m}{l_1}$

choices such that $(b(n)Y_r, X_{r,q}) \in R_j, r = r_1, \cdots, r_{K_j}$, $(b(n)Y_r, X_{r,q}) \in A_j, s = s_1, \cdots, s_{l_j}$, $j = 1, 2$, are independent. Therefore

$$\begin{aligned} P\{B_{k_1,k_2,l_1,l_2}\} \sim & \frac{n^{k_1+k_2+m}}{k_1!\, k_2!\, m!}\binom{m}{l_1}\prod_{j=1}^{2}\big(P\{(b(n)Y_t, X_{t,q}) \in R_j\}\big)^{k_j}\big(P\{(b(n)Y_t, X_{t,q}) \in A_j\}\big)^{l_j} \\ & \big(1 - P\{b(n)Y_n \in [a_1, b_1)\bigcup[a_2, b_2)\}\big)^{n-(k_1+k_2+m)(q+1)-2^q} \end{aligned}$$

This ends the proof of the lemma. ■

We remark that if the distribution of $\epsilon_1$ is continuous, which implies $F_{X^*_{1,q}}$ is also continuous, then the Poisson process $\xi^*_q$ is simple. By Theorem 4.7 in Kallenberg (1983) we only need prove for any $R_i \in \mathscr{T}$, $1 \le i \le m$,

$$P\big\{\xi_{n,q}(\bigcup_{i=1}^{m} R_i) = 0\big\} \to P\big\{\xi^*_q(\bigcup_{i=1}^{m} R_i) = 0\big\}.$$

**Proof of Lemma 2.2.** We only need prove that for any sequence of disjoint sets, $[a_i, b_i), i = 1, \cdots, m_1$, and $[c_j, d_j), j = 1, \cdots, m_2$, $R_i = [a_i, b_i) \times [a'_i, b'_i) \in \mathscr{T}$, $m_1, m_2 \ge 1$,

$$\begin{aligned} & (\xi_{n,q}, \sqrt{k}\,(V_n - EV_n))\,((R_1, \cdots, R_{m_1}), ([c_1, d_1), \cdots, [c_{m_2}, d_{m_2}))) \\ \xrightarrow{d} \; & \xi^*_q(R_1, \cdots, R_{m_1}) \times W([c_1, d_1), \cdots, [c_{m_2}, d_{m_2})), \end{aligned} \tag{5.8}$$

or, equivalently, convergence of corresponding characteristic function. To ease notation we will only prove (5.8) with $m_1 = 2$ and use the same notations as those in the proof of lemma 5.2. For any $a^*_1, a^*_2, b^*_j \in \mathbb{R}$, $1 \le j \le m_2$, since

$$\begin{aligned} & E\exp\{i(\sum_{p=1}^{2} a^*_p \xi_{n,q}(R_p) + \sum_{j=1}^{m_2} b^*_j (V_n - EV_n)[c_j, d_j))\} \\ = & \sum_{k_1,k_2,l_1,l_2} \exp\{i\sum_{p=1}^{2} a^*_p k_p\} E\Big(I_{B_{k_1,k_2,l_1,l_2}} \exp\{i\sum_{j=1}^{m_2} b^*_j (V_n - EV_n)[c_j, d_j)\}\Big), \end{aligned}$$

by Pratt's lemma (see the argument in the proof of lemma 5.2), it is sufficient to prove that

$$E\Big(I_{B_{k_1,k_2,l_1,l_2}} \exp\{i\sum_{j=1}^{m_2} b^*_j (V_n - EV_n)[c_j, d_j)\}\Big) \to b(k_1, k_2, l_1, l_2) e^{-\sum_{j=1}^{m_2} (b^*_j)^2 (d_j - c_j)/2}.$$

Since $\{(\varepsilon_{b(n/k)Y_t} - E\varepsilon_{b(n/k)Y_t})\big([c_1^{1/\alpha}, d_1^{1/\alpha}), \cdots, [c_{m_2}^{1/\alpha}, d_{m_2}^{1/\alpha})\big)\}$ is asymptotically independent of $\{b(n)Y_t \in [a, b)\}$ for any $0 < a < b < \infty$, intuitively, the above conclusion is correct. Observe that

$$
\begin{aligned}
& \Big(EI_{\{b(n)Y_n\in[a_1,b_1)^c\bigcap[a_2,b_2)^c\}}\exp\{i\sum_{j=1}^{m_2}b_j^*(V_n-EV_n)[c_j,d_j)\}\Big)^n \\
= \quad & \Big(EI_{\{b(n)Y_n\in[a_1,b_1)^c\bigcap[a_2,b_2)^c\}}(1+o(1))\exp\{i\sum_{t=1}^{n-1}\sum_{j=1}^{m_2}b_j^*(\varepsilon_{b(n/k)Y_n}-E\varepsilon_{b(n/k)Y_n})[c_j^{1/\alpha},d_j^{1/\alpha})\}\Big)^n \\
\sim \quad & e^{-(b_1^\alpha-a_1^\alpha+b_2^\alpha-a_2^\alpha)}e^{-\sum_{j=1}^{m_2}(b_j^*)^2(d_j-c_j)/2}.
\end{aligned}
$$

Following the argument of $P\{B_{k_1,k_2,l_1,l_2}\}\sim b(k_1,k_2,l_1,l_2)$, we obtain

$$
\begin{aligned}
& E\Big(I_{B_{l_1,l_2}}\exp\{i\sum_{j=1}^{m_2}b_j^*(V_n-EV_n)[c_j,d_j)\}\Big) \\
\sim \quad & \frac{n^{k_1+k_2+m}}{k_1!\,k_2!\,m!}\binom{m}{l_1}(1+o(1))\prod_{j=1}^{2}\big(P\{(b(n)Y_t,X_{t,q})\in R_j\}\big)^{k_j}\big(P\{(b(n)Y_t,X_{t,q})\in A_j\}\big)^{l_j} \\
& \qquad \Big(EI_{\{b(n)Y_n\in[a_1,b_1)^c\bigcap[a_2,b_2)^c\}}\exp\{i\sum_{j=1}^{m_2}b_j^*(V_n-EV_n)[c_j,d_j)\}\Big)^{n-(k_1+k_2+m)(q+1)-2^q} \\
\sim \quad & b(k_1,k_2,l_1,l_2)e^{-\sum_{j=1}^{m_2}(b_j^*)^2(d_j-c_j)/2}.
\end{aligned}
$$

This completes the proof of the lemma. ■

## 5.4 References


[1] Billingsley, P. (1996). *Convergence of Probability Measures.* 2nd edition. Wiley, New York.

[2] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1989). Regular Variation, 2nd Edition. Cambridge University Press. Cambridge.

[3] Davis, R.A. and McCormick, W.P. (1989). Estimation for first-order autoregressive processes with positive or bounded innovations. *Stochastic Process. Appl.*, **31,** 237–250.

[4] Hougaard, P. (1986). A Class of Multivariate Failure Time Distributions, *Biometrika*, **73,** 671–678.

[5] Kallenberg, O. (1983). *Random Measures.* 3rd ed. Academie-Verlag, Berlin.

[6] Marshall, A.W. and Olkin, I. (1967). *A multivariate exponential distribution.* Journal of Amer. Statist. Assoc. 62, 30–44.

[7] Pratt,J. (1960). On interchanging limits and integrals. *Ann. Math. Statist.*, **31,** 74–77.

[8] Resnick, S. (1987). *Extreme Values, Regular Variation and Point Processes.* Springer-Verlag, Berlin.

[9] Resnick, S. and Stărică, C. (1997a). Smoothing the Hill estimator. *Adv. Appl. Probab.*, **29,** 271–293.

[10] Resnick, S. and Stărică, C. (1997b). Asymptotic behavior of Hill's estimator for autoregressive data. *Comm. Statist. Stochastic Models*, **13,** 703–721.

[11] Zhou, J. and Basawa, I.V. (2005). Maximum Likelihood Estimation for a First-Order Bifurcating Autoregressive Process with Exponential Errors. *J. Time Series Anal.*, **26,** 825–842.

# Chapter 6

## Conclusion and Future Work

### 6.1 Maxima of Random Walks

In Chapter 2 we studied the maxima of random walks generated by random variables with regularly varying (right) tail, and obtained the asymptotic expansion for its tail distribution. This is the first higher order result in the literature so far. Our result can be easily implement with computer algebra package such as Maple.

In next chapter, we discussed the random sum of rapidly varying subexponential random variables with (right) tail, i.e. its hazard rate function $h$ satisfying

$$\begin{aligned}&h \text{ is smoothly varying,}\\ &\lim_{t\to\infty} th(t) = +\infty,\ \lim_{t\to\infty} h(t) = 0 \qquad \text{and} \qquad \liminf_{t\to\infty} th(t)/\log t > 0\,. \end{aligned} \tag{6.1}$$

It is very natural to ask what are the asymptotic expansions for the tail distributions of maxima of random walks generated by rapidly varying subexponential random variables.

Let the renewal function $U = \sum_{n\geq 0} F_-^{\star n}$. Following the proof of Lemma 1 in Chapter 2, by Faà di Bruno's formula we can show that the hazard rate function, $h_+$, of the distribution of first ascending ladder height is

$$h_+^{(k-1)} = \Big(-\log\int \overline{F}\,\mathrm{d}U\Big)^{(k)} = \sum_{\substack{i_1+\cdots+ki_k=k\\ i_1,\cdots,i_k\geq 0}} \Big(\sum_{1\leq j\leq k} i_j - 1\Big)!\binom{k}{i}\prod_{1\leq j\leq k}\left(\frac{-\int \overline{F}^{(j)}\,\mathrm{d}U}{j!\int \overline{F}\,\mathrm{d}U}\right)^{i_j}$$

and

$$\sum_{n\geq 0} U_{F_-}^n\,\overline{F}^{(j)} = \sum_{i'_1+\cdots+ji'_j=j}\binom{j}{i'}\sum_{n\geq 0} U_{F_-}^n \prod_{1\leq l\leq j}\left(\frac{-h^{(l-1)}}{l!}\right)^{i_l}\overline{F}.$$

It can be shown that $h_+ \sim h$. Intuitively we can say $h_+^{(m)} \sim h^{(m)}$. There are so many terms in $h^{(m)}$ that it is easy to prove it. We provide the following conjecture which is the key to further work and need to be proved.

**Conjecture.** *The hazard rate function $h_+$ of strict ascending ladder height distribution $F_+$ is smoothly varying of same order as $h$.*

Under this conjecture we can get the asymptotic expansions for the tail distribution of maximum random walks generated by rapidly varying subexponential random variables such as lognormal random variables.

## 6.2 Random Difference Equation

Random difference(recurrence) equation is defined as

$$Y_n = M_n Y_{n-1} + Q_n, \tag{6.2}$$

where $\{(M_n, Q_n)\}$ is a sequence of independent and identically distributed (i.i.d.) random matrices, $M_n$ is a $d \times d$ matrix and $Q_n$ is a $d \times 1$ vector.

When $Y_n = (X_n, \cdots, X_{n-p+1})'$, $M_n = \begin{pmatrix} \phi_1 & \cdots & \phi_p \\ & I & 0 \end{pmatrix}$ and $Q_n = (\epsilon_n, 0, \cdots, 0)$, equation (6.2) is a random coefficient AR(p) autoregressive sequence.

Autoregressive Conditional Heteroscedasticity(ARCH) model introduced by Engle (1982) has form

$$\xi_n = X_n(\beta + \lambda \xi_{n-1}^2)^{1/2}.$$

Obviously $\xi_n^2$ satisfies (6.2) with $M_n = \lambda X_n^2$ and $Q_n = \beta X_n^2$.

Bollerslev (1986) proposed Generalized Autoregressive Conditional Heteroskedastic (GARCH) model as a generalization of Engle(1982). GARCH(1,1) model is the model of form

$$X_t = \sigma_t Z_t, \text{ and } \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2.$$

If $\beta_1 = 0$ above, it is a ARCH(1) model. Clearly $(X_n^2, \sigma_n^2)'$ satisfies (6.2) with $M_n = \begin{pmatrix} \alpha_1 Z_n^2 & \beta_1 Z_n^2 \\ \alpha_1 & \beta_1 \end{pmatrix}$ and $Q_n = (\alpha_0 Z_n^2, \alpha_0)$.

Kesten(1973) studied the asymptotic behavior of the limiting distribution of $R$, the solution of (6.2). When $p = 1$, i.e. $M_n, Q_n$'s are random variables, under mild moment conditions on $(M_1, Q_1)$, there exists a positive constant $c > 0$ such that

$$P\left(R = \sum_{k=1}^{\infty} M_1 \cdots M_{k-1} Q_k > t\right) \sim ct^{-\alpha},$$

as $x \to \infty$.

This is a very interesting phenomena. Although both coefficient and error have light tail, the solution may have a heavy tail. We are interested in the higher order expansions of the tail distribution of $R$.

## 6.3 Tail Index Estimation

In chapter 4 we obtain asymptotic properties of Hill's estimator for shot noise sequence

$$X_j = \sum_{i \le j} A_i h(\tau_j - \tau_i)$$

If we can only observe $X(t) = \sum_{\tau_i \le t} A_i h(t - \tau_i)$ on a lattice set $A = \{n\Delta : n \in \mathbb{N}\}$ it is not clear that whether Hill's estimator is asymptotically normal.

We study first-order bifurcating autoregressive process with Weibull type innovations in chapter 5. The joint limiting distribution of coefficient and tail index estimator is obtained. Dr. Ishwar Basawa proposed the parameter estimation problem for first-order bifurcating autoregressive process with bivariate Pareto errors, i.e.

$$X_t = \phi X_{\lfloor t/2 \rfloor} + \epsilon_t,$$

where $(\epsilon_{2t}, \epsilon_{2t+1})$'s have density function

$$f(t_1, t_2) = \alpha(\alpha+1)(\theta_1\theta_2)^{\alpha+1}(\theta_2 t_1 + \theta_1 t_2 - \theta_1\theta_2)^{-(\alpha+2)}.$$

We are hoping to use point process technique to get asymptotic properties of parameter estimation.

## Bibliography


[1] Asmussen, S. (1987). *Applied Probability and Queues.* John Wiley & Sons, Ltd., Chichester.

[2] Asmussen, S. (1997). *Ruin Probabilities.* World Scientific, River Edge.

[3] Asmussen, S. and Kalashnikov, V., Konstantinides, D., Kluppelberg, C., Tsitsiashvili,G. (2002). A local limit theorem for random walk maxima with heavy tails, *Statist. Probab. Lett.*, **64,** 399–404.

[4] Athreya, K.B. and Ney, P.E. (1972). *Branching Processes.* Springer, New York.

[5] Barbe,Ph. and McCormick, W.P. (2004). Asymptotic expansions for infinite weighted convolutions of heavy tail distributions and applications, *Mem. Amer. Math. Soc.*, to appear, http://www.arxiv. org/abs/math.PR/0412537.

[6] Barbe, Ph. and McCormick, W.P. (2005). Asymptotic expansions for convolutions of light tailed subexponential distributions, to appear in *Prob. Theory Rel. Fields*, to appear, http://www. arxiv.org/abs/math.PR/0512141.

[7] Bartlett, M.S. (1963). The spectral analysis of point processes, *J. Roy. Statist. Ser. B.*, **25,** 264–296.

[8] Beirlant, J. and Teugels, J.L. (1992). Limit distributions for compounded sums of extreme order statistics. *J. Appl. Probab.*, **3,** 557–574.

[9] Beirlant, J. and Teugels, J.L., Vynckier, P. (1996). *Practical Analysis of Extreme Values*, Leuven University Press, Leuven, Belgium.

[10] Billingsley, P. (1996). *Convergence of Probability Measures.* 2nd edition. John Wiley & Sons, Ltd., New York.

[11] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1989). *Regular Variation.* 2nd Edition. Cambridge University Press. Cambridge.

[12] Bollerslev T, (1986). Generalised Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, **31,** 307–27,

[13] Borovkov, A.A. (2002). On subexponential distributions and asymptotics of the distribution of the maximum of sequential sums, *Siberian Math. J.* 43 995C1022.

[14] Chistyakov, V.P. (1964). A theorem on sums of independent positive random variables and its applications to branching random processes. *Theory Probab. Appl.*, **9,** 640–648.

[15] Chow, Y.S. and Teicher, H. (1988). *Probability Theory, Independence, Interchangeability, Martingales.* Springer, New York.

[16] Cline, D.B.H. (1987). Convolutions of distributions with exponential and subexponential tails, *J. Austr. Math. Soc. (A)*, 43, 347–365.

[17] Cohen, J.W. (1972). On the tail of the stationary waiting-time distribution and limit theorem for M/G/1 queue, *Ann. Inst. H. Poincaré,B*, 8, 255–263.

[18] Cowan, R. (1984). Statistical concepts in the analysis of cell lineage data. *In Proceedings of the 1983 Workshop on Cell Growth and Division*, p. 18-22. Latrobe Uni., Melbourne.

[19] Cowan, R. and Staudte, R. G. (1986). The bifurcating autoregression model in cell lineage studies. *Biometrics* 42, 769–783.

[20] Datta, S. and McCormick, W.P. (1998). Inference for the tail parameters of a linear process with heavy tail innovations. *Ann. Inst. Statist. Math.*, **50,** 337–359.

[21] Davis, R.A. and McCormick, W.P. (1989). Estimation for first-order autoregressive processes with positive or bounded innovations. *Stochastic Process. Appl.*, **31,** 237–250.

[22] de Haan, L. and Resnick, S. (1998). On asymptotic normality of the Hill estimator. *Comm. Statist. Stochastic Models*, **4,** 849–866.

[23] Doney, R.A. (1980). Moments of Ladder Heights in Random Walks, *J. Appl. Prob.*, 17, 248–252.

[24] Embrechts, P. (1985). Subexponential distribution functions and their applications: a review, in *Proceedings of the Seventh Conference on Probability Theory, (Braşov, 1982)*, 125–136, VNU Sci. Press, Utrecht, 1985.

[25] Embrechts, P., and Goldie, C.M. (1982). On Covolution Tails *Stoch. Proc. Appl.*, 13, 263–278.

[26] Embrechts, P., Goldie, C.M. and Veraverbeke, N. (1979). Subexponentiality and infinite divisibility, *Z. Wahrsch. verw. Geb.*, 49, 335–347.

[27] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events.* Springer, Berlin.

[28] Embrechts, P. and Veraverbeke, N. (1982). Estimates for the probability of ruin with special emphasis on the possibility of large claims, *Insurance: Mathematics and Economics*, 1, 55–72.

[29] Engle, R. F. (1982). Autoregressive conditional heteroscedastic models with estimates of the variance of United Kingdom inflation. *Econometrica*, **50,** 987–1007

[30] Feller, W. (1971). *An Introduction to Probability Theory and its Applications*, 2nd ed., Wiley.

[31] Geluk, J.L. (1992). Second order tail behaviour of a subordinated probability distribution. *Stoch. Proc. Appl.*, 40, 325–337.

[32] Geluk, J.L. (1996). Tails of subordinated laws: The regularly varying case. *Stoch. Proc. Appl.*, 61, 147–161.

[33] Goldie, C.M. (1991). Implicit renewal theory and tails of solutions of random equations, *Ann. Appl. Probab.*, 1, 126–166.

[34] Goldie, C. and Klüppelberg, C. (1998). Subexponential distributions, in *A Practical Guide to Heavy Tails, Statistical Techniques and Applications*, R. Adler, R. Feldman and M. Taqqu eds., Birkhäuser, 435–460.

[35] Greenwood, P., Omey, E. and Teugels, J.L. (1982). Harmonic renewal measures, *Z. Wahrssch. Verw. Gebiete*, 59, 391–409.

[36] Grübel, R. (1987). On subordinated distributions and generalized renewal measures, *Ann. Probab.*, 15, 394–415.

[37] Grübel, R. (1988). Harmonic renewal sequences and the first positive sum, *J. London Math. Soc.* , (2) 38 (1988) 179–192.

[38] Hill, B. (1975). A simple approach to inference about the tail of a distribution. *Ann. Statist.*, **3,** 743–756.

[39] Hougaard, P. (1986). A Class of Multivariate Failure Time Distributions, *Biometrika*, **73,** 671–678.

[40] Hsing, T. (1991). On tail index estimation using dependent data. *Ann. Statist.*, **19,** 1547–1569.

[41] Hsing, T. and Teugels, J. L. (1989). Extremal properties of shot noise processes. *Adv. Appl. Prob.*, **21,** 513-525.

[42] Huggins, R. M. (1995). A law of large numbers for the bifurcating autoregressive process. Comm. *Statistics. Stochastic Models* 11, 273–278.

[43] Kallenberg, O. (1983). *Random Measures.* 3rd ed. Academie-Verlag, Berlin.

[44] Kesten, H. (1973). Random Difference Equations and Renewal Theory for Products of Random Matrices. *Acta Math,* 131, 207–248

[45] Klüppelberg, C., Mikosch, T., and Schärf, A. (2003). Regular variation in the mean and stable limits for Poisson shot noise. *Bernoulli*, **9,** 467–496.

[46] Lawrance, A.J. and Kottegoda, H.T. (1977). Stochastic modelling of riverflow time series. *J. R. Statist. Soc. A*, **140,** 1–14.

[47] Lewis, P.A. (1964). A branching Poisson process model for the analysis of computer failure patterns *J. Roy. Statist. Soc. Ser. B*, **26,** 398–456.

[48] Ling, S. and Peng., L. (2004). Hill's estimator for the tail index of an ARMA model. *J. Statist. Plann. Inference*, **123,** 279–293.

[49] Lund, R.B. (1996). The stability of storage models with shot noise input. *J. Appl. Prob.*, **33,** 830–839.

[50] Marić, V. (2000). *Regular Variation and Differential Equations*, *Lecture Notes in Mathematics*, 1726, Springer.

[51] Marshall, A.W. and Olkin, I. (1967). *A multivariate exponential distribution.* Journal of Amer. Statist. Assoc. 62, 30–44.

[52] McCormick, W.P. (1989). Extremal properties of shot noise processes. *Adv. Appl. Prob.*, **21,** 513–525.

[53] Omey, E. and Willekens E. (1986). Second order behaviour of the tail of a subordinated probability distribution. *Stoch. Proc. Appl.*, 2, 339–353.

[54] Omey, E. and Willekens E. (1987). Second-order behaviour of distributions subordinate to a distribution with finite mean. *Comm. Statist. Stoch. Models*, 3, 311–342.

[55] Pakes, A.G. (1975). On the tails of waiting-time distributions, *J. Appl. Probab.*, 12, 555–564.

[56] Parzen, E. (1962). *Stochastic Processes.* Holden-Day, San Francisco.

[57] Pratt,J. (1960). On interchanging limits and integrals. *Ann. Math. Statist.*, **31,** 74–77.

[58] Resnick, S. (1987). *Extreme Values, Regular Variation, and Point Processes.* Springer-Verlag, New York.

[59] Resnick, S. and Stărică, C. (1995). Consistency of Hill's estimator for dependent data. *J. Appl. Prob.*, **32,** 139–167.

[60] Resnick, S. and Stărică, C. (1997a). Smoothing the Hill estimator. *Adv. Appl. Probab.*, **29,** 271–293.

[61] Resnick, S. and Stărică, C. (1997b). Asymptotic behavior of Hill's estimator for autoregressive data. *Comm. Statist. Stochastic Models*, **13,** 703–721.

[62] Resnick, S. and Stărică, C. (1998). Tail index estimation for dependent data. *Ann. Appl. Probab.*, **8,** 1156–1183.

[63] Resnick, S. and Willekens, E. (1991). Moving averages with random coefficients and random coefficient autoregressive models. *Comm. Statist. Stochastic Models*, **7,** 511–525.

[64] Samorodnitsky, G. (1998). Tail behavior of some shot noise processes. A practical guide to heavy tails. In Alder, Feldman and Taqqu (eds), *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, 473–486. Birkhüser Boston, Boston.

[65] Roman, S. (1980). The formula of Faà di Bruno, *Amer. Math. Monthly*, 87, 805–809.

[66] Rootzen, H., Leadbetter, M.R. and de Haan, L. (1990). Tail and quantile estimation for strongly mixing stationary sequences. Technical Report 292, Center for Stochastic

Processes, Department of Statistics, University of North Carolina, Chapel Hill, NC 27599–3260.

[67] Rootzen, H., Leadbetter, M.R. and de Haan, L. (1998). On the distribution tail array sums for strongly mixing stationry sequences. *Ann. Appl. Probab.*, **8,** 868–885.

[68] Teugels, J.L. (1975). The class of subexponential distributions, *Ann. Probab.*, 3, 1000–1011.

[69] Vere-Jones, D. (1970). Stochastic models for earthquake occurrence. *J. Roy. Statist. Soc. Ser. B*, **32,** 1–62.

[70] Veraverbeke, N. (1977). Asymptotic behavior of Wiener-Hopf factors of a random walk. *Stoch. Proc. Appl.*, 5, 27–37.

[71] Willmot, G.E. and Lin, X.S. (2000). *Lundberg Approximations for Compound Distributions with Insurance Applications*, *Lecture Notes in Statistics*, 156, Springer.

[72] Zachary, S. (2004). A note on Veraverbekes theorem, *Queueing Syst.*, 46, 9–14.

[73] Zhou, J. and Basawa, I.V. (2005). Maximum Likelihood Estimation for a First-Order Bifurcating Autoregressive Process with Exponential Errors. *J. Time Series Anal.*, **26,** 825–842.