Large Fluctuations in Financial Models

Milan Borkovec

Center of Mathematical Sciences
Munich University of Technology
D-80290 Munich

to my parents

Acknowledgement

This work would not have been possible without the advice and help of many people. In particular, I wish to express my gratitude to Prof. Claudia Klüppelberg for her constant support and encouragement. I am also grateful to her for giving me the opportunity to travel and getting in touch with distinguished scientists in the field of applied probability theory.

I would also like to thank Prof. Holger Rootzén for accepting to be co-examiner of this thesis. I recall fondly the very stimulating conference on extremes in Gothenburg in this summer which was organised by him. Furthermore, I appreciate very much his continuing interest in my work.

I take pleasure in thanking Prof. Charles Goldie for many inspiring discussions, helpful comments and his splendid hospitality in Brighton. The week I spent at Sussex University was fruitful as well as entertaining. I express my gratitude to Prof. Michael Sørensen for many conversations on diffusion processes from which I learnt a lot. My special thanks go also to Prof. Thomas Mikosch who provided me much insight during many discussions. I thank Prof. Hans-Jürgen Schuh for various useful comments on Markov processes. I am also grateful to Dr. Tina Rydberg for a simulation program for hyperbolic diffusions and Dr. Alex McNeil for an Splus program which estimates the extremal index by the blocks method. I am very much indebted to Dr. Ralf Korn for various interesting discussions on stochastic processes in finance and for his constant interest and support throughout the last years. In particular, when we shared an office at the Johannes Gutenberg-University in Mainz, he always lent me an ear for my problems and stimulated my work. Last but not least I would like to thank my friends and colleagues at the Johannes Gutenberg-University in Mainz and the University of Technology in Munich whose company, advice and all-round help supported this work in many ways.

Finally, I thank my parents for their steadfast belief in my ability and their love, understanding and support.

Contents

Abstract

In this thesis we investigate the extremal behaviour of some well-known stochastic models in finance. We consider discrete-time as well as continuous-time models. The thesis is therefore divided into two parts:

In a first part we study the class of autoregressive processes with ARCH(1) errors given by the stochastic difference equation

$$X_n = \alpha X_{n-1} + \sqrt{\beta + \lambda X_{n-1}^2}\,\varepsilon_n\,, \quad n \in \mathbb{N},$$

where $(\varepsilon_n)_{n\in\mathbb{N}}$ are i.i.d. random variables. Under general and tractable assumptions we show the existence and uniqueness of a stationary distribution. We prove that the stationary distribution has a Pareto-like tail with a well-specified tail index depending on α and λ. This thesis generalises results for the ARCH(1) process (the case $\alpha = 0$) proved by Kesten (1973), Vervaat (1979) and Goldie (1991). However, we present a different method of proof invoking a Tauberian theorem. We apply these results in order to investigate the extremal behaviour of the autoregressive processes with ARCH(1) errors.

The extremes of such processes occur typically in clusters. We give an explicit formula for the extremal index and the probabilities for the length of a cluster. Autoregressive processes with ARCH(1) errors are used for financial data, in particular for exchange rates.

In a second part we investigate the extremes of diffusion processes $(X_t)_{t\in\mathbb{R}}$ given by stochastic differential equations of the form

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t\,, \quad t > 0\,, \quad X_0 = x\,,$$

where W denotes the standard Brownian motion, μ is the drift term and σ is the diffusion coefficient. Under some appropriate conditions on $(X_t)_{t\in\mathbb{R}}$ we prove that the point process of

ε–upcrossings converges in distribution to a homogeneous Poisson process. We apply our results to term structure models or asset price processes such as the Vasicek model, the Cox-Ingersoll-Ross model and the generalised hyperbolic diffusion. We also show how to construct a diffusion with pre-determined stationary density which captures any extremal behaviour. As an example we introduce a new model, the generalised inverse Gaussian diffusion.

Chapter 1

Introduction

1.1 Extreme Value Theory - from Hydrology to Finance

Extreme value theory - the theory of weak convergence of maxima and/or upcrossings of stochastic processes over a high threshold - is a very elegant and fascinating mathematical theory as well as a toolbox which can be applied to a broad class of problems in many different areas (the work of Gumbel (1958) is still an excellent reference in this context).
The following examples are situations where extreme value theory typically enters.

- Premium volumes of insurance companies have to be calculated in order to cover, with sufficiently high probability, future losses.
- Dams or dikes at locations along a river or sea must be built high enough to exceed the maximum water height.
- Sports records (athletics, cycling, skating etc.) are frequently broken.
- The detection of air-quality standards is often formulated in terms of the highest level of permitted emissions.
- Mechanical components of trains, airplanes etc. must be sufficiently strong and flexible to withstand strong forces.

All these examples have in common that they concern questions about extremal behaviour. A typical procedure is to model the observed data and to make decisions on the basis of a probabilistic model of the extreme values of the data set.

Classical extreme value theory - the extreme value theory for independent, identically distributed (i.i.d.) random variables - has its roots in the late twenties of this century. More precisely, it starts with the famous paper by Fisher and Tippett (1928). Their central result, often referred to as the *Extremal Types Theorem* and later proved in complete generality by Gnedenko (1943), is the specification of the form of a non-degenerate limit distribution for centred and normalised maxima. In fact, there are only three possible types of extreme value distributions: the Gumbel law Λ, the Fréchet law Φ_α and the Weibull law Ψ_α. In other words: if F is the underlying distribution function of the random variables then

$$\lim_{n\to\infty} F^n(a_n x + b_n) = Q(x)\,, \quad x \in \mathbb{R}\,, \tag{1.2}$$

where $a_n > 0, b_n \in \mathbb{R}$ are properly chosen constants and Q is either Λ, Φ_α or Ψ_α. For the precise form of the extreme value distributions see Appendix A1.

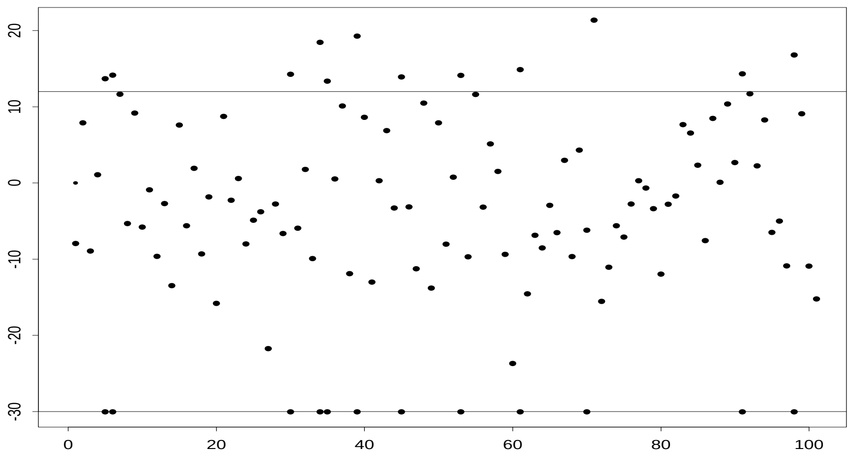

Figure 1.1: Sequence of i.i.d. random variables and associated point process of exceedances above the level 12. Exceedances are indicated at the horizontal line through -30. There is no clustering visible.

Furthermore, it is not necessary to know the detailed nature of the distribution function F in order to know which limit distribution occurs. The extreme value distribution is determined by the behaviour of the tail of $F(x)$ for large x. More precisely: (1.2) is equivalent to

$$\lim_{n\to\infty} n\,(1 - F(a_n x + b_n)) = -\ln Q(x)\,, \quad x \in \mathbb{R}\,. \tag{1.3}$$

Indeed, for $x \in \mathbb{R}$, define $S_n = \sum_{i=1}^n 1_{\{X_i > a_n x + b_n\}}$ as the number of exceedances of $a_n x + b_n$ by $X_1, ..., X_n$. As usual 1_A denotes the indicator function of the set A. Then S_n is a binomial random variable with parameters $(n, 1-F(a_n x+b_n))$. An application of Poisson's limit theorem yields $S_n \xrightarrow{d} Poi(-\ln Q(x))$, $n \to \infty$, if and only if (1.3) holds. This asymptotic Poisson property may be generalised by considering the point process N_n of exceedances of the level $a_n x + b_n$. Actually, if $N_n(\cdot) = \#\{\text{exceedances of } a_n x + b_n \text{ by } (X_i)_{1\le i\le n}\colon i/n \in \cdot\}$, $n \in \mathbb{N}$, then

$$N_n(\cdot) \xrightarrow{w} N(\cdot)\,, \quad n \to \infty\,, \tag{1.4}$$

where N is a homogeneous Poisson process with intensity $-\ln Q(x)$ and $\xrightarrow{w}$ denotes the weak convergence. In particular, the exceedances are simple (see Figure 1.1).

Since Fisher and Tippett (1928), extreme value theory has passed through an exciting theoretical development. Extending the classical results, a satisfying general theory has been developed which includes extreme value theory of dependent random sequences as well as extreme value theory of continuous parameter processes. For an introduction to classical extreme value theory we refer to Leadbetter, Lindgren and Rootzén (1983), Resnick (1987) or Embrechts, Klüppelberg and Mikosch (1997, Chapter 3).

The extremal behaviour of discrete-time stationary processes is quite well-understood. In contrast to the independent case, extremes of dependent sequences may cluster (see Figure 1.2). Suppose that we have a stationary sequence $(X_n)_{n\ge 0}$ with marginal distribution function F which satisfies some weak dependence assumptions (strong mixing or the weaker assumption $D(u_n)$ of Leadbetter (1983)). Let $M_n = \max_{1\le j\le n} X_j$ and let $a_n x + b_n$ be the sequence such that (1.2) holds and such that $\lim_{n\to\infty} P(M_n \le a_n x + b_n)$ exists. Then, there exists a constant $\theta \in [0,1]$ such that

$$\lim_{n\to\infty} P(M_n \le a_n x + b_n) = Q^{\theta}(x)\,, \quad x \in \mathbb{R}\,. \tag{1.5}$$

θ is called the *extremal index* of the sequence $(X_n)_{n\ge 0}$. This concept, originated by Newell (1964), Loynes (1965) and O'Brien (1974), was taken first as a definition by Leadbetter (1983). Since Q is computable from knowledge of the marginal distribution F, it turns out that θ is the key parameter for extending extreme value theory for i.i.d. random variables to stationary processes.

The point process convergence of exceedances is a bit more complicated than in the independent case: for fixed $x \in \text{supp}(Q)$, let N_n be again the point process of exceedances, i.e. $N_n(\cdot) = \#\{\text{exceedances of } a_n x + b_n \text{ by } (X_i)_{1\le i\le n}\colon i/n \in \cdot\}$. Let $n = r\,k$ and $I_i = ((i-1)/k, i/k]$, $i = 1, ..., k$ be a partition of $(0,1]$ into k blocks. For a suitable choice of k and r see for instance Embrechts, Klüppelberg and Mikosch (1997), Section 8.1. If $N_n(I_i) > 0$ a cluster of exceedances occurs at block i and then $\pi_n(j) = P(N_n(I_i) = j \mid N_n(I_i) > 0) = P(N_n(I_1) = j \mid N_n(I_1) > 0)$, $j \in \mathbb{N}$, is

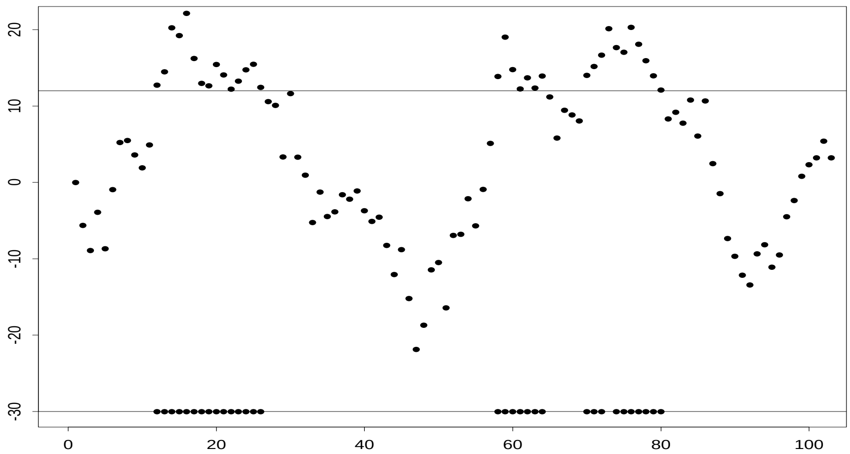

Figure 1.2: Sequence of dependent random variables and associated point process of exceedances. We have a strong clustering indicated at the horizontal line through -30.

the probability of cluster length j in the interval I_i for any $j \in \mathbb{N}$. Hsing, Hüsler and Leadbetter (1988) showed that under (1.2), (1.5) and if $\pi_n(j) \to \pi(j)$, $n \to \infty$ for any $j \in \mathbb{N}$, where $(\pi(j))_{j\in\mathbb{N}}$ is a probability distribution, then

$$N_n(\cdot) \stackrel{w}{\to} N(\cdot)\,, \quad n \to \infty\,, \tag{1.6}$$

where N is a compound Poisson process with intensity $-\theta \ln Q(x)$ and jump probabilities $(\pi(j))_{j\in\mathbb{N}}$. Moreover, under additional conditions on $\pi_n(j)$ (see Smith 1988),

$$\sum_{j=1}^{\infty} j\pi(j) = 1/\theta\,.$$

Thus, a natural interpretation of θ is that of the reciprocal of mean cluster size. In the i.i.d. case, extremes are simple and hence $\theta = 1/\theta = 1$.

For a number of special dependent sequences, the extremal behaviour has been studied in more detail. Extreme value theory for linear time series (moving average processes, autoregressive processes etc.) with heavy tailed innovations has been studied in a series of papers by Davis and Resnick, we refer to Resnick (1987), Chapter 4.5. A further reference is the book by Leadbetter, Lindgren and Rootzén (1983). Rootzén (1988) and Leadbetter, Rootzén (1988) and Asmussen (1998) have investigated regenerative and Markov sequences. Extremes of the ARCH(1) process have been studied by de Haan, Resnick, Rootzén and de Vries (1989), Perfekt (1994) and Hooghiemstra and Meester (1995).

The theory of extremes for continuous parameter processes started with the work of Rice (1939) for mean square differentiable normal processes leading to a celebrate formula for the mean number of upcrossings per unit. Briefly, for a constant u the process (X_t) has an *upcrossing* at t_0 if for some $\varepsilon > 0, X_t \leq u$ in $(t_0 - \varepsilon, t_0)$ and $X_t \geq u$ in $(t_0, t_0 + \varepsilon)$. If $N_u((0,1))$ denotes the number of upcrossings of u by (X_t) in the interval $(0,1)$ and if $\mu(u) = EN_u((0,1)) < \infty$, then the upcrossings form a stationary point process N_u with intensity $\mu(u)$. Rice could show that for mean square differentiable Gaussian processes $\mu(u) = \sqrt{C/2}\pi^{-1}\exp(-u^2/2)$, where the constant C comes from (1.7). The results of Rice were extended in various ways. Two review papers on the theory of extremes of Gaussian processes and related problems are Leadbetter, Lindgren and Rootzén (1983) and Leadbetter and Rootzén (1988). Further work on Gaussian processes can be found in Adler (1990) and Berman (1992). In particular, under certain restrictions there exist corresponding results to (1.2) and (1.4) in the continuous Gaussian case. For a standardised, continuous stationary normal process (X_t) with convariance function $r(t)$ such that Berman's condition

$$r(t)\ln(t) \to 0\,, \quad t \to \infty\,,$$

holds and

$$r(t) = 1 - C\,|t|^\alpha + o(|t|^\alpha)\,, \quad t \to 0\,, \tag{1.7}$$

for some $\alpha \in (0,2]$ is satisfied, then

$$P(\max_{0<u\leq t} X_u \leq a_t x + b_t) \to \exp(-e^{-x})\,, \quad t \to \infty\,,$$

for certain known deterministic functions $a_t > 0$ and $b_t \in \mathbb{R}$. Furthermore, we have

$$\left.\begin{array}{ll} \text{for } \alpha = 2: & N_{a_t x + b_t}(t\,\cdot) \\ \text{for } \alpha < 2: & N_{\varepsilon, a_t x + b_t}(t\,\cdot) \end{array}\right\} \xrightarrow{w} N(\,\cdot\,)\,, \quad t \to \infty\,, \tag{1.8}$$

where $N_{a_t x+b_t}(t\cdot) = \#\{\text{upcrossings of } a_t x + b_t \text{ by } (X_s)_{0\leq s\leq t} : s/t \in \cdot\}$, $N_{\varepsilon,a_t x+b_t}(t\cdot) = \#\{\varepsilon\text{-upcrossings of } a_t x + b_t \text{ by } (X_s)_{0\leq s\leq t} : s/t \in \cdot\}$, N is a homogeneous Poisson process with intensity e^{-x} and $\xrightarrow{w}$ denotes weak convergence. The notion of ε-upcrossings is discussed in detail in Chapter 3 and is needed since normal processes with $\alpha < 2$ are not mean square differentiable and hence the mean number of upcrossings need not be finite. An ε-upcrossing is always an upcrossing while the converse does not hold.

The theory of Gaussian processes was generalised to a broader class of continuous parameter processes by Leadbetter and Rootzén (1982) and Leadbetter, Lindgren and Rootzén (1983). Their results apply in particular to two cases: first to normal processes and then to stationary processes with finite upcrossing intensities.

Newell (1962), Berman (1964), Mandl (1968) and Davis (1982) investigate extremes of diffusion processes. Proposition 3.2.1 in Chapter 3 of this thesis quotes an important result of their work: the limit relation (1.2) also holds for diffusion processes under certain assumptions whereas the point process convergence of upcrossings is not known in the literature. We present the analogue of (1.4) for diffusion processes in Theorem 3.2.4 in Chapter 3.

This thesis aims at applications in finance and econometrics: loosely speaking, we consider financial time series models of changing variance and covariance over time. These models are often denoted as volatility models. There are numerous volatility models used for financial instruments. A logical conceptual division of such models results into continuos-time and discrete-time models.

Continuous-time volatility models are natural models for physicists and mathematicians with an analytic background from stochastic analysis. They are typically given by a stochastic differential equation of the form

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t\,, \quad t > 0\,, \quad X_0 = x\,, \tag{1.9}$$

where W is standard Brownian motion, μ is the *drift term* and σ is the *diffusion coefficient or volatility*. Stochastic processes which are defined by (1.9) are diffusion processes. They are often used as models for interest rates or price processes. As a first approximation, the statistical fit of diffusion processes to data may be reasonable. Choosing the volatility σ in (1.9) appropriately provides a large variety of models ranging from Gaussian processes to models which capture large fluctuations in real data (see Chapter 3). However, there are also various arguments against such models. These include that real world processes are not continuous in time. Furthermore, diffusions with arbitrary volatility are in general untractable with respect to computations. Numerical methods can then be helpful to solve such problems.

Discrete-time versions of stochastic volatility models have their roots in time series analysis and econometrics. They are usually referred to as conditional heteroskedastic models. The simplest examples of such models can be written by random recurrence equations of the form

$$X_n = \mu_n + \sigma_n\, \varepsilon_n\,, \quad n \in \mathbb{N}, \tag{1.10}$$

where ε_n are i.i.d. innovations with mean zero, μ_n is the conditional expectation of X_n (which may or may not depend on n) and the volatility σ_n describes the change of (conditional) variance. Empirical work has confirmed that such models fit many types of financial data (log-returns, exchange rates, etc.). The following empirical stylised facts of financial data can be modelled by such discrete-time stochastic volatility models:

- heavy-tailedness
- clustering in the extremes
- large fluctuations
- data are uncorrelated, but not independent

Processes which are most popular in econometrics and which belong to the class (1.10) are ARCH (autoregressive conditional heteroskedastic) models and GARCH (generalised ARCH) models.

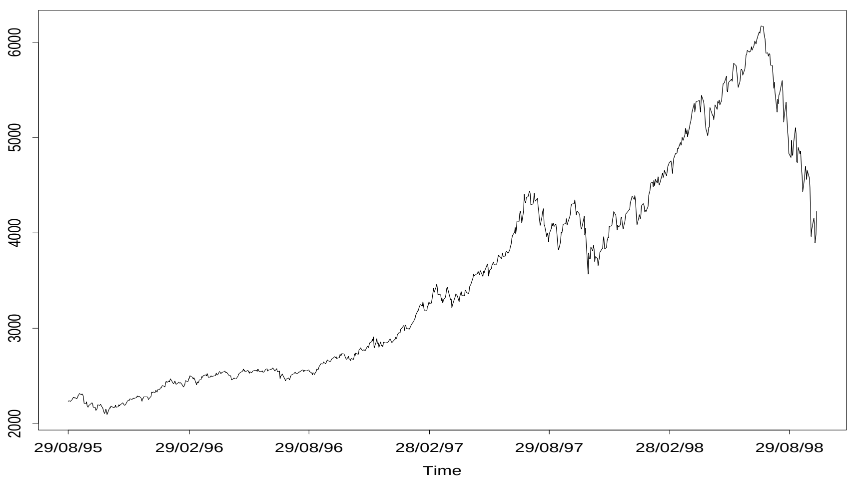

Figure 1.3: DAX (the German stock market index) closing prices during 29/8/95-9/10/98.

In this thesis we investigate the extremal behaviour of stochastic volatility models which are either time-discrete or time-continuous. Whereas extreme value theory has a long history of applications in engineering, climatology and in particular hydrology, it has only more recently entered into the financial world (see e.g. Embrechts, Klüppelberg and Mikosch (1997)). There is a growing interest in this subject in the insurance and banking world. Catastrophic losses have been rising during the last decade for the reinsurance industry. 1990 (the year of the winter storms Daria and Vivian) and 1992 (the year of hurricane Andrew) have caused an extremely high damage and put significant financial demands on society, for details on these, see Sigma (1995)-(1998).

Within the finance context, extremal events can be observed whenever stock market crashes like the one this year occur (see Figures 1.3 and 1.4). Other examples are the losses within the realm of derivatives such as the collapse of Barings Bank or the losses of the Metallgesellschaft.

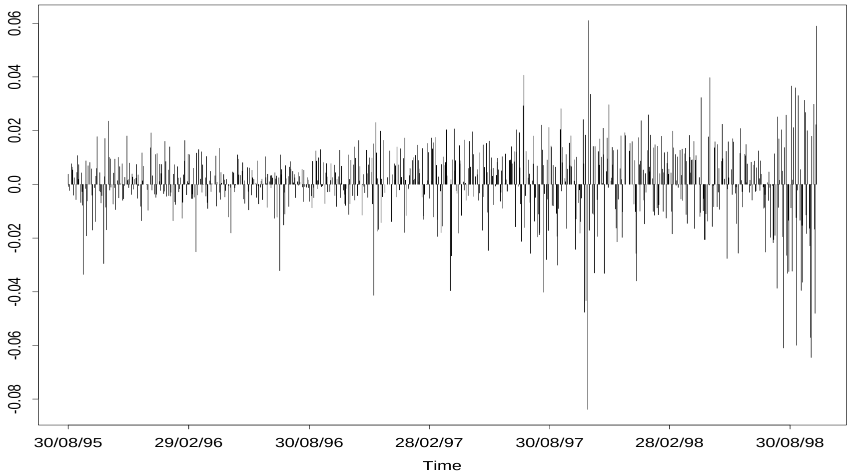

Figure 1.4: Log-returns of the DAX closing prices during 29/8/95-9/10/98.

Due to the increase over the recent past in both frequency as well as size of casualities, risk management has become a key issue in any financial institution or corporation. Loosely speaking, risk management needs precise estimates of tail probabilities and quantiles of profit-loss distributions , and indeed of general financial data. Extreme value theory yields methods for quantifying events of great losses and their consequences in a statistically optimal way. It gives the best estimates of extremal events and represents the most honest approach to measuring the uncertainity inherent in the problem. Interesting papers from a practical point of view on such problems are Embrechts, Resnick and Samorodnitsky (1997), McNeil (1998) and Emmer, Klüppelberg and Trüstedt (1998).

In order to quantify the risk of financial products a theory of the extremal behaviour of stochastic processes used in finance is required. This thesis is a contribution to the mathematical problems in this area. It lays the foundation for practical applications in finance, in particular the increasingly important area of risk management.

1.2 An outline of the thesis

As indicated in Section 1.1 and in the table of contents this thesis consists of two main parts. In Chapter 2 we consider a class of autoregressive (AR) processes with ARCH(1) errors given by the random recurrence equation

$$X_n = \alpha X_{n-1} + \sqrt{\beta + \lambda X_{n-1}^2}\varepsilon_n\,, \quad n \in \mathbb{N}\,, \tag{2.11}$$

where the innovations $(\varepsilon_n)_{n\in\mathbb{N}}$ are i.i.d. random variables and $\alpha \in \mathbb{R}$, $\beta, \lambda > 0$. As mentioned before, processes defined by such a stochastic difference equation are suitable models for log-returns of stock prices and exchange rates because of their non-constant volatilities (see also for instance Duan (1996)).

In Section 2.1 we investigate this model in detail. We introduce some assumptions on the innovations $(\varepsilon_n)_{n\in\mathbb{N}}$: the general conditions guarantee the existence and uniqueness of a stationary version of $(X_n)_{n\in\mathbb{N}}$ whereas $(D.1) - (D.3)$ allow us to describe the tail behaviour of the stationary distribution and the extremal behaviour of $(X_n)_{n\in\mathbb{N}}$. The normal distribution, for instance, satisfies all these assumptions.

In Section 2.2 we determine the parameter set of stationarity for model (2.11) and the tail of the stationary distribution. Theorem 2.2.3 collects some probabilistic properties of $(X_n)_{n\in\mathbb{N}}$, in particular the existence and uniqueness of a stationary distribution. The results are an extension of the results of Diebolt and Guégan (1990) and Maercker (1997). The main result of Section 2.2 is given in Theorem 2.2.11. Under the general conditions and $(D.1) - (D.3)$ the tail of the stationary distribution of $(X_n)_{n\in\mathbb{N}}$, which is the distribution function of a random variable X, behaves asymptotically like

$$P(X > x) \sim c\,x^{-\kappa}\,, \quad x \to \infty\,, \tag{2.12}$$

where $c = c(\alpha, \beta, \kappa, \varepsilon)$ and $\kappa = \kappa(\alpha, \lambda, \varepsilon)$ are well-specified constants depending on $\alpha, \beta, \kappa, \varepsilon$ and $\alpha, \lambda, \varepsilon$, respectively, and ε is a generic random variable with the same distribution as the innovations $(\varepsilon_n)_{n\in\mathbb{N}}$. For $\alpha = 0$ the asymptotic relation (2.12) coincides with the corresponding result in Goldie (1991). We extend his result to the larger class of processes (2.11) to which his idea of proof does not apply. Our method of proof uses the Tauberian theorem of Drasin and Shea which was proven first in Jordan (1974). The theorem takes its name from Drasin and Shea (1976). Loosely speaking, Tauberian theory draws conclusions from the asymptotic behaviour of some transform to the asymptotic behaviour of a kernel density or distribution tail (see Bingham, Goldie, Teugels (1987)).

In Section 2.3 we investigate the extremal behaviour of the AR(1) process with ARCH(1) errors (2.11). In order to do this we investigate the related process $(Z_n)_{n\in\mathbb{N}} = (\ln(X_n^2))_{n\in\mathbb{N}}$. The process $(Z_n)_{n\in\mathbb{N}}$ is crucial for the study of the extremal behaviour of $(X_n)_{n\in\mathbb{N}}$. We show in Lemma 2.3.1 that $(Z_n)_{n\in\mathbb{N}}$ behaves above a high threshold asymptotically like a random walk with negative drift, which can be completely specified. Subsection 2.3.2 contains the main results (Theorem 2.3.5) concerning the extremal behaviour of $(X_n)_{n\in\mathbb{N}}$. An explicit formula for the extremal index is given and the probability distribution for the length of a cluster is calculated. We interpret these results and present some simulations. The proof of our results invokes the work of Perfekt (1995) where the extremal behaviour of real-valued, stationary Markov chains is studied under certain assumptions.

In Chapter 3 we investigate the extremal behaviour of diffusion processes which are given by stochastic differential equations of the form (1.9) with $\mu(t, X_t) = \mu(X_t)$ and $\sigma(t, X_t) = \sigma(X_t)$, i.e. we consider always homogeneous diffusion processes which can be completely characterised by their associated scale function and speed measure. Although diffusion processes are idealised models for financial data (continuous trading is not possible in the real world), some of them capture quite well empirical observations in real data.

In Section 3.1 we present the framework for the results about the extremal behaviour of diffusion processes to follow. We shall require certain properties of the speed measure and scale function of $(X_t)_{t\geq 0}$, which we explain and summarise in the so-called usual conditions. They guarantee in particular that the diffusion process $(X_t)_{t\geq 0}$ is ergodic and has inaccessible boundaries .

Section 3.2 presents some results on extreme value theory for diffusion processes which were already mentioned. We show that, provided the properly normalised maxima M_t^X of a diffusion process up to time t have a weak non-degenerate limit as $t \to \infty$, then, under weak additional conditions, the point processes of ε-upcrossings converge to a homogeneous Poisson process (Theorem 3.2.4). This result is comparable to (1.8) for $\alpha < 2$ in the Gaussian case. Furthermore, we derive the limit distribution of M_t^X (suitably normalised) under simple conditions on the drift term and the diffusion coefficient (Theorem 3.2.7). Finally, we show how to construct a diffusion with pre-determined stationary density which captures any extremal behaviour (Theorem 3.2.8).

In Section 3.3 we apply these results in order to derive the extremal behaviour of such diffusions as the Vasicek model, the Cox-Ingersoll-Ross (CIR) model, including a generalised version, and the generalised hyperbolic diffusion. They are all standard models in finance. Depending on the choice of parameters the generalised CIR model allows for large fluctuations in the data. This is captured by the limit distribution of M_t^X and the intensity of the limit point process of ε-upcrossings.

In Section 3.4 we present a new model, the generalised inverse Gaussian diffusion, which is constructed with the pre-determined generalised inverse Gaussian stationary density and a pre-determined diffusion coefficient. If we choose the diffusion coefficient as in the CIR model we obtain a further generalisation of this important model. Whereas in Section 3.3 we mainly present results without explicit calculations, for this new model we derive certain quantities in detail.

The Appendix is made up of four different parts, A1-A4. Appendix A1 is concerned with classical extreme value theory. We describe the maximum domains of attraction of the Fréchet and Gumbel distribution and present how to compute the centring and normalising constants. In Appendix A2 we derive the normalising constants $a_t > 0$ and $b_t \in \mathbb{R}$ for the maxima of the Vasicek diffusion and the generalised Cox-Ingersoll-Ross diffusion for $1/2 < \gamma < 1$. Finally, some additional extreme value theory for Markov chains and some general Markov chain theory are provided in Appendix A3 and A4, respectively.

Chapter 2

The Autoregressive Process with ARCH(1) Errors

Recently there has been considerable interest in nonlinear time series models (see e.g. Priestley (1988), Tong (1990), Taylor (1995)). Many of these models were introduced to allow the conditional variance (conditional heteroskedasticity) of a time series model to depend on past information. It has turned out that such models fit very well to many types of financial data. Empirical work (see e.g. Mandelbrot (1963), Fama (1965)) has shown that large changes in equity returns and exchange rates, with high sampling frequency, tend to be followed by large changes setting down after some time to a more normal behaviour. This observation leads to models of the form

$$X_n = \sigma_n \, \varepsilon_n \,, \quad n \in \mathbb{N}, \tag{0.1}$$

where the innovations $(\varepsilon_n)_{n\in\mathbb{N}}$ are i.i.d. symmetric random variables with mean zero, and the volatility σ_n describes the change of (conditional) variance.
One of the specifications of (0.1) are the autoregressive conditionally heteroskedastic (ARCH) models where the conditional variance σ_n^2 is a linear function of the squared past observations. ARCH(p) models introduced by Engle (1982) are defined by

$$\sigma_n^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j X_{n-j}^2 \,, \quad \alpha_0 > 0, \, \alpha_1, ..., \alpha_p \geq, \, \alpha_p > 0, \; n \in \mathbb{N}, \tag{0.2}$$

where p is the order of the ARCH process.
In a series of papers, the ARCH model has been analysed, generalised and used to test for time-varying risk premia in the financial market. We refer for instance to the survey article by

Bollerslev, Chou and Kroner (1992) and the statistical review paper by Shephard (1996). The most famous generalisation to so-called generalised ARCH (GARCH) processes was proposed in Bollerslev (1986). The volatility σ_n is now a linear function in X_{n-1}^2, $X_{n-2}^2, \ldots$ and σ_{n-1}^2, $\sigma_{n-2}^2, \ldots$. ARCH and GARCH models are widely used to model financial time series since they capture exactly the empirical observation in financial data, namely the tendency for volatility clustering and the fact that unconditional price and return distributions tend to have fatter tails than the normal distribution.

Another extension are the class of autoregressive (AR) models with ARCH errors introduced by Weiss (1984). These models are also called SETAR-ARCH models (self-exciting autoregressive). They are defined by

$$X_n = f(X_{n-1}, \ldots, X_{n-k}) + \sigma_n \, \varepsilon_n, \quad n \geq k \,, \tag{0.3}$$

where f is again a linear function in its arguments and σ_n is given by (0.2). This model combines the advantages of an AR model which targets more on the conditional mean of X_n given the past and an ARCH model which concentrates on the conditional variance of X_n (given the past).

The class of models defined by (0.3) embodies various nonlinear models. In this chapter we focus on the AR(1) process with ARCH(1) errors, i.e. $f(X_{n-1}, \ldots, X_{n-k}) = \alpha \, X_{n-1}$ for some $\alpha \in \mathbb{R}$ and σ_n is given in (0.2) with $p = 1$. Note that in the special case $\alpha = 0$ we get just the ARCH(1) model of Engle (1982). This Markovian model is analytically tractable and may serve as a prototype for the larger class of models (0.3).

The purpose of this work is to investigate the tail of the stationary distribution of the AR(1) process with ARCH(1) errors $(X_n)_{n \in \mathbb{N}}$. The model has also been considered for instance by Diebolt and Guégan (1990) and Maercker (1997). For $\lambda = 0$ the process is an AR(1) process whose stationary distribution is determined by the innovations $(\varepsilon_n)_{n \in \mathbb{N}}$, for ε_n normal it is a Gaussian process. In the ARCH(1) case (the case when $\alpha = 0$) the tail is known (see e.g. Goldie (1991) or Embrechts, Klüppelberg, Mikosch (1997)). The result was obtained by considering the squared ARCH(1) process which leads to a stochastic difference equation which fits in the setting of Kesten (1973) and Vervaat (1979). This approach is, however, in general not possible or at least not obvious for $\alpha \neq 0$. Nevertheless for ε_n normal, provided a stationary distribution exists, a characteristic function argument transforms the model such that the results by Kesten (1973), Vervaat (1979) and Goldie (1991) may be applied. We refer to Remark 2.2.21 for further details. For the general case we present another technique using the Drasin-Shea Tauberian theorem which can be found for instance in Bingham, Goldie, Teugels (1987). This method may also be applied to other models given by a stochastic difference equation but falling out of the framework of Kesten (1973), Vervaat (1979) and Goldie (1991). In section 2.3 we investigate the extremal behaviour of the AR(1) process with ARCH(1) errors $(X_n)_{n\in\mathbb{N}}$ extending the work by de Haan, Resnick, Rootzén and de Vries (1989) .

2.1 Assumptions on the model

In this section we present the model and introduce the required assumptions on the innovations $(\varepsilon_n)_{n\in\mathbb{N}}$. They are assumed to hold from now on if it is not stated otherwise.
We consider throughout this chapter an autoregressive model of order 1 with autoregressive conditionally heteroskedastic errors of order 1 (AR(1) model with ARCH(1) errors) which is defined by the stochastic difference equation

$$X_n = \alpha X_{n-1} + \sqrt{\beta + \lambda X_{n-1}^2}\varepsilon_n \,, \quad n \in \mathbb{N}, \tag{1.1}$$

where $(\varepsilon_n)_{n\in\mathbb{N}}$ are i.i.d. random variables with mean zero , $\alpha \in \mathbb{R}$, β, $\lambda > 0$ and X_0 independent of $(\varepsilon_n)_{n\in\mathbb{N}}$. Let ε be a generic random variable with the same distribution as ε_n. Throughout this chapter, we assume that the following *general conditions* for ε are in force:

$$\begin{aligned} \varepsilon \text{ has full support } \mathbb{R}\,, \\ \varepsilon \text{ is symmetric with continuous Lebesgue density } p\,, \\ \text{the second moment of } \varepsilon \text{ exists}\,. \end{aligned} \tag{1.2}$$

Note that the process is evidently a homogeneous Markov chain with state space $\mathbb{R}$ equipped with the Borel σ-algebra. The transition kernel density is given by

$$P(X_1 \in dy \mid X_0 = x) = \frac{1}{\sqrt{\beta + \lambda x^2}}\, p(\frac{y - \alpha x}{\sqrt{\beta + \lambda x^2}})dy\,, \quad x \in \mathbb{R}. \tag{1.3}$$

Under appropriate conditions on α and λ, Theorem 2.2.3 in Section 2 guarantees the existence and uniqueness of a stationary distribution π of $(X_n)_{n\in\mathbb{N}}$. In the following F denotes the distribution function of π and X is a random variable with distribution function F. From the stochastic difference equation (1.1) it is straightforward that X satisfies the fixpoint equation

$$X \stackrel{d}{=} \alpha X + \sqrt{\beta + \lambda X^2}\,\varepsilon\,, \tag{1.4}$$

where ε is a random variable with probability density p, independent of X. In order to determine the tail of the stationary distribution function F we need some additional technical assumptions on p:

(D.1) $p(x) \geq p(x')$ for any $0 \leq x < x'$.

(D.2) For any $c \geq 0$ there exists a constant $q = q(c) \in (0,1)$ and functions $f_+(c,\cdot)$, $f_-(c,\cdot)$ with $f_+(c,x), f_-(c,x) \to 1$ as $x \to \infty$ such that for any $x > 0$ and $t > x^q$

$$p(\frac{x+c+\alpha t}{\sqrt{\beta+\lambda t^2}}) \geq p(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}})\, f_+(c,x)\,,$$
$$p(\frac{x+c-\alpha t}{\sqrt{\beta+\lambda t^2}}) \geq p(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}})\, f_-(c,x)\,.$$

(D.3) There exists a constant $\eta > 0$ such that

$$p(x) = o(x^{-(N+1+\eta+3q)/(1-q)})\,, \quad \text{as } x \to \infty\,,$$

where $N := \inf\{u \geq 0 \,|\, E(|\sqrt{\lambda}\varepsilon|^u) > 2\}$ and q is the constant in $(D.2)$.

The general conditions (1.2) and assumption $(D.1)$ are fairly general and can be checked easily, wheras $(D.2) - (D.3)$ seem to be quite technical and untractable. Nevertheless, numerous densities satisfy these assumptions, one being the normal (see Example 2.2.13).

2.2 The tail of the stationary distribution of an AR(1) process with ARCH(1) errors

In this section we want to determine the parameter set of stationarity for our model and the tail of the stationary distribution. In Theorem 2.2.3 we summarize some probabilistic properties of $(X_n)_{n\in\mathbb{N}}$, in particular the existence and uniqueness of a stationary distribution. Theorem 2.2.11 is the main theorem in this section. We show that the stationary distribution has a Pareto-like tail with a well-specified tail index . For $\alpha = 0$ our result coincides with the corresponding result in Goldie (1991) whereas for $\alpha \neq 0$ the tail index is determined by the autoregressive coefficient α, the ARCH(1) parameter λ and the distribution function of the innovations $(\varepsilon_n)_{n\in\mathbb{N}}$. The proof of this result will be an application of a modification of the Drasin-Shea Tauberian theorem.

2.2.1 Existence and uniqueness of a stationary distribution

In order to determine the tail of the stationary distribution we need some properties of the process $(X_n)_{n\in\mathbb{N}}$. They are summarised in Theorem 2.2.3. In particular, the geometric ergodicity guarantees the existence and uniqueness of a stationary distribution. A short introduction to

Markov chain terminology and the proof of Theorem 2.2.3 is given in the Appendix A4. For further details we refer to Tweedie (1976) or Meyn and Tweedie (1993).

Proposition 2.2.1 *Let ε be a random variable with probability density p satisfying our general conditions. Define $h_{\alpha,\lambda} : [0,\infty) \to [0,\infty]$ for $\alpha \in \mathbb{R}, \lambda > 0$ by*

$$h_{\alpha,\lambda}(u) := E(|\alpha + \sqrt{\lambda}\,\varepsilon|^u)\,, \quad u \geq 0\,. \tag{2.1}$$

(a) The function $h_{\alpha,\lambda}(\cdot)$ is strictly convex in $[0,T)$, where $T := \inf\{u \geq 0 \,|\, E(|\sqrt{\lambda}\,\varepsilon|^u) = \infty\}$.

(b) If furthermore the parameters α and λ are chosen such that

$$h'_{\alpha,\lambda}(0) = E(\ln|\alpha + \sqrt{\lambda}\,\varepsilon|) < 0\,, \tag{2.2}$$

then there exists a unique solution $\kappa = \kappa(\alpha,\lambda) > 0$ to the equation $h_{\alpha,\lambda}(u) = 1$. Moreover, under $h'_{\alpha,\lambda}(0) < 0$,

$$\kappa(\alpha,\lambda) \begin{cases} > 2 & , \quad \alpha^2 + \lambda E(\varepsilon^2) < 1 \\ = 2 & , \quad \alpha^2 + \lambda E(\varepsilon^2) = 1 \\ < 2 & , \quad \alpha^2 + \lambda E(\varepsilon^2) > 1 \end{cases} . \tag{2.3}$$

Proof. The function $h_{\alpha,\lambda}(\cdot)$ has derivatives of all orders in $[0,T)$. In particular, for $u \in [0,T)$,

$$\begin{aligned} h'_{\alpha,\lambda}(u) &= E(|\alpha + \sqrt{\lambda}\,\varepsilon|^u \ln(|\alpha + \sqrt{\lambda}\,\varepsilon|))\,, \\ h''_{\alpha,\lambda}(u) &= E(|\alpha + \sqrt{\lambda}\,\varepsilon|^u (\ln|\alpha + \sqrt{\lambda}\,\varepsilon|)^2) \;>\; 0\,. \end{aligned} \tag{2.4}$$

Statement (a) follows from (2.4). Because of the symmetry of ε we may assume w.l.o.g. $\alpha \geq 0$ and hence

$$\begin{aligned} h_{\alpha,\lambda}(u) \;&\geq\; E(1_{\{\varepsilon \geq 2/\sqrt{\lambda}\}}|\alpha + \sqrt{\lambda}\,\varepsilon|^u) \geq E(1_{\{\varepsilon \geq 2/\sqrt{\lambda}\}}|\sqrt{\lambda}\,\varepsilon|^u) \\ &=\; 2^u\, E(1_{\{\varepsilon \geq 2/\sqrt{\lambda}\}}) \to \infty\,, \quad u \to \infty\,. \end{aligned}$$

The latter fact, together with $h_{\alpha,\lambda}(0) = 1$, assumption (2.2) and the strict convexity of $h_{\alpha,\lambda}$ implies that there exists a unique solution $\kappa > 0$ such that $h_{\alpha,\lambda}(\kappa) = 1$. Finally,

$$h_{\alpha,\lambda}(2) = \alpha^2 + \lambda\, E(\varepsilon^2)\,,$$

which finishes the proof. □

$\|\alpha\|$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7
λ	(0,3.56]	(0,3.55]	(0,3.52]	(0,3.47]	(0,3.39]	(0,3.30]	(0,3.18]	(0,3.04]
$\|\alpha\|$	0.8	0.9	1	1.1	1.2	1.25	1.27	1.27805
λ	(0,2.87]	(0,2.66]	(0,2.42]	(0.17,2.11]	(0.38,1.69]	(0.58,1.38]	(0.75,1.19]	(0.94,0.96]

Table 2.1: Numerical domain of λ dependent on $|\alpha|$ such that $h'_{\alpha,\lambda}(0) < 0$ in the case $\varepsilon \sim N(0,1)$.

λ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\|\alpha\|$	1.05	1.11	1.16	1.20	1.23	1.25	1.26	1.27	1.28
λ	1	1.1	1.2	1.5	2	2.5	3	3.5	3.56
$\|\alpha\|$	1.28	1.27	1.27	1.23	1.13	0.97	0.72	0.24	0.04

Table 2.2: Numerical supremum of $|\alpha|$ dependent on λ such that $h'_{\alpha,\lambda}(0) < 0$ in the case $\varepsilon \sim N(0,1)$.

Remark 2.2.2 (a) By Jensen's inequality $\alpha^2 + \lambda\, E(\varepsilon^2) < 1$ implies $h'_{\alpha,\lambda}(0) < 0$.

(b) Proposition 2.2.1 holds in particular for a standard normal distributed random variable ε. In this case $T = \infty$.

(c) In general, it is not possible to determine explicitly which parameters α and λ satisfy (2.2). If $\alpha = 0$ and $\varepsilon \sim N(0,1)$ (i.e. in the ARCH(1) case) (2.2) is fulfilled if and only if $\lambda \in (0, 2e^{\gamma})$, where γ is Euler's constant (see Embrechts et al. (1997), Section 8.4). For $\alpha \neq 0$, Tables 2.1 and 2.2 show numerical domains of α and λ for $\varepsilon \sim N(0,1)$.

(d) Note that κ is a function of α and λ. Since ε is symmetric κ does not depend on the sign of α. For $\varepsilon \sim N(0,1)$ we can show: for fixed λ, κ is decreasing in $|\alpha|$. See also Table 3.

Proof. W.l.o.g. $\alpha \geq 0$. Let $\varphi(\cdot \,|\, \mu, \sigma^2)$ denote the normal density with mean μ and variance σ^2. Then, by symmetry of φ,

$$\begin{aligned}
\frac{\partial h_{\alpha,\lambda}(u)}{\partial \alpha} &= \frac{1}{\lambda} \int_{-\infty}^{\infty} |y|^u (y-\alpha)\varphi(y|\alpha,\lambda)dy \\
&= \frac{1}{\lambda}\left(\int_{-\infty}^{0} (-y)^u (y-\alpha)\varphi(y|\alpha,\lambda)dy + \int_{0}^{\infty} y^u (y-\alpha)\varphi(y|\alpha,\lambda)dy \right)
\end{aligned}$$

$\lvert\alpha\rvert$ \ λ	0.2	0.4	0.6	0.8	1.0	1.2	1.5	2.0	2.5	3.0	3.5
0	12.85	6.09	3.82	2.67	1.99	1.54	1.07	0.61	0.33	0.15	0.01
0.2	11.00	5.49	3.52	2.51	1.89	1.46	1.03	0.59	0.32	0.13	0.01
0.4	8.12	4.28	2.87	2.10	1.61	1.26	0.90	0.51	0.27	0.10	-
0.6	5.41	3.03	2.12	1.60	1.25	0.99	0.71	0.39	0.19	0.05	-
0.8	3.00	1.85	1.37	1.07	0.85	0.68	0.48	0.25	0.09	-	-
1.0	0.96	0.83	0.70	0.57	0.47	0.37	0.25	0.09	-	-	-
1.2	-	0.01	0.01	0.01	0.01	0.01	0.01	-	-	-	-

Table 2.3: Numerical solution of $h_{\alpha,\lambda}(\kappa) = 1$ for $\kappa = \kappa(\alpha, \lambda)$ dependent on α and λ in the case $\varepsilon \sim N(0,1)$. For $\alpha = 0$ a similar table can be found in de Haan et al. (1989).

$$= \quad u \int_0^\infty y^{u-1} \left(\varphi(y|\alpha,\lambda) - \varphi(y|-\alpha,\lambda)\right) dy > 0\,, \quad u \geq 0\,,$$

where the last line follows by partial integration with respect to y. We may therefore conclude that, if $\alpha' > \alpha$ then $h_{\alpha,\lambda}(u) < h_{\alpha',\lambda}(u)$ for any λ, u. Assume $\kappa(\alpha) \leq \kappa(\alpha')$. Then we have by Proposition 2.2.1(b) and Hölder's inequality that

$$1 = h_{\alpha,\lambda}(\kappa(\alpha)) < h_{\alpha',\lambda}(\kappa(\alpha)) \leq h_{\alpha',\lambda}(\kappa(\alpha'))^{\kappa(\alpha)/\kappa(\alpha')} = 1\,,$$

which is a contradiction. □

We are now ready to state the following theorem.

Theorem 2.2.3 *Consider the process $(X_n)_{n\in\mathbb{N}}$ in (1.1) with $(\varepsilon_n)_{n\in\mathbb{N}}$ satisfying the general conditions and with parameters α and λ satisfying (2.2). Then the following assertions hold:*

(a) Let ν be the normalised Lebesgue-measure $\nu(\cdot) := \lambda(\cdot \cap [-M,M])/\lambda([-M,M])$. Then $(X_n)_{n\in\mathbb{N}}$ is an aperiodic positive ν-recurrent Harris chain with regeneration set $[-M,M]$ for M large enough.

(b) $(X_n)_{n\in\mathbb{N}}$ is geometric ergodic. In particular, $(X_n)_{n\in\mathbb{N}}$ has a unique stationary distribution and satisfies the strong mixing condition with geometric rate of convergence. The stationary distribution is continuous and symmetric.

(c) If $\alpha^2 + \lambda\, E(\varepsilon^2) < 1$, then the stationary distribution of $(X_n)_{n\in\mathbb{N}}$ has finite second moment.

Remark 2.2.4 (a) Statements (a) and (b) are basically a collection of results of Diebolt and Guégan (1990) and Maercker (1997). They assume $\alpha^2 + \lambda\, E(\varepsilon^2) < 1$ and hence only cover the finite variance case.

(b) When we study the stationary distribution of $(X_n)_{n\in\mathbb{N}}$ we may w.l.o.g. assume that $\alpha \geq 0$. For a justification, consider the process $(\widetilde{X}_n)_{n\in\mathbb{N}} = ((-1)^n X_n)_{n\in\mathbb{N}}$ which satisfies to the stochastic difference equation

$$\widetilde{X}_n = -\alpha \widetilde{X}_{n-1} + \sqrt{\beta + \lambda \widetilde{X}_{n-1}^2}\, \varepsilon_n\,, \quad n \in \mathbb{N},$$

where $(\varepsilon_n)_{n\in\mathbb{N}}$ are the same random variables as in (1.1) and $\widetilde{X}_0 = X_0$. If $\alpha < 0$, because of the symmetry of the stationary distribution we may hence study the new process $(\widetilde{X}_n)_{n\in\mathbb{N}}$.

(c) By statement (c), the assumption $\alpha^2 + \lambda\, E(\varepsilon^2) < 1$ is sufficient for the existence of the second moment. We will see in Remark 2.2.12(c) that it is also necessary.

(d) Theorem 2.2.3 is crucial for investigating the extremal behaviour of $(X_n)_{n\in\mathbb{N}}$. The strong mixing property includes automatically that the sequence $(X_n)_{n\in\mathbb{N}}$ satisfies the conditions $D(u_n)$ and $\Delta(u_n)$. The condition $D(u_n)$ is a frequently used mixing condition due to Leadbetter et al. (1983) whereas the slightly stronger condition $\Delta(u_n)$ was introduced by Hsing (1984). Loosely speaking, $D(u_n)$ and $\Delta(u_n)$ give the "degree of independence" of extremes situated far apart from each other.

Proof. Because of the strict positivity and continuity of the transition density the process $(X_n)_{n\in\mathbb{N}}$ is a ν-irreducible Feller chain. By Feigin and Tweedie (1985), p.3, this implies that every compact set of the state space is small and thus $[-M, M]$ for arbitrary $M > 0$ is small. Finally, by Proposition 5.3 of Tweedie (1976), $[-M, M]$ is a status set for $(X_n)_{n\in\mathbb{N}}$.

(a) Because of Proposition 2.2.1, for $\alpha \in \mathbb{R}$ and $\lambda > 0$ such that $h'_{\alpha,\lambda}(0) < 0$ there exists a $\kappa > 0$ such that $h_{\alpha,\lambda}(u) < 1$ for any $u \in (0, \kappa)$ and $h_{\alpha,\lambda}(0) = h_{\alpha,\lambda}(\kappa) = 1$. Now choose $\eta \in (0, \min(\kappa, 2))$ and $\delta \in (0, 1 - h_{\alpha,\lambda}(\eta))$ arbitrary. For any such η and δ there exists a constant $C = C(\eta, \delta) \in (0, 1)$ such that

$$h_{\alpha,\lambda}(\eta) + \delta \leq 1 - 2\,C\,. \tag{2.5}$$

Define $g(x) := 1 + |x|^\eta \geq 1$ for any $x \in \mathbb{R}$. For M large enough and $|x| > M$ we have by continuity of $h_{\alpha,\lambda}$ in α

$$\left| h_{\alpha x/\sqrt{x^2+\beta/\lambda},\lambda}(\eta) - h_{\alpha,\lambda}(\eta) \right| < \delta \tag{2.6}$$

and

$$C\, g(x) \geq 1 + (h_{\alpha,\lambda}(\eta) \pm \delta)(-1 + O(|x|^{\eta-2})) \tag{2.7}$$

since $\eta < 2$, $h_{\alpha,\lambda}(\eta) - \delta$ is independent of x and g increases to ∞. From (1.3) we obtain for $x \to \infty$

$$\begin{aligned}
\int_{(-\infty,\infty)} g(y) P(X_1 \in dy \,|\, X_0 = x) &= 1 + (\beta + \lambda x^2)^{\eta/2} E(|\frac{\alpha x}{\sqrt{\lambda x^2 + \beta}} + \varepsilon|^\eta) \\
&= 1 + (\frac{\beta}{\lambda} + x^2)^{\eta/2} h_{\alpha x/\sqrt{x^2+\beta/\lambda},\lambda}(\eta) \\
&= 1 + (1 + O(x^{-2}))|x|^\eta\, h_{\alpha x/\sqrt{x^2+\beta/\lambda},\lambda}(\eta) \\
&= 1 + O(|x|^{\eta-2})\, h_{\alpha x/\sqrt{x^2+\beta/\lambda},\lambda}(\eta) + |x|^\eta\, h_{\alpha x/\sqrt{x^2+\beta/\lambda},\lambda}(\eta) \\
&= 1 + \left(-1 + O(|x|^{\eta-2})\right)\, h_{\alpha x/\sqrt{x^2+\beta/\lambda},\lambda}(\eta) + g(x)\, h_{\alpha x/\sqrt{x^2+\beta/\lambda},\lambda}(\eta)\,,
\end{aligned} \tag{2.8}$$

where the third line follows from Taylor expansion. From (2.5)-(2.8), we obtain for any $x \in \mathbb{R}$ with $|x| > M$,

$$\int_{(-\infty,\infty)} g(y) P(X_1 \in dy \,|\, X_0 = x) \leq C\, g(x) + (1 - 2C) g(x) = (1 - C) g(x)\,. \tag{2.9}$$

Define

$$\tau_{[-M,M]} := \inf\{n \geq 1 \,|\, X_n \in [-M, M]\}$$

and let $x \in \mathbb{R}$ be arbitrary. Then we have

$$\begin{aligned}
E(\tau_{[-M,M]} \,|\, X_0 = x) &= E(1_{\{X_1 \in [-M,M]\}} E(\tau_{[-M,M]} | X_1) | X_0 = x) \\
&\qquad + E(1_{\{X_1 \in [-M,M]^c\}} E(\tau_{[-M,M]} | X_1) | X_0 = x) \\
&\leq 1 + E(1_{\{X_1 \in [-M,M]^c\}} E(\tau_{[-M,M]} | X_1) | X_0 = x) \\
&\leq 1 + \int_{[-M,M]^c} E(\tau_{[-M,M]} | X_1 = y) P(X_1 \in dy | X_0 = x)\,.
\end{aligned}$$

By (2.9), Lemma A4.1 holds and we obtain for all $x \in \mathbb{R}$,

$$\begin{aligned}
E(\tau_{[-M,M]} | X_0 = x) &\leq 1 + \int_{[-M,M]^c} \frac{g(y)}{C} P(X_1 \in dy | X_0 = x) \\
&\leq 1 + \frac{1}{C} + E\left(\left| \alpha x + \sqrt{\lambda x^2 + \beta}\varepsilon \right|^\eta \right) < \infty
\end{aligned} \tag{2.10}$$

and thus $[-M, M]$ is Harris recurrent. Since the transition density of $(X_n)_{n\in\mathbb{N}}$ is strictly positive on $[-M, M]$ we know from Asmussen (1987), p.151, that there exists some constant $\widetilde{C} \in (0,1)$ such that

$$P(X_1 \in B \,|\, X_0 = x) \geq \widetilde{C}\, \nu(B) \tag{2.11}$$

for any $x \in [-M, M]$ and any Borel-measurable set B, i.e. $(X_n)_{n\in\mathbb{N}}$ is a Harris chain with regeneration set $[-M, M]$. Finally, by Theorem A4.2, (2.9) and the fact that $[-M, M]$ is a status set, $(X_n)_{n\in\mathbb{N}}$ is positive Harris ν-recurrent.

(b) Note that

$$\sup_{x\in[-M,M]} \int_{\mathbb{R}} g(y) P(X_1 \in dy|X_0 = x) = 1 + \sup_{x\in[-M,M]} E\left(\left|\alpha x + \sqrt{\lambda x^2 + \beta}\varepsilon\right|^\eta\right) < \infty\,. \tag{2.12}$$

Thus the geometric ergodicity follows from Theorem A4.3 and the same arguments as in the proof of statement (a) of this theorem. The process is therefore strongly mixing with a geometric rate. The symmetry of the stationary distribution follows from the ergodicity and the fact that the processes $(X_n)_{n\in\mathbb{N}}$ and $(-X_n)_{n\in\mathbb{N}}$ have the same transition probabilities, hence the same unique stationary distribution. Finally, because of the continuity of the transition probabilities, the stationary distribution function is continuous as well.

(c) Define now the small set $A := \{x \in \mathbb{R} \,|\, x^2 \leq \max\{1, \frac{\beta}{(1-2\delta) - (\alpha^2 + \lambda\, E(\varepsilon^2))}\}\}$ with $\delta > 0$ such that

$$(1 - 2\,\delta) - (\alpha^2 + \lambda\, E(\varepsilon^2)) > 0\,.$$

Choose $g(x) = 1 + x^2$. Note that for any $x \in A^c$,

$$\begin{aligned}
\int_{\mathbb{R}} g(y) P(X_1 \in dy \,|\, X_0 = x) &= 1 + x^2 \left(\alpha^2 + \lambda\, E(\varepsilon^2) + \frac{\beta}{x^2}\right) \\
&\leq 1 + x^2\,(1 - 2\delta) \\
&= 1 - x^2\delta + x^2(1-\delta) \\
&\leq 1 - \delta + x^2(1-\delta) \;=\; g(x)\,(1-\delta)\,.
\end{aligned} \tag{2.13}$$

By (2.13), (2.12) for $\eta = 2$ and A instead of $[-M, M]$, Theorem A4.4 holds and the second moment of the stationary distribution is finite. □

2.2.2 The Pareto-like tail of the stationary distribution

In this subsection we investigate the tail $\overline{F}(x) = 1 - F(x)$ of the stationary distribution for large x of the AR(1) process $(X_n)_{n\in\mathbb{N}}$ with ARCH(1) errors defined in (1.1). It turns out that the stationary distribution has a Pareto-like tail. We completely specify this tail. To start with, we show that even if the building blocks $(\varepsilon_n)_{n\in\mathbb{N}}$ have moments of all orders not all moments of the stationary distribution are finite.

Proposition 2.2.5 *Suppose $(X_n)_{n\in\mathbb{N}}$ is given by equation (1.1) with $(\varepsilon_n)_{n\in\mathbb{N}}$ satisfying the general conditions and with parameters α and λ satisfying (2.2). Let X be the stationary limit variable of $(X_n)_{n\in\mathbb{N}}$. Choose $N > 0$ such that*

$$E(|\sqrt{\lambda}\varepsilon|^N) > 2\,. \tag{2.14}$$

Then

$$E(|X|^N) = \infty\,.$$

Proof. Assume that the N-th moment is finite. As a consequence of (1.4)

$$\begin{aligned}
E(|X|^N) &= E(|\alpha X + \sqrt{\beta + \lambda X^2}\varepsilon|^N) \\
&= E(1_{\{X<0\}}|X|^N|\alpha + \sqrt{\frac{\beta}{X^2} + \lambda}(-\varepsilon)|^N) + E(1_{\{X>0\}}|X|^N|\alpha + \sqrt{\frac{\beta}{X^2} + \lambda}\varepsilon|^N) \\
&= E(|X|^N|\alpha + \sqrt{\frac{\beta}{X^2} + \lambda}\varepsilon|^N) \\
&\geq E(|X|^N)E(1_{\{\varepsilon>0\}}|\sqrt{\lambda}\varepsilon|^N) \\
&> E(|X|^N)\,,
\end{aligned}$$

where we used in the third and forth line that X and ε are independent. The last line is a consequence of (2.14) and the symmetry of ε. $\square$

Remark 2.2.6 (a) Note that $N > 2$ if $\alpha^2 + \lambda\, E(\varepsilon^2) < 1$ since the second moment exists by Theorem 2.2.3(c).

(b) Condition (2.14) can be replaced by $E(1_{\{\varepsilon>0\}}|\alpha + \sqrt{\lambda}\varepsilon|^N) > 1$ for $\alpha \geq 0$ and $E(1_{\{\varepsilon<0\}}|\alpha + \sqrt{\lambda}\varepsilon|^N) > 1$ for $\alpha < 0$, respectively. These alternative conditions may enable us to find a smaller N.

In order to determine the tail of the distribution of X we need the following technical corollary.

Corollary 2.2.7 *Let $\overline{F}(x) = P(X > x)$, $x \geq 0$, be the right tail of the stationary distribution function . For any $C_1 > 1$, $C_2 > 0$ and $\eta > 0$, there exists some $x_0 > C_1$ such that*

$$\overline{F}(x_0) > C_2\, x_0^{-(N+\eta)}\,,$$

where N is chosen to satisfy (2.14).

Proof. Assume there exist some constants $C_1 > 1, C_2 > 0$ and $\eta > 0$ such that

$$\overline{F}(x) \leq C_2\, x^{-(N+\eta)} \quad \forall x > C_1\,. \tag{2.15}$$

Let $0 < \delta < 1$ be arbitrary. Then, by symmetry of X, using partial integration and (2.15),

$$\begin{aligned} E(1_{\{|X|>C_1\}}|X|^{N-\delta}) &= 2\int_{C_1}^{\infty} x^{N-\delta} dF(x) \\ &= -2\int_{C_1}^{\infty} x^{N-\delta} d\overline{F}(x) \\ &\leq 2C_1^N + 2(N-\delta)\int_{C_1}^{\infty} x^{N-1-\delta}\overline{F}(x)dx \\ &\leq 2C_1^N + 2NC_2\int_{C_1}^{\infty} x^{-1-\eta-\delta}dx \\ &\leq 2C_1^N + 2N\frac{C_2}{\eta} < \infty\,. \end{aligned}$$

Since the rhs is independent of δ, by the monotone convergence theorem ,

$$E(1_{\{|X|>C_1\}}|X|^N) \leq 2C_1^N + 2N\frac{C_2}{\eta} < \infty\,.$$

But this is a contradiction to Proposition 2.2.5 and hence (2.15) is false. □

Because of Proposition 2.2.5 we know that the distribution of X is heavy-tailed in the sense that not all moments exist. In the following we want to find out the precise asymptotic behaviour of its tail. We need the notion of bounded increase of a function, see Bingham, Goldie, Teugels (1987), p.71.

Definition 2.2.8 *Let $h : (c,\infty) \to [0,\infty)$ for some $c \in \mathbb{R}$ and let $\alpha(h)$ be the* upper Matuszewska index*, i.e. $\alpha(h)$ is the infimum of those $\alpha \in \mathbb{R}$ for which there exists a constant $C = C(\alpha)$ such that for each $\Lambda > 1$,*

$$h(\lambda x)/h(x) \leq C(1+o(1))\lambda^{\alpha}, \quad x \to \infty,\ \textit{uniformly in } \lambda \in [1,\Lambda]\,.$$

The function h has bounded increase *if $\alpha(h) < \infty$.*

Remark 2.2.9 Note that non-negative functions which are decreasing have bounded increase.

It turns out that the following modification of the Drasin-Shea Theorem (Bingham et al. (1987), Theorem 5.2.3, p.273) is the key to our result.

Theorem 2.2.10 *Let $k: [0,\infty) \to [0,\infty)$ be an integrable function and let (a,b) be the maximal open interval (where $a < 0$) such that*

$$\check{k}(z) = \int_{(0,\infty)} t^{-z} \frac{k(t)}{t} dt < \infty \quad \text{for } z \in (a,b)\,.$$

If $a > -\infty$, assume $\lim_{\delta \downarrow 0} \check{k}(a+\delta) = \infty$, if $b < \infty$, assume $\lim_{\delta \downarrow 0} \check{k}(b-\delta) = \infty$. Let $h: [0,\infty) \to [0,\infty)$ be locally bounded. Assume h has bounded increase. If

$$\lim_{x\to\infty} \frac{\int_{(0,\infty)} k(x/t)h(t)dt/t}{h(x)} = c > 0\,, \tag{2.16}$$

then

$$c = \check{k}(\rho) \text{ for some } \rho \in (a,b) \text{ and } h(x) \sim x^{\rho} l(x)\,,$$

where l is some slowly varying function.

We will identify h with the tail $\overline{F}$ of the distribution of X. Now we are ready to formulate our main theorem in this section.

Theorem 2.2.11 *Suppose $(X_n)_{n\in\mathbb{N}}$ is given by equation (1.1) with $(\varepsilon_n)_{n\in\mathbb{N}}$ satisfying the general conditions and $(D.1) - (D.3)$ and with parameters α and λ satisfying (2.2). Let $\overline{F}(x) = P(X > x)$, $x \geq 0$, be the right tail of the stationary distribution function. Then*

$$\overline{F}(x) \sim l(x)x^{-\kappa}\,, \quad x \to \infty\,, \tag{2.17}$$

where l is a slowly varying function and κ is given as the unique positive solution to

$$E(|\alpha + \sqrt{\lambda}\varepsilon|^{\kappa}) = 1\,. \tag{2.18}$$

Remark 2.2.12 (a) For the ARCH(1) process (i.e. the case $\alpha = 0$) this result is well-known. The slowly varying function l is then a constant, given implicitly by certain moments of the stationary distribution (see Goldie (1991)).

(b) Note that $E(|\alpha+\sqrt{\lambda}\varepsilon|^{\kappa}) = h_{\alpha,\lambda}(\kappa)$ as in Lemma 2.2.1. Recall that for $\varepsilon \sim N(0,1)$ and fixed λ, the exponent κ is decreasing in $|\alpha|$. This means that the distribution of X gets heavier tails. In particular, our new model has for $\alpha \neq 0$ heavier tails than the ARCH(1) process (see also

Table 2.3).

(c) Theorem 2.2.11 together with Lemma 2.2.1 implies that the second moment of the stationary distribution exists if and only if $\alpha^2 + \lambda\, E(\varepsilon^2) < 1$.

Example 2.2.13 We give three different distributions for $(\varepsilon_n)_{n\in\mathbb{N}}$ which satisfy the general conditions and $(D.1) - (D.3)$.

(a) The normal distribution with mean 0 and variance σ^2:
From Remark 2.2.2(b) it is straightforward that the general conditions and $(D.1), (D.3)$ hold. It remains to show $(D.2)$. Choose $c \geq 0$, $q \in (1/2, 1)$, $x > 0$ and $t > x^q$ arbitrary. Then

$$\begin{aligned} p(\frac{x+c-\alpha t}{\sqrt{\beta+\lambda t^2}}) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{(x-\alpha t)^2}{2\sigma^2(\beta+\lambda t^2)} - \frac{(x-\alpha t)c}{\sigma^2(\beta+\lambda t^2)} - \frac{c^2}{2\sigma^2(\beta+\lambda t^2)}\Big) \\ &\geq \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{(x-\alpha t)^2}{2\sigma^2(\beta+\lambda t^2)} - \frac{c}{\lambda\sigma^2} x^{1-2q} - \frac{c^2}{2\sigma^2(\beta+\lambda x^{2q})}\Big) \\ &= p(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}}) \exp\Big(-\frac{c}{\lambda\sigma^2} x^{1-2q} - \frac{c^2}{2\sigma^2(\beta+\lambda x^{2q})}\Big)\,. \end{aligned}$$

Similarly, we obtain

$$p(\frac{x+c+\alpha t}{\sqrt{\beta+\lambda t^2}}) \geq p(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}}) \exp\Big(-\frac{c}{\lambda\sigma^2} x^{1-2q}(1+\alpha x^{q-1}) - \frac{c^2}{2\sigma^2(\beta+\lambda x^{2q})}\Big)\,.$$

(b) The Laplace (double exponential) distribution:
Consider the probability density $p(x) = \frac{1}{2\theta}\exp(-\frac{|x|}{\theta})$, $x \in \mathbb{R}$, $\theta > 0$. Again it is obvious that the general conditions and $(D.1), (D.3)$ are satisfied. $(D.2)$ holds since for any $c \geq 0$, $q \in (0,1)$, $x > 0$ and $t > x^q$

$$p(\frac{x+c-\alpha t}{\sqrt{\beta+\lambda t^2}}) \geq p(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}}) \exp(-\frac{c}{\theta\sqrt{\lambda}\, x^q})$$

and

$$p(\frac{x+c+\alpha t}{\sqrt{\beta+\lambda t^2}}) \geq p(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}}) \exp(-\frac{c}{\theta\sqrt{\lambda} x^q})\,.$$

(c) The Student's t distribution with $\nu > 2$ degrees of freedom:

$$p_\nu(x) = \frac{\Gamma(\frac{1}{2}(\nu+1))}{\sqrt{\pi\nu}\,\Gamma(\frac{1}{2}\nu)} \left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2}\,.$$

It is well-known that for $\nu \geq 4$

$$E(|\varepsilon|^4) = 3 + \frac{6}{\nu - 4} \tag{2.19}$$

and for any $r \geq \nu$

$$E(|\varepsilon|^r) = \infty \tag{2.20}$$

(see e.g. Johnson, Kotz, Balakrishnan (1995)). The general conditions and $(D.1)$ are clearly fulfilled. Now note that for any $c \geq 0$, $q \in (0,1)$, $x > 0$ and $t > x^q$

$$\begin{aligned} p_\nu(\frac{x+c-\alpha t}{\sqrt{\beta+\lambda t^2}}) &\geq p_\nu(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}}) \left(\frac{x-\alpha\, t}{x+c-\alpha\, t}\right)^{\nu+1} \\ &\geq p_\nu(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}}) \,(1+x^{-1}\frac{c}{1-\alpha x^{q-1}})^{-(\nu+1)} \end{aligned}$$

and similarly,

$$p_\nu(\frac{x+c+\alpha t}{\sqrt{\beta+\lambda t^2}}) \geq p_\nu(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}}) \,(1+x^{-1}\frac{c}{1+\alpha x^{q-1}})^{-(\nu+1)} \,.$$

It remains to check $(D.3)$. Because of (2.20) with $r = \nu$, there exist constants $\eta > 0$ and $q \in (0,1)$ such that

$$\nu > N + \eta + q(\nu+1) + 3q \,, \tag{2.21}$$

where $N = \inf\{u \geq 0 \,|\, E(|\sqrt{\lambda}\,\varepsilon|^u) > 2\}$. For x large enough we thus have

$$p_\nu(x) x^{(N+1+\eta+3q)/(1-q)} \quad \leq \quad const\, x^{(N+\eta-\nu+q(\nu+1)+3q)/(1-q)} \,. \tag{2.22}$$

By (2.21), the exponent in (2.22) is strictly negative and hence $(D.3)$ holds.

The proof of Theorem 2.2.11 will be an application of Theorem 2.2.10. Proposition 2.2.14 presents an implicit formula for the right tail $\overline{F}(x)$, $x > 0$, of the distribution of X. We will need the formula to show that assumption (2.16) is fulfilled. In the following all assumptions of Theorem 2.2.11 hold. Recall that we may w.l.o.g. assume that $\alpha \geq 0$.

Proposition 2.2.14

$$1 = \frac{\overline{H}(x/\sqrt{\beta})}{\overline{F}(x)} + \int_0^\infty f(x,t)dt + \int_0^\infty h(x,t)dt \,, \quad x \in \mathbb{R}, \tag{2.23}$$

where $\overline{H} = 1 - H$ denotes the tail of the distribution function of ε and for $x \in \mathbb{R}$, $t > 0$,

$$\begin{aligned} f(x,t) &:= \Big(p(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}}) + p(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}})\Big) \frac{x\lambda t^2}{(\beta+\lambda t^2)^{3/2}} \frac{\overline{F}(t)}{\overline{F}(x)} \frac{1}{t} \,, \\ h(x,t) &:= \Big(p(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}}) - p(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}})\Big) \frac{\alpha\beta t}{(\beta+\lambda t^2)^{3/2}} \frac{\overline{F}(t)}{\overline{F}(x)} \frac{1}{t} \,. \end{aligned}$$

Proof. By (1.4) and the symmetry of X, we have

$$\begin{aligned}
\overline{F}(x) &= \int_{-\infty}^{\infty} P(\alpha X + \sqrt{\beta+\lambda X^2}\,\varepsilon > x \,|\, X = t) dF(t) \\
&= \int_{-\infty}^{0} P(\alpha t + \sqrt{\beta+\lambda t^2}\varepsilon > x) dF(t) + \int_{0}^{\infty} P(\alpha t + \sqrt{\beta+\lambda t^2}\varepsilon > x) dF(t) \\
&= -\int_{0}^{\infty} P(-\alpha t + \sqrt{\beta+\lambda t^2}\varepsilon > x) dF(-t) + \int_{0}^{\infty} P(\alpha t + \sqrt{\beta+\lambda t^2}\varepsilon > x) dF(t) \\
&= -\int_{0}^{\infty} P(-\alpha t + \sqrt{\beta+\lambda t^2}\varepsilon > x) d\overline{F}(t) + \int_{0}^{\infty} P(\alpha t + \sqrt{\beta+\lambda t^2}\varepsilon > x) dF(t) \\
&= -\int_{0}^{\infty} \Big(\overline{H}(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}}) + \overline{H}(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}})\Big) d\overline{F}(t)\,,
\end{aligned}$$

where we used in the fourth line that the distribution of X is symmetric. By partial integration and again symmetry, we obtain

$$\begin{aligned}
\overline{F}(x) &= \overline{H}(\frac{x}{\sqrt{\beta}}) - \int_0^\infty \Big(p(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}})\frac{\alpha(\beta+\lambda t^2)-(x+\alpha t)\lambda t}{(\beta+\lambda t^2)^{3/2}} \\
&\qquad + p(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}})\frac{-\alpha(\beta+\lambda t^2)-(x-\alpha t)\lambda t}{(\beta+\lambda t^2)^{3/2}}\Big)\overline{F}(t)dt \\
&= \overline{H}(\frac{x}{\sqrt{\beta}}) + \int_0^\infty \Big(p(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}}) + p(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}})\Big)\frac{x\lambda t^2}{(\beta+\lambda t^2)^{3/2}}\overline{F}(t)\frac{dt}{t} \qquad (2.24)\\
&\qquad + \int_0^\infty \Big(p(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}}) - p(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}})\Big)\frac{\alpha\beta t}{(\beta+\lambda t^2)^{3/2}}\overline{F}(t)\frac{dt}{t}\,.
\end{aligned}$$

This finishes the proof. □

We investigate now (2.23). Together with Lemma 2.2.5 and Corollary 2.2.7 we collect some results in the following lemmas and corollaries. These results will be crucial in applying Theorem 2.2.10.

Lemma 2.2.15 $\lim_{x\to\infty}\int_0^\infty f(x,t)dt = 1\,.$

Proof. Let $q \in (0,1)$ and $\eta > 0$ be the constants in $(D.2)$ and $(D.3)$, respectively. Because of assumption $(D.3)$, there exist constants x_0, $D > 0$ such that for all $x > x_0$

$$p(x) \le D\, x^{-(N+1+\eta+3q)}$$

and thus we have for all $x > x_0$

$$\overline{H}(x/\sqrt{\beta}) \le \frac{D\,\beta^{(N+\eta+3q)/2}}{N+\eta+3q}\, x^{-(N+\eta+3q)}\,. \qquad (2.25)$$

By Corollary 2.2.7 and (2.25), there exists therefore a sequence $(x_n)\uparrow\infty$ and a constant $\widetilde{D} > 0$ such that for any $x_n > \sqrt{\beta}\,x_0$

$$\frac{\overline{H}(x_n/\sqrt{\beta})}{\overline{F}(x_n)} \le \frac{D\,\beta^{(N+\eta+3q)/2}}{\widetilde{D}\,(N+\eta+3q)}\, x_n^{-3q}\,. \qquad (2.26)$$

Note now that

$$|h(x,t)| \le \frac{\alpha\beta}{x\lambda} f(x,t), \quad \text{for any t} \ge 1 \text{ and } x \ge 0. \tag{2.27}$$

Hence,

$$\left| \int_1^\infty h(x,t)dt \right| \le \frac{\alpha\beta}{x\lambda} \int_1^\infty f(x,t)dt\,. \tag{2.28}$$

Because of (2.28) and the fact that due to assumption $(D.1)$

$$h(x,t) \ge 0 \tag{2.29}$$

for x large enough and $t \in [0,1]$, we have

$$\limsup_{x\to\infty} \int_0^\infty f(x,t)dt \le 1\,. \tag{2.30}$$

Fix now some $t \ge 0$. By Corollary 2.2.7 there exists a sequence $(x_n)_{n\in\mathbb{N}} \uparrow \infty$ such that for some function $c(t) \ge 0$ and any $\eta > 0$

$$0 \le f(x_n,t) \le c(t)\Big(p(\frac{x_n - \alpha t}{\sqrt{\beta + \lambda t^2}}) + p(\frac{x_n + \alpha t}{\sqrt{\beta + \lambda t^2}})\Big) x_n^{N+1+\eta}\,,$$

and therefore by condition $(D.3)$

$$\liminf_{x\to\infty} f(x,t) = 0\,, \qquad \text{for any } t \in (0,\infty)\,.$$

With similar arguments we derive also

$$\liminf_{x\to\infty} h(x,t) = 0\,, \qquad \text{for any } t \in (0,\infty)\,.$$

Because of the continuity of the functions f and h we have thus that for any fixed $T \ge 0$

$$\liminf_{x\to\infty} \int_0^T f(x,t)dt = 0 \quad \text{ and } \quad \liminf_{x\to\infty} \int_0^T h(x,t)dt = 0\,. \tag{2.31}$$

Thus again with (2.26), (2.28) and (2.31) with $T = 1$ we obtain

$$\liminf_{x\to\infty} \int_0^\infty f(x,t)dt = 1\,,$$

which finishes the proof. □

The next corollary is a consequence of Lemma 2.2.15 and its proof. The result supports our supposition that the stationary distribution has a Pareto-tail.

Corollary 2.2.16 $\lim_{x\to\infty} \overline{F}(x+c)/\overline{F}(x) = 1$ *for any* $c \in (0,\infty)$.

Proof. By monotonicity, $\limsup_{x\to\infty} \overline{F}(x+c)/\overline{F}(x) \le 1$. Furthermore, due to Lemma 2.2.15,

$$\begin{aligned}
&\liminf_{x\to\infty} \frac{\overline{F}(x+c)}{\overline{F}(x)} \\
&= \liminf_{x\to\infty} \frac{\overline{F}(x+c)}{\overline{F}(x)} \int_0^\infty f(x+c,t)dt \\
&\ge \liminf_{x\to\infty} \int_0^\infty \Big(p(\frac{x+c-\alpha t}{\sqrt{\beta+\lambda t^2}}) + p(\frac{x+c+\alpha t}{\sqrt{\beta+\lambda t^2}})\Big) \frac{(x+c)\lambda t^2}{(\beta+\lambda t^2)^{3/2}} \frac{\overline{F}(t)}{\overline{F}(x)} \frac{dt}{t} \\
&\ge \liminf_{x\to\infty} \int_{x^q}^\infty \Big(p(\frac{x+c-\alpha t}{\sqrt{\beta+\lambda t^2}}) + p(\frac{x+c+\alpha t}{\sqrt{\beta+\lambda t^2}})\Big) \frac{x\lambda t^2}{(\beta+\lambda t^2)^{3/2}} \frac{\overline{F}(t)}{\overline{F}(x)} \frac{dt}{t}, \qquad (2.32)
\end{aligned}$$

where $q \in (0,1)$ is the constant in $(D.2)$. Now choose $\delta > 0$ arbitrary. Because of condition $(D.2)$ (2.32) may be estimated below by

$$\begin{aligned}
&\liminf_{x\to\infty} \frac{\overline{F}(x+c)}{\overline{F}(x)} \\
&\ge (1-\delta) \liminf_{x\to\infty} \int_{x^q}^\infty \Big(p(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}}) + p(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}})\Big) \frac{x\lambda t^2}{(\beta+\lambda t^2)^{3/2}} \frac{\overline{F}(t)}{\overline{F}(x)} \frac{dt}{t} \\
&= (1-\delta) \liminf_{x\to\infty} \int_{x^q}^\infty f(x,t)dt\,. \qquad (2.33)
\end{aligned}$$

Next consider

$$\begin{aligned}
&\liminf_{x\to\infty} \int_0^\infty f(x,t)dt \\
&= \liminf_{x\to\infty} \int_0^T f(x,t)dt + \liminf_{x\to\infty} \int_T^{x^q} f(x,t)dt + \liminf_{x\to\infty} \int_{x^q}^\infty f(x,t)dt \\
&=: J_1 + J_2 + J_3\,. \qquad (2.34)
\end{aligned}$$

We showed in the proof of Lemma 2.2.15 that $J_1 = 0$. Furthermore, by assumption $(D.1)$,

$$\begin{aligned}
J_2 &\le \liminf_{x\to\infty} \Big(p(\frac{x-\alpha x^q}{\sqrt{\beta+\lambda x^{2q}}}) + p(\frac{x}{\sqrt{\beta+\lambda x^{2q}}})\Big) \frac{x^{2q+1}\lambda}{(\beta+\lambda T^2)^{3/2}T} \frac{1}{\overline{F}(x)} (x^q - T) \\
&\le \liminf_{x\to\infty} \Big(p(\frac{x-\alpha x^q}{\sqrt{\beta+\lambda x^{2q}}}) + p(\frac{x}{\sqrt{\beta+\lambda x^{2q}}})\Big) \frac{x^{3q+1}\lambda}{(\beta+\lambda T^2)^{3/2}T} \frac{1}{\overline{F}(x)} \\
&= \liminf_{x\to\infty} \Big(p(\frac{x^{1-q}(1-\alpha x^{q-1})}{\sqrt{\beta/x^{2q}+\lambda}}) + p(\frac{x^{1-q}}{\sqrt{\beta/x^{2q}+\lambda}})\Big) \frac{x^{3q+1}\lambda}{(\beta+\lambda T^2)^{3/2}T} \frac{1}{\overline{F}(x)}\,.
\end{aligned}$$

From Corollary 2.2.7 and the assumption $(D.3)$ we conclude $J_2 = 0$. Plugging all this together we get from (2.33) that

$$\liminf_{x\to\infty} \frac{\overline{F}(x+c)}{\overline{F}(x)} \ge (1-\delta) \liminf_{x\to\infty} \int_0^\infty f(x,t)dt = 1-\delta\,,$$

where the last line follows from Lemma 2.2.15. Because δ was arbitrary the corollary is proven.

□

Lemma 2.2.17 $\lim_{x\to\infty}\int_0^\infty g(x,t)dt = 1$, *where*

$$g(x,t) := \left(p(\frac{x-\alpha t}{\sqrt{\lambda t}}) + p(\frac{x+\alpha t}{\sqrt{\lambda t}})\right)\frac{x\lambda t^2}{(\lambda t^2)^{3/2}}\frac{\overline{F}(t)}{\overline{F}(x)}\frac{1}{t}\,.$$

Proof. By the general conditions and assumption $(D.1)$, for any $x,t\geq 0$

$$p(\frac{x\pm\alpha t}{\sqrt{\lambda}\,t}) \leq p(\frac{x\pm\alpha t}{\sqrt{\beta+\lambda t^2}})$$

and hence with Lemma 2.2.15 and the same arguments as in the proof of Corollary 2.2.16 we get

$$\begin{aligned}\limsup_{x\to\infty}\int_0^\infty g(x,t)dt &= \limsup_{x\to\infty}\int_{x^q}^\infty g(x,t)dt\\ &\leq \limsup_{x\to\infty}\int_{x^q}^\infty f(x,t)dt\\ &= \limsup_{x\to\infty}\int_0^\infty f(x,t)dt = 1\,,\end{aligned}$$

where $q\in(0,1)$ is the constant in $(D.2)$. It remains to show the converse inequality for the limes inferior. Note that for any fixed $T\geq 0$

$$\liminf_{x\to\infty}\int_0^T g(x,t)dt = 0\,. \tag{2.35}$$

By (2.35), the general conditions and assumptions $(D.1),(D.2)$ and substitution $t=\sqrt{\beta/\lambda+s^2}$, we have

$$\begin{aligned}&\liminf_{x\to\infty}\int_0^\infty g(x,t)dt = \liminf_{x\to\infty}\int_{\sqrt{\beta/\lambda}}^\infty g(x,t)dt\\ &= \liminf_{x\to\infty}\int_0^\infty\Big(p(\frac{x-\alpha\sqrt{\beta/\lambda+s^2}}{\sqrt{\beta+\lambda s^2}}) + p(\frac{x+\alpha\sqrt{\beta/\lambda+s^2}}{\sqrt{\beta+\lambda s^2}})\Big)\frac{x\lambda s^2}{(\beta+\lambda s^2)^{3/2}}\frac{\overline{F}(\sqrt{s^2+\beta/\lambda})}{\overline{F}(x)}\frac{ds}{s}\\ &\geq \liminf_{x\to\infty}\int_0^\infty 1_{\{x-\alpha\sqrt{\beta/\lambda+s^2}\geq 0\}}p(\frac{x-\alpha s}{\sqrt{\beta+\lambda s^2}})\frac{x\lambda s^2}{(\beta+\lambda s^2)^{3/2}}\frac{\overline{F}(\sqrt{s^2+\beta/\lambda})}{\overline{F}(x)}\frac{ds}{s}\\ &\quad+\liminf_{x\to\infty}\int_0^\infty 1_{\{x-\alpha\sqrt{\beta/\lambda+s^2}<0\}}p(\frac{x-\alpha\sqrt{\beta/\lambda}-\alpha s}{\sqrt{\beta+\lambda s^2}})\frac{x\lambda s^2}{(\beta+\lambda s^2)^{3/2}}\frac{\overline{F}(\sqrt{s^2+\beta/\lambda})}{\overline{F}(x)}\frac{ds}{s}\\ &\quad+\liminf_{x\to\infty}\int_0^\infty p(\frac{x+\alpha\sqrt{\beta/\lambda}+\alpha s}{\sqrt{\beta+\lambda s^2}})\frac{x\lambda s^2}{(\beta+\lambda s^2)^{3/2}}\frac{\overline{F}(\sqrt{s^2+\beta/\lambda})}{\overline{F}(x)}\frac{ds}{s}\\ &\geq \liminf_{y\to\infty}\int_{(y+\alpha\sqrt{\beta/\lambda})^q}^\infty 1_{\{y+\alpha\sqrt{\beta/\lambda}-\alpha\sqrt{\beta/\lambda+s^2}\geq 0\}}p(\frac{y+\alpha\sqrt{\beta/\lambda}-\alpha s}{\sqrt{\beta+\lambda s^2}})\\ &\qquad\qquad\frac{y\lambda s^2}{(\beta+\lambda s^2)^{3/2}}\frac{\overline{F}(\sqrt{s^2+\beta/\lambda})}{\overline{F}(y+\alpha\sqrt{\beta/\lambda})}\frac{ds}{s}\\ &\quad+\liminf_{y\to\infty}\int_{(y+\alpha\sqrt{\beta/\lambda})^q}^\infty 1_{\{y+\alpha\sqrt{\beta/\lambda}-\alpha\sqrt{\beta/\lambda+s^2}<0\}}p(\frac{y-\alpha s}{\sqrt{\beta+\lambda s^2}})\end{aligned}$$

$$
\begin{aligned}
& \qquad\qquad \frac{y\lambda s^2}{(\beta+\lambda s^2)^{3/2}} \frac{\overline{F}(\sqrt{s^2+\beta/\lambda})}{\overline{F}(y+\alpha\sqrt{\beta/\lambda})} \frac{ds}{s} \\
& + \liminf_{y\to\infty} \int_{(y+\alpha\sqrt{\beta/\lambda})^q}^{\infty} f_+(\alpha\sqrt{\beta/\lambda}, y+\alpha\sqrt{\beta/\lambda})\, p(\frac{y+\alpha\sqrt{\beta/\lambda}+\alpha s}{\sqrt{\beta+\lambda s^2}}) \\
& \qquad\qquad \frac{y\lambda s^2}{(\beta+\lambda s^2)^{3/2}} \frac{\overline{F}(\sqrt{s^2+\beta/\lambda})}{\overline{F}(y+\alpha\sqrt{\beta/\lambda})} \frac{ds}{s} \\
\geq\ & \liminf_{y\to\infty} \int_{(y+\alpha\sqrt{\beta/\lambda})^q}^{\infty} 1_{\{y+\alpha\sqrt{\beta/\lambda}-\alpha\sqrt{\beta/\lambda+s^2}\geq 0\}} f_-(\alpha\sqrt{\beta/\lambda}, y)\, p(\frac{y-\alpha s}{\sqrt{\beta+\lambda s^2}}) \\
& \qquad\qquad \frac{y\lambda s^2}{(\beta+\lambda s^2)^{3/2}} \frac{\overline{F}(\sqrt{s^2+\beta/\lambda})}{\overline{F}(y)} \frac{ds}{s} \\
& + \liminf_{y\to\infty} \int_{(y+\alpha\sqrt{\beta/\lambda})^q}^{\infty} 1_{\{y+\alpha\sqrt{\beta/\lambda}-\alpha\sqrt{\beta/\lambda+s^2}< 0\}} p(\frac{y-\alpha s}{\sqrt{\beta+\lambda s^2}}) \\
& \qquad\qquad \frac{y\lambda s^2}{(\beta+\lambda s^2)^{3/2}} \frac{\overline{F}(\sqrt{s^2+\beta/\lambda})}{\overline{F}(y)} \frac{ds}{s} \\
& + \liminf_{y\to\infty} \int_{(y+\alpha\sqrt{\beta/\lambda})^q}^{\infty} f_+(\alpha\sqrt{\beta/\lambda}, y+\alpha\sqrt{\beta/\lambda})\, f_+(\alpha\sqrt{\beta/\lambda}, y)\, p(\frac{y+\alpha s}{\sqrt{\beta+\lambda s^2}}) \\
& \qquad\qquad \frac{y\lambda s^2}{(\beta+\lambda s^2)^{3/2}} \frac{\overline{F}(\sqrt{s^2+\beta/\lambda})}{\overline{F}(y)} \frac{ds}{s}\,,
\end{aligned} \tag{2.36}
$$

where $q \in (0,1)$ is the constant in $(D.2)$. Now choose any $\delta > 0$ and T so large that for any $s, y \geq T$

$$f_-(\alpha\sqrt{\beta/\lambda}, y) > (1-\delta)^{1/3}\,, \tag{2.37}$$

$$f_+(\alpha\sqrt{\beta/\lambda}, y) > (1-\delta)^{1/3}\,, \tag{2.38}$$

$$\frac{\overline{F}(\sqrt{s^2+\beta/\lambda})}{\overline{F}(s)} > (1-\delta)^{1/3}\,. \tag{2.39}$$

(2.39) holds because of Corollary 2.2.16. Plugging (2.37)-(2.39) in (2.36) we get

$$
\begin{aligned}
& \liminf_{x\to\infty} \int_0^{\infty} g(x,t)dt \\
\geq\ & (1-\delta)^{2/3} \liminf_{y\to\infty} \int_{(y+\alpha\sqrt{\beta/\lambda})^q}^{\infty} 1_{\{y+\alpha\sqrt{\beta/\lambda}-\alpha\sqrt{\beta/\lambda+s^2}\geq 0\}} p(\frac{y-\alpha s}{\sqrt{\beta+\lambda s^2}}) \frac{y\lambda s^2}{(\beta+\lambda s^2)^{3/2}} \frac{\overline{F}(s)}{\overline{F}(y)} \frac{ds}{s} \\
& +(1-\delta)^{1/3} \liminf_{y\to\infty} \int_{(y+\alpha\sqrt{\beta/\lambda})^q}^{\infty} 1_{\{y+\alpha\sqrt{\beta/\lambda}-\alpha\sqrt{\beta/\lambda+s^2}< 0\}} p(\frac{y-\alpha s}{\sqrt{\beta+\lambda s^2}}) \frac{y\lambda s^2}{(\beta+\lambda s^2)^{3/2}} \frac{\overline{F}(s)}{\overline{F}(y)} \frac{ds}{s} \\
& +(1-\delta) \liminf_{y\to\infty} \int_{(y+\alpha\sqrt{\beta/\lambda})^q}^{\infty} p(\frac{y+\alpha s}{\sqrt{\beta+\lambda s^2}}) \frac{y\lambda s^2}{(\beta+\lambda s^2)^{3/2}} \frac{\overline{F}(s)}{\overline{F}(y)} \frac{ds}{s} \\
\geq\ & (1-\delta) \liminf_{y\to\infty} \int_{(y+\alpha\sqrt{\beta/\lambda})^q}^{\infty} f(y,t)dt = (1-\delta)\,,
\end{aligned}
$$

where the last line follows from Lemma 2.2.15 and the proof of Corollary 2.2.16. □

With these lemmas we are now able to prove Theorem 2.2.11.

Proof of Theorem 2.2.11. The proof is just an application of Theorem 2.2.10. Choose

$$k(x) = \frac{x}{\sqrt{\lambda}}\Big(p(\frac{x-\alpha}{\sqrt{\lambda}}) + p(\frac{x+\alpha}{\sqrt{\lambda}})\Big)\,, \quad x > 0\,, \tag{2.40}$$

and

$$h(x) = \overline{F}(x)\,, \quad x > 0\,. \tag{2.41}$$

One can readily see that k is non-negative, h is non-negative, locally bounded and of bounded increase because of Remark 2.2.9. Note that for any $z \in (-\infty, \infty)$

$$\begin{aligned}
\check{k}(z) &= \int_0^\infty t^{-z}\frac{k(t)}{t}dt \\
&= \int_0^\infty t^{-z}\frac{1}{\sqrt{\lambda}}p(\frac{t-\alpha}{\sqrt{\lambda}})dt + \int_{-\infty}^0 (-t)^{-z}\frac{1}{\sqrt{\lambda}}p(\frac{t-\alpha}{\sqrt{\lambda}})dt \\
&= E(|\alpha + \sqrt{\lambda}\varepsilon|^{-z})\,.
\end{aligned}$$

Let (a, b) be the maximal open interval such that

$$\check{k}(z) < \infty \quad \text{for } z \in (a, b)\,.$$

Note that $a = -T = -\inf\{u \geq 0 \,|\, h_{\alpha,\lambda}(u) = \infty\} < 0$ and $b = 1$ because of Proposition 2.2.1 and the fact that for $z \geq 0$

$$\int_1^\infty t^{-z}\frac{k(t)}{t}dt \leq \int_1^\infty \frac{1}{\sqrt{\lambda}}\Big(p(\frac{t-\alpha}{\sqrt{\lambda}}) + p(\frac{t+\alpha}{\sqrt{\lambda}})\Big)dt < \infty$$

and

$$\int_0^1 t^{-z}\frac{k(t)}{t}dt \leq const \int_0^1 t^{-z}dt = \begin{cases} < \infty\,, & z < 1 \\ = \infty\,, & z \geq 1 \end{cases}.$$

Furthermore, by the dominated and monotone convergence theorem, respectively,

$$\begin{aligned}
\lim_{\delta\downarrow 0}\check{k}(a+\delta) &= \lim_{\delta\downarrow 0} E\left(1_{\{|\alpha+\sqrt{\lambda}\varepsilon|\leq 1\}}|\alpha + \sqrt{\lambda}\,\varepsilon|^{-(a+\delta)}\right) \\
&\qquad + \lim_{\delta\downarrow 0} E\left(1_{\{|\alpha+\sqrt{\lambda}\varepsilon|> 1\}}|\alpha + \sqrt{\lambda}\,\varepsilon|^{-(a+\delta)}\right) \\
&= E\left(1_{\{|\alpha+\sqrt{\lambda}\varepsilon|\leq 1\}}|\alpha + \sqrt{\lambda}\,\varepsilon|^{T}\right) \,+\, E\left(1_{\{|\alpha+\sqrt{\lambda}\varepsilon|> 1\}}|\alpha + \sqrt{\lambda}\,\varepsilon|^{T}\right) \\
&= h_{\alpha,\lambda}(T) \,=\, \infty
\end{aligned}$$

and

$$\begin{aligned}
\lim_{\delta\downarrow 0}\check{k}(b-\delta) &= \lim_{\delta\downarrow 0}\int_0^\infty t^{-(1-\delta)}\frac{1}{\sqrt{\lambda}}\Big(p(\frac{t-\alpha}{\sqrt{\lambda}}) + p(\frac{t+\alpha}{\sqrt{\lambda}})\Big)dt \\
&\geq const \lim_{\delta\downarrow 0}\int_0^1 t^{-(1+\delta)}dt \,=\, const \lim_{\delta\downarrow 0}\frac{1}{\delta} \,=\, \infty\,.
\end{aligned}$$

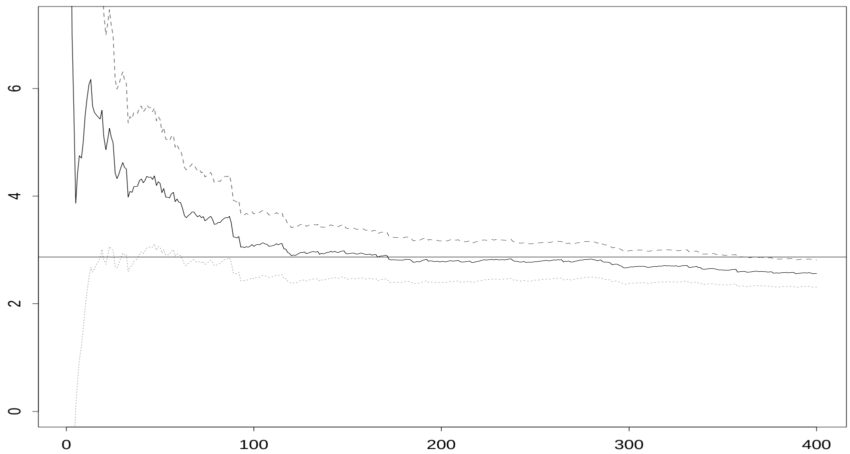

Figure 2.1: The Hill estimator for the stationary distribution of the autoregressive process with ARCH(1) errors with length $n = 10000$ and parameters $\alpha = 0.4, \lambda = 0.6$ and $\varepsilon \sim N(0,1)$. We calculate the 5 and 95 percent empirical quantiles (dotted lines) and the empirical median (solid line). The horizontal line indicates the numerical solution of κ in (2.18). From Table 2.3 we know that $\kappa = 2.87$.

Finally, by Lemma 2.2.17, we have

$$\lim_{x\to\infty} \frac{\int_0^\infty k(x/t)\overline{F}(t)dt/t}{\overline{F}(x)} = \lim_{x\to\infty} \int_0^\infty g(x,t)dt = 1$$

and hence condition (2.16) is fulfilled with $c = 1$. Therefore all assumptions of Theorem 2.2.10 are satisfied and we derive (setting $\kappa = -\rho$) that

$$\overline{F}(x) \sim x^{-\kappa} l(x)\,, \tag{2.42}$$

where l is some slowly varying function and κ is determined by the equation

$$E(|\alpha + \sqrt{\lambda}\varepsilon|^\kappa) = 1\,, \quad \text{for some } \kappa \in (-1, T)\,. \tag{2.43}$$

Since the tail of the stationary distribution function is decreasing the solution κ in (2.43) has to be strictly positive and hence by Theorem 2.2.10 there exists a solution $\kappa \in (0, T)$ in (2.43) which is unique because of Lemma 2.2.1. $\square$

Theorem 2.2.11 is a generalisation of the result of the ARCH(1) process as a consequence of using a different kind of technique for the proof. However, the Drasin-Shea Tauberian theorem guarantees only a regularly varying tail, i.e. an unknown slowly varying function l appears whereas in the ARCH(1) case this is a constant which was calculated explicitly by Goldie (1991). In Goldie (1991), the distribution tail of a random recurrence equation is derived by renewal arguments. As mentioned before, the AR(1) process with ARCH(1) errors does not fit into this setting. However, we introduce another process $(Y_n)_{n\in\mathbb{N}}$ which has the same stationary distribution as the process $(|X_n|)_{n\in\mathbb{N}}$ and which satisfies the conditions needed to fit into Goldie's framework. To check Goldie's assumptions leading to a Pareto-like tail we rely on the results which we have worked out so far.

We investigate the process $(Y_n)_{n\in\mathbb{N}}$ given by the stochastic difference equation

$$Y_n = |\alpha Y_{n-1} + \sqrt{\beta + \lambda Y_{n-1}^2}\varepsilon_n| , \quad n \geq 1 , \tag{2.44}$$

where $(\varepsilon_n)_{n\in\mathbb{N}}$ are the same i.i.d. random variables as in Theorem 2.2.11, the constants are the same as for the process $(X_n)_{n\in\mathbb{N}}$ and Y_0 equals $|X_0|$ a.s. The following lemma manifests that the processes $(Y_n)_{n\in\mathbb{N}}$ and $(|X_n|)_{n\in\mathbb{N}}$ have trivially the same stationary distribution.

Lemma 2.2.18 *If $|X_0| = Y_0$ a.s. then $(|X_n|)_{n\in\mathbb{N}} \stackrel{d}{=} (Y_n)_{n\in\mathbb{N}}$.*

Proof. Note first that the processes $(|X_n|)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ are homogeneous Markov processes. It suffices therefore to show that the one-dimensional transition probabilities are the same. For any $x \in \mathbb{R}$,

$$\begin{aligned}
P(Y_1 \leq x \,|\, Y_0) &= P\left(\left|\alpha|X_0| + \sqrt{\beta + \lambda X_0^2}\varepsilon_1\right| \leq x \,\Big|\, |X_0|\right) \\
&= P(-x \leq \alpha X_0 + \sqrt{\beta + \lambda X_0^2}\varepsilon_1 \leq x, X_0 \geq 0 \,|\, |X_0|) \\
&\qquad + P(-x \leq \alpha(-X_0) + \sqrt{\beta + \lambda X_0^2}\varepsilon_1 \leq x, X_0 < 0 \,|\, |X_0|) \\
&= P(-x \leq \alpha X_0 + \sqrt{\beta + \lambda X_0^2}\varepsilon_1 \leq x, X_0 \geq 0 \,|\, |X_0|) \\
&\qquad + P(-x \leq \alpha X_0 + \sqrt{\beta + \lambda X_0^2}\varepsilon_1 \leq x, X_0 < 0 \,|\, |X_0|) \\
&= P(|\alpha X_0 + \sqrt{\beta + \lambda X_0^2}\varepsilon_1| \leq x \,|\, |X_0|) \\
&= P(|X_1| \leq x \,|\, |X_0|)
\end{aligned}$$

This finishes the proof. $\square$

Corollary 2.2.19 *The slowly varying function l in Theorem 2.2.11 can be chosen as the constant*

$$c = \frac{1}{2\kappa} \frac{E\left(\left|\alpha|X| + \sqrt{\beta + \lambda X^2}\varepsilon\right|^\kappa - \left|(\alpha + \sqrt{\lambda}\varepsilon)|X|\right|^\kappa\right)}{E\left(|\alpha + \sqrt{\lambda}\varepsilon|^\kappa \ln|\alpha + \sqrt{\lambda}\varepsilon|\right)} . \tag{2.45}$$

Remark 2.2.20 In the ARCH(1)-case this result is known (see Goldie (1991)). But there, the result was proven differently by investigating $(X_n^2)_{n\in\mathbb{N}}$.

Proof. The proof is an application of Corollary 2.4 of Goldie (1991). Consider the process in (2.44). Define $M := |\alpha + \sqrt{\lambda}\varepsilon|$ and choose κ as in Theorem 2.2.11. Due to our assumptions on ε the conditions of Corollary 2.4 of Goldie (1991) on M are satisfied. By Lemma 2.2.18 the stationary distribution of the process $(Y_n)_{n\in\mathbb{N}}$ is the same as for $(|X|_n)_{n\in\mathbb{N}}$. In particular, by Theorem 2.2.11,

$$E(Y^{\kappa-1}) < \infty\,, \tag{2.46}$$

where Y has the stationary distribution of $(Y_n)_{n\in\mathbb{N}}$. Finally,

$$\begin{aligned}
&E\left(\left|(|\alpha Y + \sqrt{\beta+\lambda Y^2}\,\varepsilon|)^\kappa - (|\alpha + \sqrt{\lambda}\,\varepsilon|Y)^\kappa\right|\right)\\
&\le\; E\left(\kappa|(\alpha Y + \sqrt{\beta+\lambda Y^2}\varepsilon) - (\alpha Y + \sqrt{\lambda}Y\varepsilon)|\max\{|\alpha Y + \sqrt{\beta+\lambda Y^2}\varepsilon|, |\alpha+\sqrt{\lambda}\varepsilon|Y\}^{\kappa-1}\right)\\
&\le\; E\left(\kappa\sqrt{\beta}|\varepsilon|(|\alpha Y + \sqrt{\beta+\lambda Y^2}\varepsilon| + |\alpha+\sqrt{\lambda}\varepsilon|Y)^{\kappa-1}\right)\\
&\le\; E\left(\kappa\sqrt{\beta}|\varepsilon|(const\,\alpha^{\kappa-1}Y^{\kappa-1} + const\,(\beta+\lambda Y^2)^{(\kappa-1)/2}|\varepsilon|^{\kappa-1} + const\,\lambda^{(\kappa-1)/2}|\varepsilon|^{\kappa-1}Y^{\kappa-1})\right)\\
&\le\; const\,\kappa\sqrt{\beta}\alpha^{\kappa-1}E(|\varepsilon|)E(Y^{\kappa-1}) + const\,\kappa\beta^{\kappa/2}E(|\varepsilon|^\kappa)\\
&\quad +const\,\kappa\sqrt{\beta}\lambda^{(\kappa-1)/2}E(Y^{\kappa-1}) + const\,\kappa\sqrt{\beta}\lambda^{(\kappa-1)/2}E(|\varepsilon|^\kappa)E(Y^{\kappa-1})\,,
\end{aligned}$$

where the second line follows from the fact that

$$||x|^r - |y|^r| \le r|x-y|\max\{|x|,|y|\}^{r-1}, \quad \text{for any } x,y\in\mathbb{R} \text{ and } 1<r<\infty\,,$$

and in the forth and fifth line we used that

$$|x+y|^r \le const\,(|x|^r + |y|^r), \quad \text{for any } x,y\in\mathbb{R} \text{ and } r>0\,.$$

Together with (2.46) we have

$$E\left(\left|(|\alpha Y + \sqrt{\beta+\lambda Y^2}\,\varepsilon|)^\kappa - (|\alpha + \sqrt{\lambda}\,\varepsilon|Y)^\kappa\right|\right) < \infty\,.$$

Hence all assumptions of Corollary 2.4 of Goldie (1991) are satisfied and the result follows using the facts that Y and $|X|$ have the same distribution and X is symmetric. □

Remark 2.2.21 In the proof of Corollary 2.2.19 we did not use the whole knowledge of the tail of the stationary distribution of the process $(X_n)_{n\in\mathbb{N}}$ from Theorem 2.2.11. We only used

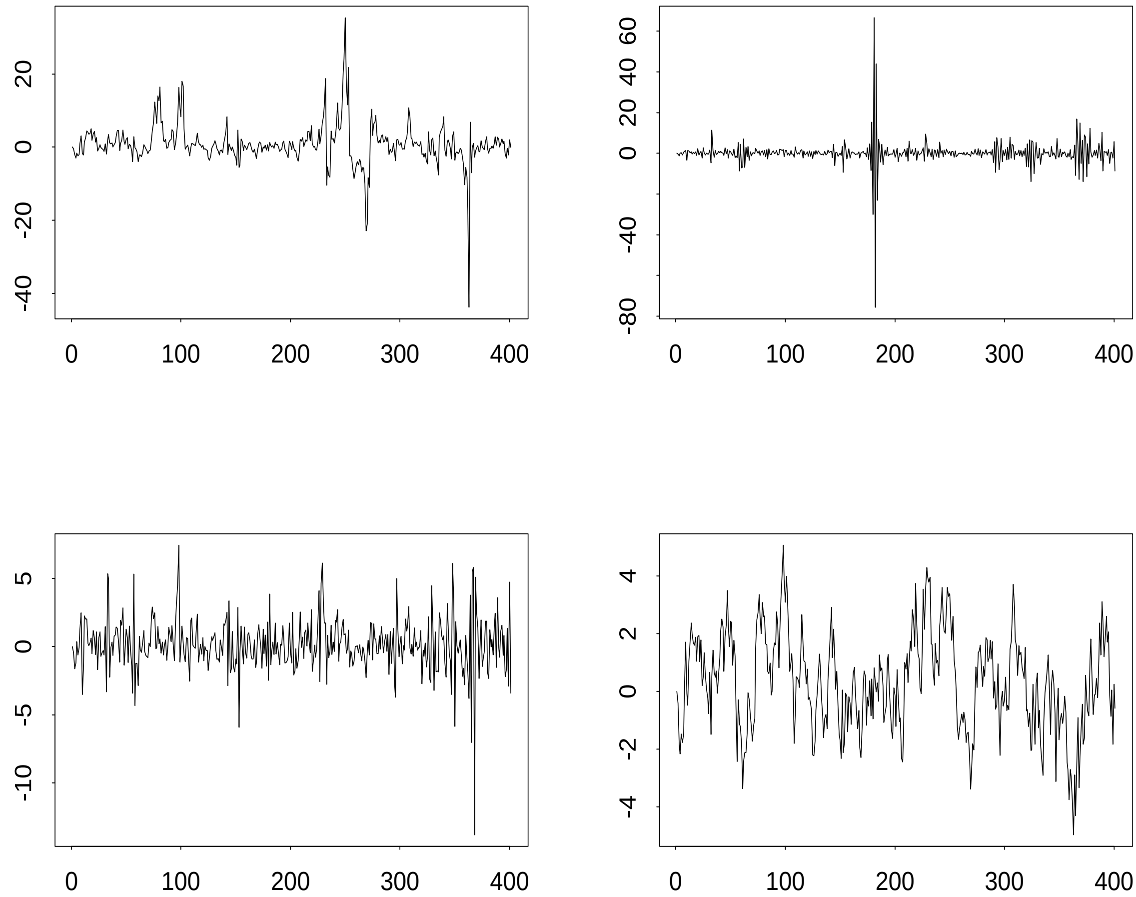

Figure 2.2: Simulated sample path of the AR(1) process with ARCH(1) errors $(X_n)_{n\in\mathbb{N}}$ with parameters $\alpha = 0.8, \beta = 1, \lambda = 0.8$ (top, left), with $\alpha = -0.8, \beta = 1, \lambda = 0.8$ (top, right), with $\alpha = 0, \beta = 1, \lambda = 0.8$ (bottom, left), and with $\alpha = 0.8, \beta = 1, \lambda = 0$ (bottom, right) in the case $\varepsilon \sim N(0,1)$. The innovations $(\varepsilon_n)_{n\in\mathbb{N}}$ are the same in all four pictures. It appears that the marginal distributions of the nondegenerate AR(1) processes with ARCH(1) errors (top) have clearly fatter tails than the ones of the ARCH(1) process and of the AR(1) process.

that the $(\kappa-1)$-th moment of the stationary distribution exists. It might be supposed that this result could be found easier. In the case $\varepsilon \sim N(0,1)$ this is indeed possible. Recall that the random variable X which has the stationary distribution function is characterized by the fixpoint equation

$$X \stackrel{d}{=} \alpha X + \sqrt{\beta+\lambda X^2}\,\varepsilon\,. \tag{2.47}$$

Now note that for any $t \in \mathbb{R}$

$$\begin{aligned} E(e^{itX}) &= E(e^{it\alpha X}\,E(e^{it\sqrt{\beta+\lambda X^2}\,\varepsilon}\mid X)) \\ &= e^{-\beta t^2/2}\,E(e^{it\alpha X - t^2\lambda X^2/2}) \\ &= E(e^{it\sqrt{\beta}N_1})\,E(e^{it(\alpha X+\sqrt{\lambda}X\,N_2)})\,, \end{aligned} \tag{2.48}$$

where N_1 and N_2 are independent standard normal random variables, independent of X. From (2.48) we obtain the fixpoint equation

$$X \stackrel{d}{=} \sqrt{\beta}N_1 + (\alpha+\sqrt{\lambda}N_2)X\,.$$

Hence X is limit variable of the ergodic process $(\widetilde{X}_n)_{n\in\mathbb{N}}$ given by the stochastic difference equation

$$\widetilde{X}_n = \sqrt{\beta}N_{1,n} + (\alpha+\sqrt{\lambda}N_{2,n})\widetilde{X}_{n-1}\,, \tag{2.49}$$

where $(N_{1,n})_{n\in\mathbb{N}}$ and $(N_{2,n})_{n\in\mathbb{N}}$ are two independent sequences of iid normal distributed random variables. The stationary distribution of the process $(\widetilde{X}_n)_{n\in\mathbb{N}}$ follows from Goldie (1991, Corollary 2.4), see also Embrechts et al. (1997), Section 8.4.

2.3 Extremal behaviour of the AR(1) process with ARCH(1) errors

In the present section we study the extremal behaviour of AR processes with ARCH errors. We again focus on the AR(1) process with ARCH(1) errors, i.e. $f(X_{n-1},...,X_{n-k}) = \alpha X_{n-1}$ for some $\alpha \in \mathbb{R}$ and σ_n is given in (0.2) with $p=1$. Our results for the extremes will be an extension of the results in de Haan, Resnick, Rootzén and de Vries (1989).

Extremal behaviour of a Markov process $(X_n)_{n\in\mathbb{N}}$ is for instance manifested in the asymptotic behaviour of the maxima

$$M_n = \max_{1\le k\le n} X_k\,, \quad n \ge 1\,.$$

The limit behaviour of M_n is a well-studied problem in extreme value theory. Two review paper on this and related problems are Rootzén (1988) and Perfekt (1994). For a general overview of extremes of Markov processes, see also Leadbetter, Lindgren and Rootzén (1983), Leadbetter and Rootzén (1988) and the references therein. Loosly speaking, under quite general mixing conditions, one can show that for n and x large

$$P(M_n \leq x) \approx F^{n\theta}(x)\,, \tag{3.1}$$

where F is the stationary distribution function of $(X_n)_{n\in\mathbb{N}}$ and $\theta \in (0,1)$ is a constant called *extremal index*. A natural interpretation of θ is that of the reciprocal of mean cluster size (see e.g. Embrechts, Klüppelberg and Mikosch (1997, Chapter 6) and the references therein). The practical implication of (3.1) is that dependence in data does often not invalidate the application of classical extreme value theory. There are many methods for determing the extremal index. However, most are very technical and often useless in practice. An alternative is then to estimate θ from the data.

For the AR(1) process with ARCH(1) errors we derive an explicit formula for the extremal index. We furthermore investigate the point process of exceedances of a high threshold u of $(X_n)_{n\in\mathbb{N}}$ which characterizes the extremal behaviour of the process in detail. This point process converges in distribution to a compound Poisson process with a well-specified intensity and a well-specified distribution of the size of the jumps.

2.3.1 Preliminaries

In order to study the extremal behaviour of $(X_n)_{n\in\mathbb{N}}$ and $(X_n^2)_{n\in\mathbb{N}}$ we define the auxiliary process $(Z_n)_{n\in\mathbb{N}} := (\ln(Y_n^2))_{n\in\mathbb{N}}$. Since $(Y_n)_{n\in\mathbb{N}}$ follows (2.44) the process $(Z_n)_{n\in\mathbb{N}}$ satisfies the stochastic difference equation

$$Z_n = Z_{n-1} + \ln\Big((\alpha + \sqrt{\beta\, e^{-Z_{n-1}} + \lambda}\,\varepsilon_n)^2\Big)\,, \quad n \in \mathbb{N}, \tag{3.2}$$

where $(\varepsilon_n)_{n\in\mathbb{N}}$ are i.i.d. random variables that satisfy the general conditions and $(D.1) - (D.3)$, the constants are the same as in our old process $(X_n)_{n\in\mathbb{N}}$ and Z_0 equals $\ln(X_0^2)$ a.s.. Note that $(Z_n)_{n\in\mathbb{N}} \stackrel{d}{=} (\ln(X_n^2))_{n\in\mathbb{N}}$ and thus the process $(Z_n)_{n\in\mathbb{N}}$ is again regenerative and strongly mixing. Moreover, $(Z_n)_{n\in\mathbb{N}}$ does not dependent on the sign of the parameter α since ε_n is symmetric. In the following we assume therefore that $\alpha \geq 0$. We will see that $(Z_n)_{n\in\mathbb{N}}$ can be bounded by two random walks $(S_n^{l,a})_{n\in\mathbb{N}}$ and $(S_n^{u,a})_{n\in\mathbb{N}}$ from below and above, respectively. This result together with $(Z_n)_{n\in\mathbb{N}} \stackrel{d}{=} (\ln(X_n^2))_{n\in\mathbb{N}}$ appears to be the key to the description of the extremal behaviour of $(X_n)_{n\in\mathbb{N}}$. Via results for $(Z_n)_{n\in\mathbb{N}}$, we prove for instance that the regenerative process $(X_n)_{n\in\mathbb{N}}$ has finite mean recurrence times which allow us to consider only the extremal behaviour of the stationary process $(X_n)_{n\in\mathbb{N}}$. The process $(Z_n)_{n\in\mathbb{N}}$ will be also important in the proof of Lemma 2.3.7. For the construction of the two random walks $(S_n^{l,a})_{n\in\mathbb{N}}$ and $(S_n^{u,a})_{n\in\mathbb{N}}$ we need some more definitions. With the same notation as before, let

$$A_a := \{\omega \,|\, \frac{-\alpha}{\sqrt{\beta\, e^{-a}+\lambda}-\sqrt{\beta}\, e^{-a/2}} \le \varepsilon(\omega) \le \frac{-\alpha}{\sqrt{\beta\, e^{-a}+\lambda}+\sqrt{\beta}\, e^{-a/2}}\}\,, \tag{3.3}$$

$$p(a,\alpha,\beta,\lambda,\varepsilon) := \ln\Big((\alpha+\sqrt{\beta\, e^{-a}+\lambda}\,\varepsilon)^2\Big)\,,$$

$$q(a,\alpha,\beta,\lambda,\varepsilon) := \ln\Big(1-\frac{2\alpha\sqrt{\beta}e^{-a/2}\varepsilon}{(\alpha+\sqrt{\beta\, e^{-a}+\lambda}\,\varepsilon)^2}1_{\{\varepsilon<0\}}\Big)\,, \tag{3.4}$$

$$r(a,\alpha,\beta,\lambda,\varepsilon) := \ln\Big(1-\frac{\beta\varepsilon^2 e^{-a}}{(\alpha+\sqrt{\beta\, e^{-a}+\lambda}\,\varepsilon)^2}1_{\{\varepsilon<0\}}\Big)\,.$$

Note that $q(a,\alpha,\beta,\lambda,\varepsilon), r(a,\alpha,\beta,\lambda,\varepsilon) \to 0$ a.s. for $a\to\infty$. Now define

$$S_n^{l,a} := \sum_{j=1}^n U_j^a \quad \text{and} \quad S_n^{u,a} := \sum_{j=1}^n V_j^a, \quad n\in\mathbb{N}, \tag{3.5}$$

where

$$\begin{aligned} U_j^a \;\; := \;\; -\infty\cdot 1_{A_a} &+ \Big(p(a,\alpha,\beta,\lambda,\varepsilon_j)+r(a,\alpha,\beta,\lambda,\varepsilon_j)\Big)\cdot 1_{A_a^c\cap\{\varepsilon_j<0\}} \\ &+ \ln(\alpha+\sqrt{\lambda}\varepsilon_j)^2\cdot 1_{\{\varepsilon_j\ge 0\}} \end{aligned} \tag{3.6}$$

and

$$V_j^a := p(a,\alpha,\beta,\lambda,\varepsilon_j)+q(a,\alpha,\beta,\lambda,\varepsilon_j) \tag{3.7}$$

for some $a\ge 0$. The following lemma shows that the random walks defined in (3.5)-(3.7) are really upper and lower bounds for $(Z_n)_{n\in\mathbb{N}}$ above a high level.

Lemma 2.3.1 *Let a be large enough, $N_a := \inf\{j\ge 1\,|\, Z_j\le a\}$ and $Z_0>a$. Then*

$$Z_0+S_k^{l,a}\le Z_k\le Z_0+S_k^{u,a} \quad \textit{for any } k\le N_a \textit{ a.s.} \tag{3.8}$$

Proof. We prove only the lower bound. The proof of the upper bound is similar but easier. Let $x \geq a$ be arbitrary. If $\varepsilon \geq 0$ it is obvious that

$$(\alpha + \sqrt{\beta\, e^{-x} + \lambda}\,\varepsilon)^2 \geq (\alpha + \sqrt{\lambda}\,\varepsilon)^2\,. \tag{3.9}$$

Consider now $\varepsilon < 0$, then

$$\begin{aligned}
&(\alpha + \sqrt{\beta\, e^{-x} + \lambda}\,\varepsilon)^2 - (\alpha + \sqrt{\beta\, e^{-a} + \lambda}\,\varepsilon)^2 \\
&= \; 2\alpha(-\varepsilon)\Big(\sqrt{\beta\, e^{-a} + \lambda} - \sqrt{\beta\, e^{-x} + \lambda}\Big) - \beta(e^{-a} - e^{-x})\,\varepsilon^2 \\
&\geq \; -\beta\, e^{-a}\,\varepsilon^2\,.
\end{aligned} \tag{3.10}$$

Note that we have a non-trivial lower bound of $(\alpha + \sqrt{\beta\, e^{-x} + \lambda}\,\varepsilon)^2$ if and only if

$$(\alpha + \sqrt{\beta\, e^{-a} + \lambda}\,\varepsilon)^2 - \beta\, e^{-a}\,\varepsilon^2 > 0\,. \tag{3.11}$$

It is straightforward that (3.11) is equivalent to

$$\varepsilon > \frac{-\alpha}{\sqrt{\beta\, e^{-a} + \lambda} + \sqrt{\beta}\, e^{-a/2}} \quad \text{or} \quad \varepsilon < \frac{-\alpha}{\sqrt{\beta\, e^{-a} + \lambda} - \sqrt{\beta}\, e^{-a/2}}\,. \tag{3.12}$$

From (3.9), (3.10) and (3.12), we obtain

$$\Big(\alpha + \sqrt{\beta\, e^{-x} + \lambda}\,\varepsilon(\omega)\Big)^2 \geq \begin{cases} (\alpha + \sqrt{\lambda}\,\varepsilon(\omega))^2\,, & \omega \in \{\varepsilon \geq 0\} \\ (\alpha + \sqrt{\beta\, e^{-a} + \lambda}\,\varepsilon(\omega))^2 - \beta\, e^{-a}\,\varepsilon(\omega)^2\,, & \omega \in A_a^c \cap \{\varepsilon < 0\} \\ 0\,, & \omega \in A_a \end{cases} \tag{3.13}$$

Now take logarithms and use the additive structure (3.2) of $(Z_n)_{n\in\mathbb{N}}$. $\square$

Remark 2.3.2 (a) If a is large enough then $S_n^{u,a}$ and $S_n^{l,a}$ are random walks with negative drift.

Proof. Note that

$$\begin{aligned}
E(V_1^a) &= E\left(p(a,\alpha,\beta,\lambda,\varepsilon_1) + q(a,\alpha,\beta,\lambda,\varepsilon_1)\right) \\
&= E\Big(\ln\Big((\alpha + \sqrt{\beta\, e^{-a} + \lambda}\,\varepsilon_1)^2 + 2\alpha\sqrt{\beta}\, e^{-a/2}\,(-\varepsilon_1)\, 1_{\{\varepsilon_1 < 0\}}\Big)\Big) \\
&\to E(\ln(\alpha + \sqrt{\lambda}\varepsilon_1)^2) < 0\,, \quad \text{as } a \to \infty\,,
\end{aligned}$$

where we used the dominated convergence theorem and (2.2) in the last step . Hence for a large enough the statement follows. $\square$

(b) Let $(S_n)_{n\in\mathbb{N}} := \left(\sum_{j=1}^n \ln\Big((\alpha+\sqrt{\lambda}\varepsilon_j)^2\Big)\right)_{n\in\mathbb{N}}$. For $a\uparrow\infty$ we have

$$S_k^{l,a} \xrightarrow{P} S_k \quad \text{and} \quad S_k^{u,a} \xrightarrow{a.s.} S_k\,, \tag{3.14}$$

for any $k\in\mathbb{N}$, i.e. both random walks converge at least in probability to the same random walk. Furthermore,

$$\sup_{k\geq 1} S_k^{l,a} \xrightarrow{d} \sup_{k\geq 1} S_k \quad \text{and} \quad \sup_{k\geq 1} S_k^{u,a} \xrightarrow{a.s.} \sup_{k\geq 1} S_k\,. \tag{3.15}$$

Proof. The a.s. convergence of $(S_n^{u,a})_{n\in\mathbb{N}}$ and $\sup_{k\geq 1} S_k^{u,a}$ is straightforward since p,q and r converge a.s.. Consider therefore the lower random walk $(S_n^{l,a})_{n\in\mathbb{N}}$. Note that for $a\uparrow\infty$

$$P(A_a)\to 0$$

and hence

$$1_{A_a^c\cap\{\varepsilon<0\}} \xrightarrow{P} 1_{\{\varepsilon<0\}} \quad \text{and} \quad 1_{A_a\cap\{\varepsilon<0\}} \xrightarrow{P} 0\,. \tag{3.16}$$

Furthermore,

$$p(a,\alpha,\beta,\lambda,\varepsilon_1) + r(a,\alpha,\beta,\lambda,\varepsilon_1) \xrightarrow{a.s.} \ln\left((\alpha+\sqrt{\lambda}\varepsilon_1)^2\right)\,, \tag{3.17}$$

and therefore (3.14) holds. Finally we note that

$$\begin{aligned} E\max(0,U_1^a) &= E\max\Big(0,\Big(p(a,\alpha,\beta,\lambda,\varepsilon_1)+r(a,\alpha,\beta,\lambda,\varepsilon_1)\Big)\,1_{A_a^c\cap\{\varepsilon_1<0\}}\Big) \\ &\qquad + E\max\Big(0,\ln(\alpha+\sqrt{\lambda}\varepsilon_1)^2\,1_{\{\varepsilon_1\geq 0\}}\Big) \\ &\to E\max(0,\ln(\alpha+\sqrt{\lambda}\varepsilon_1)^2), \quad \text{as } a\to\infty\,, \end{aligned} \tag{3.18}$$

where we used (3.16), (3.17) and the dominated convergence theorem. By Borovkov (1976) , Theorem 22, p.53, (3.14) and (3.18) we derive that

$$\sup_{k\geq 1} S_k^{l,a} \xrightarrow{d} \sup_{k\geq 1} S_k\,.$$

□

Lemma 2.3.1 characterizes the behaviour of the process $(Z_n)_{n\in\mathbb{N}}$ above a high treshold a and hence also the behaviour of $(X_n^2)_{n\in\mathbb{N}}$. This is the key to what follows: the process $(S_n)_{n\in\mathbb{N}}$ will determine completely the extremal behaviour of (X_n^2). Recall from Theorem 2.2.3 that $(X_n)_{n\in\mathbb{N}}$ is Harris recurrent with regeneration set $[-e^{a/2}, e^{a/2}]$ for a large enough. Thus there exists in particular a renewal point process T_0, T_1, T_2, ... which describes the regenerative structure of $(X_n)_{n\in\mathbb{N}}$.

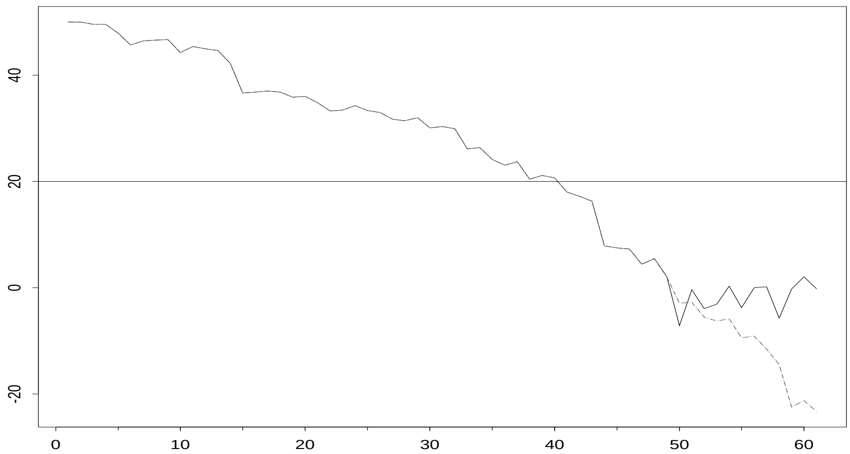

Figure 2.3: Simulated sample path of $(Z_n)_{n\in\mathbb{N}}$ with parameters $\alpha = 0.6, \beta = 1, \lambda = 0.4$ and starting point $Z_0 = 50$ (solid line) and the corresponding random walks $(S_n^{l,a})_{n\in\mathbb{N}}$ and $(S_n^{u,a})_{n\in\mathbb{N}}$ with $a = 20$ (dotted lines), respectively. Note that the random walks are hardly distinguishable from each other and $(Z_n)_{n\in\mathbb{N}}$ for $n \leq 47$. Hence they are extremely good bounds above the level $a = 20$. If the process falls far below the level 20 they are still very close, but are no longer bounds for $(Z_n)_{n\in\mathbb{N}}$. The picture also confirms our statement that the random walks have negative drift and converge to the same limit.

Corollary 2.3.3 *The renewal point process $(T_n)_{n\in\mathbb{N}_0}$ which describes the regenerative structure of $(X_n)_{n\in\mathbb{N}}$ is aperiodic and has finite mean recurrence times $C_0 = T_0$ and $C_1 = T_1 - T_0$.*

Proof. The renewal process can be constructed in the following way (see e.g. Asmussen (1987), Section VI.3 for some background on regenerative processes):
Define

$$\tau_1 := \inf\{k \geq 1 \,|\, X_k \in [-e^{a/2}, e^{a/2}]\} \stackrel{d}{=} \inf\{k \geq 1 \,|\, Z_k \leq a\} = N_a$$

and $\tau_{i+1} := \inf\{k > \tau_i \,|\, X_k \in [-e^{a/2}, e^{a/2}]\} \stackrel{d}{=} \inf\{k > \tau_i \,|\, Z_k \leq a\}$ for $i = 1, 2, 3, \ldots$. Since, above level a, $(Z_n)_{n\in\mathbb{N}}$ is dominated by the random walk with negative drift $(S_n^{u,a})_{n\in\mathbb{N}}$ and

$$\sup_{x\in(-\infty,a]} E(\max(0, Z_1) \,|\, Z_0 = x) < \infty\,, \tag{3.19}$$

it follows that $\tau_1, \tau_2, \tau_3, \ldots$ are well defined and have finite expectations. Now let $M_1 := \inf\{i \geq 1 \,|\, I_{\tau_i} = 1\}$ and $M_{j+1} := \inf\{i > M_j \,|\, I_{\tau_i} = 1\}$ for $j = 1, 2, 3, \ldots$ with $P(I_1 = 1) = 1 - P(I_1 = 0) = \widetilde{C}$ and independent of $(X_n)_{n\in\mathbb{N}}$ where $\widetilde{C}$ is the constant in (2.11). Note that

$$P(M_j - M_{j-1} = i) = \widetilde{C}(1 - \widetilde{C})^{i-1} \quad \text{for } i, j = 1, 2, \ldots \text{ and } M_0 = 0\,. \tag{3.20}$$

From Asmussen (1987), p.151 and (2.11), the renewal process $(T_n)_{n\geq 0}$ is now given by

$$T_n := \tau_{M_{n+1}} + 1\,, \quad n \geq 0\,,$$

and hence, by (3.20)

$$E(C_0) = E(T_0) \leq E(\tau_{M_1+1}) \leq const\, E(M_1 + 1) < \infty\,.$$

Similar calculation shows that $E(C_1) < \infty$ as well. Since the transition density of $(Z_n)_{n\in\mathbb{N}}$ is positive and continuous it follows finally that C_1 is aperiodic. □

As a consequence of Corollary 2.3.4 we may suppose in the following that the process $(X_n)_{n\in\mathbb{N}}$ is stationary.

Corollary 2.3.4 *For any probability measure μ and any sequence $(u_n)_{n\in\mathbb{N}}$*

$$\left|P^\mu\Big(\max_{1\leq k\leq n} X_k \leq u_n\Big) - P^\pi\Big(\max_{1\leq k\leq n} X_k \leq u_n\Big)\right| \to 0, \quad \text{as } n\to\infty\,,$$

where P^μ denotes the probability law for $(X_n)_{n\in\mathbb{N}}$ when X_0 starts with distribution μ and π is the stationary distribution.

Proof. The proof invokes a coupling argument. Let $X = (X_n)_{n\in\mathbb{N}}$ be the AR(1) process with ARCH(1) errors with arbitrary initial probability μ and let $X' = (X'_n)_{n\in\mathbb{N}}$ be a parallel process, governed by the same transition probabilities as and independent of X, and with initial distribution π. Now define T as the first common renewal time of X and X', i.e.

$$T := \inf\{\, n \in \mathbb{N} \,|\, T_n = T'_n \,\}\,.$$

From Asmussen (1987) and (2.11) we get in particular that

$$X_T \stackrel{d}{=} X'_T \sim \nu\,. \tag{3.21}$$

Defining $M_{l,r} := \max_{l\leq j\leq r} X_j$ and $M_r := M_{1,r}$ we get that

$$\begin{aligned}
&\left|P\Big(\max_{1\leq k\leq n} X_k \leq u_n\Big) - P\Big(\max_{1\leq k\leq n} X'_k \leq u_n\Big)\right| \\
&= \left|E\Big(P(M_n \leq u_n \,|\, T, M_T, M'_T) - P(M'_n \leq u_n \,|\, T, M_T, M'_T)\Big)\right| \\
&\leq \left|E\Big(1_{\{T\leq n, M_T\leq M_n, M'_T\leq M'_n\}}\Big(P(M_n \leq u_n \,|\, T, M_T, M'_T) - P(M'_n \leq u_n \,|\, T, M_T, M'_T)\Big)\Big)\right| \\
&\quad + \left|E\Big(1_{\{M_T > M_n\}\cup\{M'_T > M'_n\}}\Big(P(M_n \leq u_n \,|\, T, M_T, M'_T) - P(M'_n \leq u_n \,|\, T, M_T, M'_T)\Big)\Big)\right|
\end{aligned}$$

$$
\begin{aligned}
&= \Big|E\Big(1_{\{T\le n, M_T\le M_n, M'_T\le M'_n\}}\Big(P(M_{T,n}\le u_n\,|\,T,M_T,M'_T) - P(M'_{T,n}\le u_n\,|\,T,M_T,M'_T)\Big)\Big)\Big| \\
&\quad +\Big|E\Big(1_{\{M_T>M_n\}\cup\{M'_T>\le M'_n\}}\Big(P(M_n\le u_n\,|\,T,M_T,M'_T) - P(M'_n\le u_n\,|\,T,M_T,M'_T)\Big)\Big)\Big| \\
&= \Big|E\Big(1_{\{T\le n, M_T\le M_n, M'_T\le M'_n\}}\cdot 0\Big)\Big| \\
&\quad +\Big|E\Big(1_{\{M_T>M_n\}\cup\{M'_T>\le M'_n\}}\Big(P(M_n\le u_n\,|\,T,M_T,M'_T) - P(M'_n\le u_n\,|\,T,M_T,M'_T)\Big)\Big)\Big| \\
&\le E\Big(1_{\{M_T>M_n\}\cup\{M'_T>\le M'_n\}}\Big|P(M_n\le u_n\,|\,T,M_T,M'_T) - P(M'_n\le u_n\,|\,T,M_T,M'_T)\Big|\Big) \\
&\le 2P(\{M_T>M_n\}\cup\{M'_T>M'_n\}) \\
&\le 2E\Big(P(M_T>M_n\,|\,T) + P(M'_T>M'_n\,|\,T)\Big) \\
&= 2E\Big(1_{\{T>n\}}\Big(P(M_T>M_n\,|\,T) + P(M'_T>M'_n\,|\,T)\Big)\Big) \\
&\le 4P(T>n)\,,
\end{aligned}
$$

where we used in the forth line that $\{T>n\}\subseteq\{M_T>M_n\}$ and in the seventh line (3.21) and the Markov structure of X and X'. Hence, if we prove that T is almost sure finite we are finished. But by Corollary 2.3.4 the process $(X_n)_{n\in\mathbb{N}}$ is regenerative and the embedded renewal process is aperiodic and has finite mean recurrence time. From Lindvall (1992), p.23 the statement follows.

□

2.3.2 Limit distribution of the normalised maximum and cluster probabilities of the exceedances

In this section we present the main results concerning the extremal behaviour of the AR(1) process with ARCH(1) errors and the associated squared process. Let $(\widehat{X}_n)_{n\in\mathbb{N}}$ be the associated independent process of $(X_n)_{n\in\mathbb{N}}$, i.e. $\widehat{X}_1, \widehat{X}_2, ...$ are i.i.d. random variables with the stationary distribution function of $(X_n)_{n\in\mathbb{N}}$. From (2.17), Corollary 2.2.19 and classical extreme value theory we obtain

$$
\lim_{n\to\infty} P(n^{-1/\kappa}\max_{1\le k\le n}\widehat{X}_k\le x) = \exp(-c\,x^{-\kappa})\,,\quad x\ge 0\,, \tag{3.22}
$$

hence the maximum of the associated independent process $(\widehat{X}_n)_{n\in\mathbb{N}}$ belongs to the domain of attraction of a Fréchet distribution. In the dependent case we prove a similar result. The limit distribution is still a Fréchet distribution but a constant θ occurs in the exponent. θ is called the *extremal index* of the process $(X_n)_{n\in\mathbb{N}}$ and is a measure of local dependence amongst the exceedances over a high threshold by the process $(X_n)_{n\in\mathbb{N}}$. It has a natural interpretation as the reciprocal of the mean cluster size. In order to describe the extremes in more detail, we also consider the point process $(N_n)_{n\in\mathbb{N}}$ of exceedances of an appropriately chosen high threshold u_n given by

$$N_n(\cdot) := \#\{k/n \in \cdot \,|\, X_k > u_n\,,\, k \in \{1,...,n\}\,\} \tag{3.23}$$

and show that this point process converges to a compound Poisson process N. We derive the intensity and the distribution of the jumps which we denote by $(\pi_k)_{k\in\mathbb{N}}$. Note that in the extreme value theory for strong mixing processes the jumps equal the lengths of clusters of exceedances. For further background we refer to Leadbetter et al. (1983), Rootzén (1988) or Embrechts et al. (1997, Section 8.1). For the ARCH(1) process it was convenient to investigate first the squared process. This is not the case for our model since we have a completely different structure due to the autoregressive part of $(X_n)_{n\in\mathbb{N}}$. Nevertheless, only for the squared process $(X_n^2)_{n\in\mathbb{N}}$ a comparison with results in the ARCH(1) case (see de Haan et al. (1989)) is possible. The following theorem collects our results.

Theorem 2.3.5 *(a) Suppose $(X_n)_{n\in\mathbb{N}}$ is given by equation (1.1) with $(\varepsilon_n)_{n\in\mathbb{N}}$ satisfying the general conditions (1.2) and $(D.1)-(D.3)$ with parameters α and λ satisfying (2.2) and $X_0 \sim \mu$. Then*

$$\lim_{n\to\infty} P^\mu(n^{-1/\kappa} \max_{1\le j\le n} X_j \le x) = \exp(-c\theta x^{-\kappa})\,, \quad x \ge 0\,, \tag{3.24}$$

where P^μ denotes the law for $(X_n)_{n\in\mathbb{N}}$ when X_0 starts with the distribution μ, κ solves the equation $E(|\alpha+\lambda\varepsilon|^\kappa) = 1$, c is defined by (2.45) and

$$\theta = \kappa \int_1^\infty P(\sup_{k\ge1} \prod_{i=1}^k (\alpha+\sqrt{\lambda}\varepsilon_i) \le y^{-1}) y^{-\kappa-1} dy\,.$$

For $x\in\mathbb{R}$, let N_n be the point process of exceedances of the threshold $u_n = n^{1/\kappa}x$ by $X_1,...,X_n$ given by (3.23). Then

$$N_n \quad \xrightarrow{d} \quad N, \quad n\to\infty\,,$$

where N is a compound Poisson process with intensity $c\theta x^{-\kappa}$ and cluster probabilities

$$\pi_k = \frac{\theta_k - \theta_{k+1}}{\theta}\,, \quad k\in\mathbb{N}, \tag{3.25}$$

where

$$\theta_k = \kappa \int_1^\infty P(\#\{j\ge1 \,|\, \prod_{i=1}^j (\alpha+\sqrt{\lambda}\varepsilon_i) > y^{-1}\} = k-1) y^{-\kappa-1} dy\,, \quad k\in\mathbb{N}.$$

In particular, $\theta_1 = \theta$.

(b) Let $(X_n)_{n\in\mathbb{N}}$ *be the AR(1) process with ARCH(1) errors in (a) and* $(X_n^2)_{n\in\mathbb{N}}$ *the squared process. Then*

$$\lim_{n\to\infty} P^\mu(n^{-2/\kappa} \max_{1\le j\le n} X_j^2 \le x) = \exp(-2c\theta^{(2)}x^{-\kappa/2})\,, \quad x \ge 0\,, \tag{3.26}$$

where κ, c *are the same constants as in* (a) *and*

$$\theta^{(2)} = \frac{\kappa}{2}\int_1^\infty P(\sup_{k\ge 1}\prod_{i=1}^{k}(\alpha+\sqrt{\lambda}\varepsilon_i)^2 \le y^{-1})y^{-\frac{\kappa}{2}-1}dy\,.$$

For $x \in \mathbb{R}$*, let* $N_n^{(2)}$ *be the point process of exceedances of the threshold* $u_n = n^{2/\kappa}x$ *by* $X_1^2, \ldots, X_n^2$*. Then*

$$N_n^{(2)} \quad \stackrel{d}{\to} \quad N^{(2)}, \quad n \to \infty\,,$$

where $N^{(2)}$ *is a compound Poisson process with intensity* $2c\theta^{(2)}x^{-\kappa/2}$ *and cluster probabilities*

$$\pi_k^{(2)} = \frac{\theta_k^{(2)} - \theta_{k+1}^{(2)}}{\theta^{(2)}}\,, \quad k \in \mathbb{N}\,, \tag{3.27}$$

where

$$\theta_k^{(2)} = \frac{\kappa}{2}\int_1^\infty P(\#\{j\ge 1 \,|\, \prod_{i=1}^{j}(\alpha+\sqrt{\lambda}\varepsilon_i)^2 > y^{-1}\} = k-1)y^{-\frac{\kappa}{2}-1}dy\,, \quad k\in\mathbb{N}.$$

In particular, $\theta_1^{(2)} = \theta^{(2)}$.

Remark 2.3.6 (a) Theorem 2.3.5 is a generalisation of the result of de Haan et al. (1989) in the ARCH(1) case (i.e. $\alpha = 0$). They use a different approach which does not extend to the general case because of the autoregressive part of $(X_n)_{n\in\mathbb{N}}$.

(b) Note that for the squared process one can describe the extremal index and the cluster probabilities by the random walk $(S_n)_{n\in\mathbb{N}}$, namely

$$\theta_k^{(2)} = \frac{\kappa}{2}\int_0^\infty P(\#\{j\ge 1 \,|\, S_j > -x\} = k-1)\, e^{-\frac{\kappa}{2}x}\,dx\,, \quad k\in\mathbb{N}.$$

The description of the extremal behaviour of $(X_n^2)_{n\in\mathbb{N}}$ by the random walk $(S_n)_{n\in\mathbb{N}}$ is to be expected since by Lemma 2.3.1 and Remark 2.3.2 the process $(Z_n)_{n\in\mathbb{N}} = (\ln(X_n^2))_{n\in\mathbb{N}}$ behaves above a high threshold asymptotically like $(S_n)_{n\in\mathbb{N}}$. Unfortunately, this link fails for $(X_n)_{n\in\mathbb{N}}$. Another possibility for proving statement (b) is to follow the work of Hooghiemstra and

Meester (1995) using the regenerative structure of $(Z_n)_{n\in\mathbb{N}}$, Lemma 2.3.1, Remark 2.3.2(b) and Corollary 2.3.4.

(c) Analogous to de Haan et al. (1989) we may construct "estimators" for the extremal indices $\theta^{(2)}$ and $\theta_k^{(2)}$ of $(X_n^2)_{n\in\mathbb{N}}$, respectively, by

$$\widehat{\theta}^{(2)} = \frac{1}{N}\sum_{i=1}^{N} 1_{\{\sup_{1\le j\le m} S_j^{(i)} \le -E_{\kappa/2}^{(i)}\}}$$

and

$$\widehat{\theta}_k^{(2)} = \frac{1}{N}\sum_{i=1}^{N} 1_{\{\sum_{j=1}^{m} 1_{\{S_j^{(i)} > -E_{\kappa/2}^{(i)}\}} = k-1\}}\,, \quad \text{for } k \in \mathbb{N},$$

where N denotes the number of simulated sample paths of $(S_n)_{n\in\mathbb{N}}$, $E_{\kappa/2}^{(i)}$ are i.i.d. exponential random variables with intensity κ and m is chosen large enough. These estimators can be studied as in the case $\alpha = 0$ and $\varepsilon \sim N(0,1)$ in de Haan et al. (1989). In particular,

$$\frac{\theta^{(2)} - \widehat{\theta}^{(2)}}{(\theta^{(2)}(1-\theta^{(2)})/N)^{1/2}}$$

is approximately $N(0,1)$ distributed. Because of Remark 2.3.6(b) this approach is not possible for $(X_n)_{n\in\mathbb{N}}$. We choose as "estimators" for θ and θ_k for $(X_n)_{n\in\mathbb{N}}$

$$\widehat{\theta} = \frac{1}{N}\sum_{i=1}^{N} 1_{\{\sup_{1\le j\le m}\prod_{l=1}^{j}(\alpha+\sqrt{\lambda}\,\varepsilon_l^{(i)}) \le 1/P_\kappa^{(i)}\}} \tag{3.28}$$

and

$$\widehat{\theta}_k = \frac{1}{N}\sum_{i=1}^{N} 1_{\{\sum_{j=1}^{m} 1_{\{\prod_{l=1}^{j}(\alpha+\sqrt{\lambda}\,\varepsilon_l^{(i)}) > 1/P_\kappa^{(i)}\}} = k-1\}}\,, \quad \text{for } k \in \mathbb{N}, \tag{3.29}$$

where N denotes the number of simulated paths of $(\prod_{l=1}^{n}(\alpha+\sqrt{\lambda}\,\varepsilon_l))_{n\in\mathbb{N}}$, $P_\kappa^{(i)}$ are i.i.d. Pareto-distributed random variables with intensity κ, i.e. with distribution function $G(x) = 1 - x^{-\kappa}$, $x \ge 0$, and m is large enough. These are suggestive estimators since $\prod_{l=1}^{n}(\alpha+\sqrt{\lambda}\,\varepsilon_l) \to 0$ a.s. as $n \to \infty$ because of assumption (2.2).

(d) Note that the extremal index θ of $(X_n)_{n\in\mathbb{N}}$ is not symmetric in the parameter α (see Table 2.4). This observation is intuitively obvious since for $\alpha > 0$ the clustering is stronger by the autoregressive part than for $\alpha < 0$.

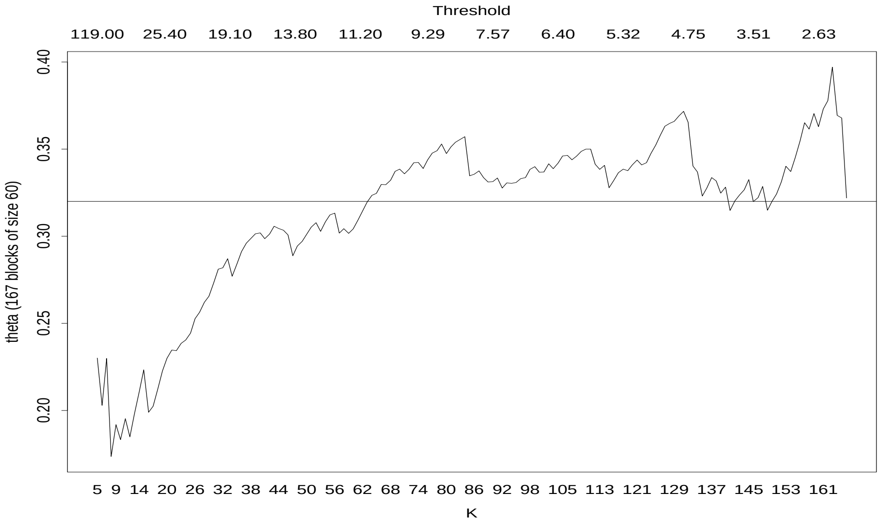

Figure 2.4: Estimated extremal index of a simulated sample path of $(X_n)_{0\le n\le 10000}$ with parameters $\alpha = 0.8, \beta = 1, \lambda = 0.6$ and $\varepsilon \sim N(0,1)$ using the blocks method for the data (see Embrechts et al. (1997), Section 8.1) . The length of a block is chosen as 60. The solid line is the numerically computed extremal index using (3.28), see also Table 2.4.

α \ λ	0.2	0.4	0.6	0.8	1.0	1.2	1.5	2.0	2.5	3.0	3.5
-1.2	-	0.001	0.001	0.003	0.004	0.001	0.000	-	-	-	-
-1	0.15	0.19	0.19	0.16	0.13	0.09	0.05	0.01	-	-	-
-0.8	0.56	0.47	0.41	0.34	0.26	0.21	0.13	0.05	0.01	-	-
-0.6	0.86	0.71	0.61	0.50	0.41	0.33	0.22	0.10	0.03	0.00	-
-0.4	0.96	0.85	0.71	0.60	0.50	0.40	0.30	0.14	0.06	0.01	-
-0.2	0.98	0.89	0.77	0.65	0.56	0.47	0.33	0.18	0.07	0.02	0.00
0	0.98	0.89	0.78	0.65	0.55	0.45	0.33	0.18	0.08	0.02	0.00
0.2	0.94	0.82	0.72	0.61	0.52	0.43	0.32	0.18	0.07	0.02	0.00
0.4	0.85	0.72	0.63	0.53	0.45	0.37	0.28	0.13	0.06	0.01	-
0.6	0.68	0.55	0.48	0.41	0.35	0.29	0.21	0.10	0.03	0.00	-
0.8	0.39	0.34	0.32	0.27	0.22	0.19	0.12	0.05	0.01	-	-
1.0	0.09	0.14	0.13	0.13	0.11	0.08	0.04	0.01	-	-	-
1.2	-	0.000	0.001	0.003	0.004	0.001	0.000	-	-	-	-

Table 2.4: "Estimated" extremal index θ of $(X_n)_{n\in\mathbb{N}}$ in the case $\varepsilon \sim N(0,1)$. We chose $N = m = 2000$. Note that the extremal index decreases as $|\alpha|$ increases and that we have no symmetry in α.

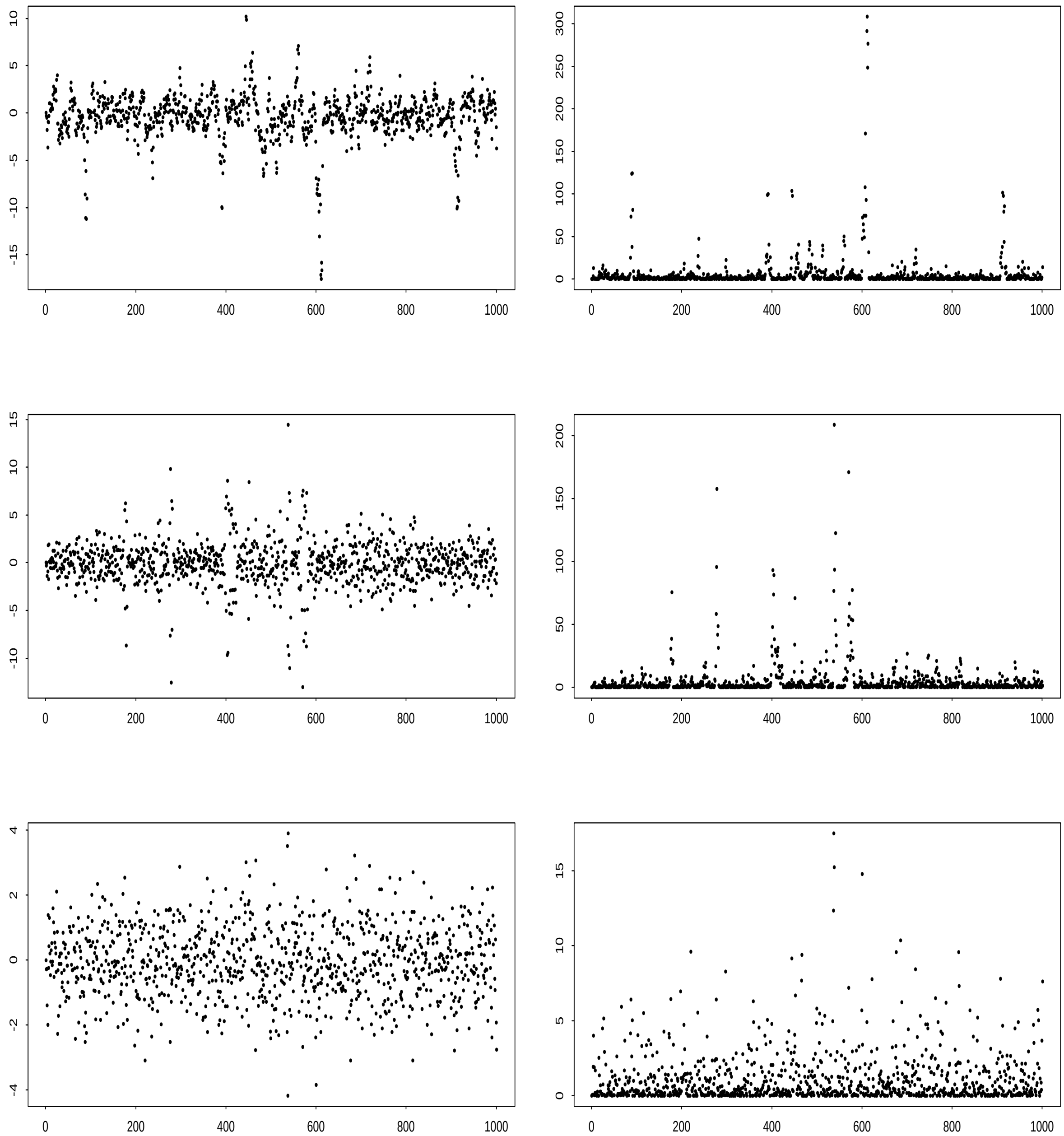

Figure 2.5: Simulated sample path of $(X_n)_{n\in\mathbb{N}}$ with parameters $\alpha = 0.8, \beta = 1, \lambda = 0.2$ (top, left), of $(X_n^2)_{n\in\mathbb{N}}$ with the same parameters (top, right), of $(X_n)_{n\in\mathbb{N}}$ with parameters $\alpha = -0.8, \beta = 1, \lambda = 0.2$ (middle, left), of $(X_n^2)_{n\in\mathbb{N}}$ with the same parameters (middle, right), of $(X_n)_{n\in\mathbb{N}}$ with parameters $\alpha = 0, \beta = 1, \lambda = 0.2$ (bottom, left) and of $(X_n^2)_{n\in\mathbb{N}}$ with the same parameters (bottom,right) in the case $\varepsilon \sim N(0, 1)$. All simulations are based on the same simulated noise sequence $(\varepsilon_n)_{n\in\mathbb{N}}$.

$\|\alpha\|$ \ λ	0.2	0.4	0.6	0.8	1.0	1.2	1.5	2.0	2.5	3.0	3.5
0	0.95	0.80	0.65	0.52	0.41	0.31	0.22	0.11	0.04	0.01	0.00
0.2	0.94	0.77	0.62	0.49	0.38	0.31	0.22	0.10	0.04	0.01	0.00
0.4	0.84	0.67	0.55	0.43	0.35	0.26	0.19	0.08	0.03	0.01	-
0.6	0.67	0.52	0.41	0.34	0.25	0.18	0.14	0.06	0.02	0.00	-
0.8	0.38	0.31	0.26	0.20	0.16	0.13	0.08	0.03	0.00	-	-
1.0	0.09	0.12	0.11	0.10	0.07	0.05	0.03	0.01	-	-	-
1.2	-	0.000	0.001	0.001	0.000	0.000	0.000	-	-	-	-

Table 2.5: "Estimated" extremal index $\theta^{(2)}$ of $(X_n^2)_{n\in\mathbb{N}}$ dependent on $|\alpha|$ and λ in the case $\varepsilon \sim N(0,1)$. We chose $N = m = 2000$. Note that the extremal index decreases as $|\alpha|$ increases.

2.3.3 Proof of Theorem 2.3.5

The proof of Theorem 2.3.5 will be an application of results in Perfekt (1994) (see also Appendix A3). In order to apply these results we need to check the assumptions in Theorem A3.1 and A3.2. The next lemma provides a technical property for the squared AR(1) process with ARCH(1) errors $(X_n^2)_{n\in\mathbb{N}}$. It is the most restrictive assumption in Perfekt (1994).

Lemma 2.3.7 *Let $(p_n)_{n\in\mathbb{N}}$ be an increasing sequence such that*

$$\frac{p_n}{n} \to 0 \quad \textit{and} \quad \frac{n\gamma(\sqrt{p_n})}{p_n} \to 0 \quad \textit{as } n \to \infty\,, \tag{3.30}$$

where γ is the mixing function of $(X_n)_{n\in\mathbb{N}}$, i.e. for any $m \in \mathbb{N}$

$\gamma(m) = \sup\{\,|P(A\cap B) - P(A)\,P(B)| \,:\, A \in \sigma(X_j,\, 1 \le j \le k),\, B \in \sigma(X_j,\, j \ge k+m),\, k \in \mathbb{N}\}$.

Then for $u_n = n^{2/\kappa}x$

$$\lim_{p\to\infty} \limsup_{n\to\infty} P(\max_{p\le j\le p_n} X_j^2 > u_n \,|\, X_0^2 > u_n) = 0\,. \tag{3.31}$$

Remark 2.3.8 (a) The strong mixing condition is a property of the underlying $\sigma-$field of a process. Hence γ is also the mixing function of $(X_n^2)_{n\in\mathbb{N}}$ and $(Z_n)_{n\in\mathbb{N}}$ and we may work in all these cases with the same sequence $(p_n)_{n\in\mathbb{N}}$. Note that because of Theorem 2.2.3(b) there exist constants $\eta \in (0,1)$ and $c > 0$ such that $\gamma(m) \le c\,\rho^m$ for any $m \in \mathbb{N}$.

(b) In the case of a strong mixing process, conditions (3.30) are sufficient to guarantee that $(p_n)_{n\in\mathbb{N}}$ is a $\Delta(u_n)$-separating sequence. This is a straightforward consequence of the fact that

α	λ	θ	π_1	π_2	π_3	π_4	π_5	π_6
0	0.2	0.974	0.973	0.027	0.000	0.000	0.000	0.000
0	0.4	0.889	0.895	0.088	0.013	0.003	0.001	0.000
0	0.6	0.781	0.799	0.147	0.036	0.012	0.005	0.001
0	0.8	0.664	0.702	0.175	0.087	0.013	0.011	0.009
0	1	0.549	0.607	0.188	0.107	0.036	0.034	0.017
-0.4	0.2	0.962	0.962	0.037	0.001	0.000	0.000	0.000
0.4	0.2	0.853	0.867	0.103	0.026	0.002	0.002	0.000
-0.4	0.4	0.837	0.860	0.110	0.024	0.006	0.001	0.000
0.4	0.4	0.717	0.734	0.186	0.048	0.018	0.009	0.001
-0.4	0.6	0.715	0.747	0.168	0.048	0.026	0.006	0.002
0.4	0.6	0.624	0.676	0.182	0.066	0.040	0.019	0.012
-0.4	0.8	0.595	0.623	0.220	0.097	0.018	0.016	0.014
0.4	0.8	0.539	0.611	0.167	0.111	0.045	0.036	0.018
-0.4	1	0.497	0.540	0.210	0.115	0.075	0.040	0.004
0.4	1	0.445	0.533	0.185	0.080	0.109	0.032	0.017
-0.8	0.2	0.572	0.626	0.185	0.111	0.026	0.033	0.001
0.8	0.2	0.386	0.470	0.172	0.148	0.062	0.068	0.006
-0.8	0.4	0.488	0.559	0.193	0.107	0.067	0.020	0.016
0.8	0.4	0.331	0.429	0.184	0.099	0.066	0.62	0.057
-0.8	0.6	0.414	0.520	0.159	0.134	0.072	0.043	0.016
0.8	0.6	0.314	0.443	0.156	0.110	0.087	0.073	0.041
-0.8	0.8	0.338	0.392	0.219	0.130	0.090	0.053	0.030
0.8	0.8	0.266	0.358	0.158	0.132	0.140	0.068	0.000
-0.8	1	0.273	0.429	0.137	0.126	0.106	0.016	0.012
0.8	1	0.224	0.346	0.132	0.114	0.129	0.045	0.004

Table 2.6: "Estimated" extremal index θ and cluster probabilities $(\pi_k)_{1\le k\le 6}$ of $(X_n)_{n\in\mathbb{N}}$ dependent on α and λ in the case $\varepsilon \sim N(0,1)$. We chose $N = m = 2000$. Note that the extremal index for $\alpha > 0$ is much larger than for $\alpha < 0$.

$\|\alpha\|$	λ	θ	π_1	π_2	π_3	π_4	π_5	π_6
0	0.2	0.954	0.959	0.037	0.004	0.000	0.000	0.000
0	0.4	0.803	0.819	0.137	0.029	0.014	0.001	0.000
0	0.6	0.651	0.682	0.186	0.092	0.018	0.010	0.008
0	0.8	0.521	0.578	0.215	0.103	0.036	0.027	0.019
0	1	0.406	0.455	0.233	0.135	0.054	0.044	0.023
0.4	0.2	0.844	0.853	0.122	0.018	0.004	0.002	0.001
0.4	0.4	0.664	0.686	0.203	0.069	0.026	0.008	0.004
0.4	0.6	0.553	0.610	0.201	0.095	0.054	0.015	0.008
0.4	0.8	0.423	0.506	0.219	0.084	0.074	0.028	0.023
0.4	1	0.342	0.431	0.216	0.107	0.066	0.045	0.023
0.8	0.2	0.378	0.445	0.184	0.159	0.071	0.057	0.011
0.8	0.4	0.309	0.423	0.143	0.131	0.097	0.060	0.018
0.8	0.6	0.255	0.328	0.202	0.145	0.088	0.012	0.045
0.8	0.8	0.208	0.301	0.186	0.092	0.077	0.077	0.048
0.8	1	0.152	0.237	0.178	0.099	0.092	0.053	0.010

Table 2.7: "Estimated" extremal index $\theta^{(2)}$ and cluster probabilities $(\pi_k^{(2)})_{1\le k\le 6}$ of $(X_n^2)_{n\in\mathbb{N}}$ dependent on α and λ in the case $\varepsilon \sim N(0,1)$. We chose $N = m = 2000$.

$\sigma(\{X_j \le u_n\}, 1 \le j \le k) \subseteq \sigma(X_j, 1 \le j \le k)$, $\sigma(\{X_j \le u_n\}, j \ge l_n + k) \subseteq \sigma(X_j, j \ge l_n + k)$ and choosing additionally $l_n = \sqrt{p_n}$. The notion of a $\Delta(u_n)-$separating sequence was first introduced by O'Brian (1989) and describes somehow the interval length needed to accomplish asymptotic independence of extremal events over a high level u_n in separate intervals. For a definition see also Perfekt (1994). Note that $(p_n)_{n\in\mathbb{N}}$ is in the case of a strong mixing process independent of $(u_n)_{n\in\mathbb{N}}$.

Proof. Note that

$$\begin{aligned} P(\max_{p\le j\le p_n} X_j^2 > u_n \,|\, X_0^2 > u_n) &= P(\tau_1 < p,\, \max_{p\le j\le p_n} X_j^2 > u_n \,|\, X_0^2 > u_n) \\ &\quad + P(p \le \tau_1 < p_n,\, \max_{p\le j\le p_n} X_j^2 > u_n \,|\, X_0^2 > u_n) \\ &\quad + P(\tau_1 \ge p_n,\, \max_{p\le j\le p_n} X_j^2 > u_n \,|\, X_0^2 > u_n) \\ &=: I_1 + I_2 + I_3 , \end{aligned} \tag{3.32}$$

where $\tau_1 = \inf\{j \ge 1 \,|\, X_j^2 \le e^a\} \overset{d}{=}\in \{j \ge 1 \,|\, Z_j \le a\} = N_a$ as in Lemma 2.3.1 and Corollary 2.3.4. In order to get upper bounds of I_1, I_2 and I_3 we show first that there exist constants $C > 0$ and $N \in \mathbb{N}$ such that for any $n > N$, $x \in [e^{-n}, e^a]$ and $k \in \mathbb{N}$

$$n\, P(X_k^2 > u_n \,|\, X_0^2 = x) \le C . \tag{3.33}$$

Assume that (3.33) does not hold. Choose C, $N > 0$ arbitrary and $\eta > 0$ small. Because of the continuity of the transition probability (i.e. equicontinuity on compact sets) , there exist $n > N$, $x \in [e^{-n}, e^a], k \in \mathbb{N}$ and $\delta = \delta(\eta) > 0$ such that for any $y \in (x-\delta, x+\delta) \cap [e^{-n}, e^a]$

$$n\, P(X_k^2 > u_n \,|\, X_0^2 = y) > C - \eta . \tag{3.34}$$

Let F_{X^2} denote the stationary distribution function of $(X_n^2)_{n\in\mathbb{N}}$. By Theorem 2.2.3 we have that

$$\lim_{n\to\infty} n\, \overline{F}_{X^2}(u_n) = 2\, c\, x^{-\kappa/2} , \tag{3.35}$$

where c is given by the formula in (2.45) and κ is the solution of (2.18). Furthermore, by (3.34) we have

$$\begin{aligned} n\, \overline{F}_{X^2}(u_n) &= \int_{(-\infty,\infty)} n\, P(X_k^2 > u_n \,|\, X_0^2 = y) dF_{X^2}(y) \\ &\ge \int_{(x-\delta,x+\delta)\cap[e^{-n},e^a]} n\, P(X_k^2 > u_n \,|\, X_0^2 = y) dF_{X^2}(y) \\ &> (C-\eta)\, P(X_0^2 \in (x-\delta, x+\delta) \cap [e^{-n}, e^a]) \\ &\ge (C-\eta)\, D , \end{aligned}$$

where $D := \inf_{z\in[0,e^a]}(F_{X^2}(z+\delta) - F_{X^2}(z)) > 0$ because F_{X^2} is continuous. Since $C > 0$ is arbitrary this is a contradiction to (3.35).

Now we estimate (3.32).

$$
\begin{aligned}
I_1 &\le \sum_{l=1}^{p-1} P\Big(\tau_1 = l,\ \max_{p\le j\le p_n} X_j^2 > u_n \,\Big|\, X_0^2 > u_n\Big) \\
&\le \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} P\Big(\tau_1 = l,\ X_j^2 > u_n \,\Big|\, X_0^2 > u_n\Big) \\
&= \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\Big(1_{\{\tau_1=l\}}\, P(X_j^2 > u_n \,|\, X_l^2) \,\Big|\, X_0^2 > u_n\Big) \qquad (3.36)\\
&= \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\Big(1_{\{\tau_1=l\}}\, 1_{\{X_l^2\ge e^{-n}\}}\, P(X_j^2 > u_n \,|\, X_l^2) \,\Big|\, X_0^2 > u_n\Big) \\
&\quad + \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\Big(1_{\{\tau_1=l\}}\, 1_{\{X_l^2< e^{-n}\}}\, P(X_j^2 > u_n \,|\, X_l^2) \,\Big|\, X_0^2 > u_n\Big) \\
&=: J_1 + J_2 .
\end{aligned}
$$

Furthermore, by (3.33),

$$
\begin{aligned}
J_1 &\le \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} \frac{1}{n} E\Big(1_{\{\tau_1=l\}}\, 1_{\{X_l^2\ge e^{-n}\}}\, n\, P(X_j^2 > u_n \,|\, X_l^2) \,\Big|\, X_0^2 > u_n\Big) \\
&\le \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} \frac{C}{n} E\Big(1_{\{\tau_1=l\}}\, 1_{\{X_l^2\ge e^{-n}\}} \,\Big|\, X_0^2 > u_n\Big) \\
&\le \sum_{j=1}^{p_n} \frac{C}{n} P(\tau_1 < p \,|\, X_0^2 > u_n) \qquad (3.37)\\
&\le C\,\frac{p_n}{n} \\
&\to 0, \quad \text{as } n\to\infty,
\end{aligned}
$$

since $p_n = o(n)$. Similarly, with $B_l := \{X_1^2 > e^a, \ldots, X_{l-1}^2 > e^a\}$ for any $l = 2, 3, 4, \ldots$ and $B_1 = \Omega$, we obtain

$$
\begin{aligned}
J_2 &\le \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\Big(1_{\{\tau_1=l\}}\, 1_{\{X_l^2< e^{-n}\}} \,\Big|\, X_0^2 > u_n\Big) \\
&= \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\Big(1_{B_l}\, P(X_l^2 < e^{-n} \,|X_{l-1}^2) \,\Big|\, X_0^2 > u_n\Big) \\
&= \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\,\Big(1_{B_l}\, P\Big((\alpha X_{l-1} + \sqrt{\beta + \lambda X_{l-1}^2}\,\varepsilon_l)^2 < e^{-n} \,\Big| X_{l-1}^2\Big) \,\Big|\, X_0^2 > u_n\Big)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\left(1_{B_l\cap\{X_{l-1}>0\}}\, P\Big(\frac{-e^{-n/2}/X_{l-1}-\alpha}{\sqrt{\beta/X_{l-1}^2+\lambda}} < \varepsilon_l < \frac{e^{-n/2}/X_{l-1}-\alpha}{\sqrt{\beta/X_{l-1}^2+\lambda}}\Big)\,\Big|\, X_0^2 > u_n\right)\\
&\quad + \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\left(1_{B_l\cap\{X_{l-1}<0\}}\, P\Big(\frac{e^{-n/2}/X_{l-1}+\alpha}{\sqrt{\beta/X_{l-1}^2+\lambda}} < \varepsilon_l < \frac{-e^{-n/2}/X_{l-1}+\alpha}{\sqrt{\beta/X_{l-1}^2+\lambda}}\Big)\,\Big|\, X_0^2 > u_n\right)\\
&= \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\left(1_{B_l\cap\{X_{l-1}>0\}}\, P\Big(\frac{-e^{-n/2-a/2}-\alpha}{\sqrt{\lambda}} < \varepsilon_l < \frac{e^{-n/2-a/2}-\alpha}{\sqrt{\lambda}}\Big)\,\Big|\, X_0^2 > u_n\right)\\
&\quad + \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\left(1_{B_l\cap\{X_{l-1}<0\}}\, P\Big(\frac{-e^{-n/2-a/2}+\alpha}{\sqrt{\lambda}} < \varepsilon_l < \frac{e^{-n/2-a/2}+\alpha}{\sqrt{\lambda}}\Big)\,\Big|\, X_0^2 > u_n\right)\\
&\le 2\,const\, p\, p_n\, e^{-n/2-a/2}\\
&\to 0\,,\quad \text{as } n\to\infty\,,
\end{aligned}
$$

and therefore with (3.37) $I_1 \to 0$ as $n\to\infty$.

Now we estimate $\limsup_{n\to\infty} I_3$. Note first that by the Markov inequality

$$
\begin{aligned}
P\Big(\max_{p\le j\le p_n} S_j^{u,a} > -z\Big) &\le \sum_{j=p}^{p_n} P\Big(e^{\frac{\kappa}{4}S_j^{u,a}} > e^{-\frac{\kappa}{4}z}\Big)\\
&= \sum_{j=p}^{p_n} P\Big(\prod_{m=1}^{j}\Big((\alpha+\sqrt{\beta e^{-a}+\lambda}\,\varepsilon_m)^2 - 2\alpha\sqrt{\beta}e^{-a/2}\varepsilon_m 1_{\{\varepsilon_m<0\}}\Big)^{\kappa/4} > e^{-\frac{\kappa}{4}z}\Big)\\
&\le e^{\frac{\kappa}{4}z}\sum_{j=p}^{p_n} E\left(\Big((\alpha+\sqrt{\beta\, e^{-a}+\lambda}\,\varepsilon_1)^2 - 2\alpha\sqrt{\beta}e^{-a/2}\varepsilon_1 1_{\{\varepsilon_1<0\}}\Big)^{\kappa/4}\right)^j\\
&\le e^{\frac{\kappa}{4}z}\sum_{j=p}^{p_n}\eta^j\,, \qquad (3.38)
\end{aligned}
$$

where $\eta<1$ such that $E\Big(\Big((\alpha+\sqrt{\beta\, e^{-a}+\lambda}\,\varepsilon_1)^2 - 2\alpha\sqrt{\beta}e^{-a/2}\varepsilon_1 1_{\{\varepsilon_1<0\}}\Big)^{\kappa/4}\Big) \le \eta$ for a large enough. This is possible because of (2.2) which implies that $E(|\alpha+\sqrt{\lambda}\,\varepsilon_1|^u) < 1$ for all $u\in(0,\kappa)$ and the fact that

$$
E\Big(\Big((\alpha+\sqrt{\beta\exp(-a)+\lambda}\,\varepsilon_1)^2 - 2\alpha\sqrt{\beta}e^{-a/2}\varepsilon_1 1_{\{\varepsilon_1<0\}}\Big)^{\kappa/4}\Big) \to E\Big(|\alpha+\sqrt{\lambda}\,\varepsilon_1|^{\kappa/2}\Big)\,,\quad a\to\infty
$$

by the dominated convergence theorem. Thus from Theorem 2.2.3, Lemma 2.3.1, (3.38) and a large enough,

$$
\begin{aligned}
\limsup_{n\to\infty} I_3 &\le \limsup_{n\to\infty} P(N_a \ge p_n,\ \max_{p\le j\le p_n} Z_0 + S_j^{u,a} > \ln u_n \mid Z_0 > \ln u_n)\\
&\le \limsup_{n\to\infty} P(\max_{p\le j\le p_n} Z_0 + S_j^{u,a} > \ln u_n \mid Z_0 > \ln u_n)\\
&= \limsup_{n\to\infty}\int_0^\infty P(\max_{p\le j\le p_n} S_j^{u,a} > -z)\frac{\kappa}{2}e^{-\frac{\kappa}{2}z}dz \qquad (3.39)
\end{aligned}
$$

$$\leq \ 2\sum_{j=p}^{\infty}\eta^j \ = \ 2\frac{\eta^{p-1}}{1-\eta}\,.$$

Finally, note that

$$\begin{aligned} I_2 \ &\leq \ P(p \leq \tau_1 < p_n,\ \max_{\tau_1 < j \leq p_n} X_j^2 > u_n \,|\, X_0^2 > u_n) + P(p \leq \tau_1 < p_n,\ \max_{p \leq j \leq \tau_1} X_j^2 > u_n \,|\, X_0^2 > u_n) \\ &=: \ K_1 \,+\, K_2\,. \end{aligned}$$

Similarly as for I_1 and I_3, respectively, we derive that

$$\limsup_{n\to\infty} K_1 = 0 \quad \text{and} \quad \limsup_{n\to\infty} K_2 \ = \ 2\frac{\eta^{p-1}}{1-\eta}\,.$$

Now plugging all together and letting $p \to \infty$ the statement follows. □

Corollary 2.3.9 *Let $(p_n)_{n\in\mathbb{N}}$ be the same sequence as in Lemma 2.3.7. Then $(p_n)_{n\in\mathbb{N}}$ is also a $\Delta(u_n)-$separating sequence for $(X_n)_{n\in\mathbb{N}}$, where $u_n = n^{1/\kappa}x$ and $x \in \mathbb{R}$ arbitrary and*

$$\lim_{p\to\infty}\limsup_{n\to\infty} P(\max_{p\leq j\leq p_n} X_j > u_n \,|\, X_0 > u_n) = 0\,. \tag{3.40}$$

Proof. Because of Remark 2.3.8(a) and (b), it is straightforward that $(p_n)_{n\in\mathbb{N}}$ is a $\Delta(u_n)-$separating sequence for $(X_n)_{n\in\mathbb{N}}$. Note furthermore that

$$\begin{aligned} P(\max_{p\leq j\leq p_n} X_j^2 > u_n^2 \,|\, X_0^2 > u_n^2) &= \frac{P(\max_{p\leq j\leq p_n} X_j^2 > u_n^2,\, X_0^2 > u_n^2)}{P(X_0^2 > u_n^2)} \\ \geq \ \frac{P(\max_{p\leq j\leq p_n} X_j > u_n,\, X_0 > u_n)}{P(X_0 > u_n) + P(X_0 < -u_n)} &= \frac{1}{2}P(\max_{p\leq j\leq p_n} X_j > u_n \,|\, X_0 > u_n) \end{aligned}$$

and hence the statement follows using Lemma 2.3.7. □

Now we are finally able to prove Theorem 2.3.5.

Proof of Theorem 2.3.5. The proof is an application of Theorem A3.2. We prove only statement (a), statement (b) follows along the same lines using Theorem A3.1. As stated already we may assume w.l.o.g. that $(X_n)_{n\in\mathbb{N}}$ is stationary. Let $x \in \mathbb{R}$ be arbitrary. Note that

$$\lim_{u\to\infty} \frac{P(X_0 > u + \frac{1}{\kappa}u\,x)}{P(X_0 > u)} = \begin{cases} \infty & , \quad 1 + \frac{1}{\kappa}x \leq 0 \\ (1+\frac{1}{\kappa}x)^{-\kappa} & , \quad 1 + \frac{1}{\kappa}x > 0 \end{cases}$$

and

$$\lim_{u\to\infty} P(\frac{X_1}{u} \leq x \,|\, X_0 = u) = P(\alpha + \sqrt{\lambda}\,\varepsilon \leq x)\,.$$

By Corollary 2.3.9 and the strong mixing property of $(X_n)_{n\in\mathbb{N}}$ all assumptions of Theorem A3.2 are fulfilled and we have that the extremal index θ is given by

$$\begin{aligned}\theta &= \int_1^\infty P(\#\{j \geq 1 \,|\, (\prod_{i=1}^{j}(\alpha+\sqrt{\lambda}\,\varepsilon_i))Y_0 > 1\} = 0 \,|\, Y_0 = y)\,\kappa\, y^{-\kappa-1}dy \\ &= \int_1^\infty P(\max_{j\geq 1}(\prod_{i=1}^{j}(\alpha+\sqrt{\lambda}\,\varepsilon_i) \leq y^{-1})\,\kappa\, y^{-\kappa-1}dy\,.\end{aligned}$$

The cluster probabilities can be determined in the same way and hence the statement follows. □

2.4 Conclusions

In this chapter we investigated the tail of the stationary distribution of the AR(1) process with ARCH(1) errors $(X_n)_{n\in\mathbb{N}}$. Our main tool was a Tauberian theorem. This approach is new as far as we know. One might expect that the method may also be applied to other models than the AR(1) model with ARCH(1) errors. Unfortunately, each model has to be studied individually in the same way as in the case of the AR(1) process with ARCH(1) errors presented in this chapter. Finally, the method does not seem to be very robust towards model changes.

After having determined the tail of the stationary distribution we studied the extremal behaviour of $(X_n)_{n\in\mathbb{N}}$. Although there exist plenty of results concerning the extremal behaviour of Markov chains, especially regenerative Markov chains, they are usually not very tractable. Checking the assumptions is a tedious and often seems even an impossible task. However, in the case of the AR(1) process with ARCH(1) errors this was possible. It appeared that the strong mixing condition and the $\Delta(u_n)$-separating sequence were crucial for the extremes of the process $(X_n)_{n\in\mathbb{N}}$.

The notion of strong mixing and $\Delta(u_n)$-separating sequence are not only known in extreme value theory but also in other areas as for instance in the theory for sample autocovariance and autocorrelation functions of heavy-tailed stationary processes (see Davis and Mikosch (1998)). Davis and Mikosch showed that if the strong mixing condition, (3.40) in a multivariate form and additionally a regular variation condition on the finite-dimensional distribution of the process hold then the weak convergence of the point processes $N_n = \sum_{j=1}^n \varepsilon_{\underline{X}_j/a_n}$ exists, where $\underline{X}_j = (X_j, ..., X_{j+m})$ for some $m \geq 0$. Finally, under some additional restrictions, even joint convergence of the sample autocovariances and autocorrelations at different lags can be established. In the case of infinite variance of $(X_n)_{n\in\mathbb{N}}$, the limits of the sample autocorrelation function are in general random. This is in contrast to infinite variance linear processes (see Davis and Resnick (1985), (1986)). Since the assumptions are fulfilled for the ARCH(1) process one might expect that they also hold for the AR(1) process with ARCH(1) errors which is simply an extension of the first. This project is part of current research and some interesting results have already been achieved. The work will be presented in a forthcoming paper.

Chapter 3

Extreme Value Theory for Diffusion Processes

Over the last decade a variety of stochastic models have been suggested as appropriate models for financial products. In a continuous time setting the dynamics of an interest rate or price process is often modelled as a diffusion process given by a stochastic differential equation (SDE)

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t\,, \quad t > 0\,, \quad X_0 = x\,, \tag{0.1}$$

where W is standard Brownian motion, μ is the *drift term* and σ is the *diffusion coefficient or volatility.* Two standard models in finance are of the above form:

(i) The Black-Scholes model: (X_t) models the price process of an asset, here $\mu(x) = \mu x$ and the volatility $\sigma(x) = \sigma x$. The resulting model for the price process is geometric Brownian motion.

(ii) The Vasicek model: the process (X_t) models an interest rate, the drift term μ is linear and the volatility $\sigma > 0$ is some constant.

Both models can be considered in the framework of Gaussian models. It has been recognised for decades that financial data like interest rates and asset prices exhibit fluctuations which cannot be modelled by Gaussian processes or simple transformations as in the two standard models above.

There are two features, heavy-tailedness and the dependence structure, that require modelling for financial data. Various models have been suggested to capture these features. For a discussion of non-linear heavy-tailed models and further references we refer to Section 7.6 of Embrechts, Klüppelberg and Mikosch (1997). There are in principle two different approaches.

A first concept replaces the Gaussian driving process in the Black-Scholes or Vasicek model (or any other traditional model) by a process with heavy-tailed marginals as for instance a stable process, a Lévy process or a discrete time counterpart as an ARMA (autoregressive-moving average) process with heavy-tailed noise (see e.g. Barndorff-Nielsen (1995), Eberlein and Keller (1995), Klüppelberg and Mikosch (1996), Mittnik and Rachev (1997)).

The second concept sticks to Brownian motion as the driving dynamic of the process, but introduces a path-dependent, time-dependent or even stochastic volatility into the model. These models are commonly referred to as volatility models, and include diffusions given by the SDE (0.1). Hence this paper is about such models. Discrete time counterparts are for instance (G)ARCH models and extensions, which have been successfully applied in econometrics. The extremal behaviour of the AR(1) process with ARCH(1) errors has been studied in Chapter 2 and is an interesting complement to the present paper.

In this chapter we study the extremal behaviour of diffusion processes defined by (0.1). The stationary distributions of the processes under investigation are well-known and one might expect that they influence the extremal behaviour of the process in some way. This is however not the case: for any pre-determined stationary distribution the process can exhibit quite different behaviour in its extremes.

Extremal behaviour of a stochastic process (X_t) is for instance manifested in the asymptotic behaviour of the maxima

$$M_t^X = \max_{0 \le s \le t} X_s \,, \quad t > 0 \,. \tag{0.2}$$

The asymptotic distribution of M_t^X for $t \to \infty$ has been studied by various authors, see Davis (1982) for detailed references. Two monographs on this and related problems are by Leadbetter, Lindgren and Rootzén (1983) and Berman (1992). It is remarkable that running maxima and minima of (X_t) are asymptotically independent and have the same behaviour as the extremes of i.i.d. random variables. In this chapter we restrict ourselves to the investigation of maxima, the mathematical treatment for minima being similar.

We furthermore investigate the point process of upcrossings (more precisely ε-upcrossings) of a high threshold u by (X_t). For fixed $\varepsilon > 0$ the process has an ε-upcrossing at t if it has remained below u on the interval $(t - \varepsilon, t)$ and is equal to u at t. Under weak conditions, the point process of ε-upcrossings, properly scaled in time and space, converges in distribution to a homogeneous Poisson process , i.e. it behaves again like i.i.d. random variables, coming however not from the stationary distribution of (X_t), but from the distribution function F which also describes the maxima M_t^X (see Theorem 3.2.4).

3.1 The usual conditions

The diffusion (X_t) given by the SDE (0.1) has state space $(l, r) \subset \mathbb{R}$, where l, r can be $-\infty$ or $+\infty$. We only consider the case when the boundaries l and r are inaccessible and (X_t) is recurrent. We require furthermore that, for all $x \in (l, r)$, $\sigma^2(x) > 0$ and there exists some $\varepsilon > 0$ such that $\int_{x-\varepsilon}^{x+\varepsilon}(1 + |\mu(t)|)/\sigma^2(t)dt < \infty$. These two conditions guarantee in particular that the SDE (0.1) has a weak solution which is unique in probability (see Karatzas and Shreve (1988), Chapter 5.5.C).

Associated with the diffusion is the scale function s and the speed measure m. The *scale function* is defined as

$$s(x) = \int_z^x \exp\left\{-2\int_z^y \frac{\mu(t)}{\sigma^2(t)}dt\right\} dy\,, \quad x \in (l, r)\,, \tag{1.1}$$

where z is any interior point of (l, r). Since the scale function is unique only up to a positive affine transformation (if $\widetilde{s}(x) = \alpha s(x) + \beta$ for some $\alpha > 0$ and $\beta \in \mathbb{R}$, then $\widetilde{s}$ is a scale function if and only if s is), in a first order approximation, the choice of z is of no importance. The scale function s defines in the usual way a measure on (l, r), the so-called *scale measure*, which is absolutely continuous with Lebesgue density

$$s'(x) = \exp\left\{-2\int_z^x \frac{\mu(t)}{\sigma^2(t)}dt\right\}\,, \quad x \in (l, r)\,. \tag{1.2}$$

For the *speed measure* m we know that $m(I) > 0$ for every non-empty open subinterval I of the interior of (l, r). We only consider diffusions with finite speed measure m and denote its total mass by $|m| = m((l, r))$. The speed measure of model (0.1) is absolutely continuous with Lebesgue density

$$m'(x) = \frac{2}{\sigma^2(x)s'(x)}\,, \quad x \in (l, r)\,. \tag{1.3}$$

In this situation (X_t) is ergodic and its stationary distribution is absolutely continuous with Lebesgue density

$$h(x) = m'(x)/|m|\,. \tag{1.4}$$

Notice that the connection between stationary distribution, speed measure, scale function, drift term and diffusion coefficient (given by (1.1)-(1.4)) allows us to construct diffusions with arbitrary stationary distribution (see Examples 3.3.4 and the generalised inverse Gaussian diffusion of Section 3.4) .

Since the process is recurrent and the boundaries l and r are inaccessible, we must have $s(u) \to \infty$ as $u \uparrow r$ and $s(u) \to -\infty$ as $u \downarrow l$. Conversely, if $s(u) \to \infty$ (resp. $-\infty$) as $u \uparrow r$ (resp. $u \downarrow l$), then l and r are inaccessible, and therefore (X_t) is recurrent.

For proofs of the above relations and further results on diffusions we refer to the monographs by Karlin and Taylor (1981), Karatzas and Shreve (1987), Rogers and Williams (1987), Revuz and Yor (1991) or any other advanced textbook on stochastic processes.

Throughout this paper, we assume that the diffusion process (X_t) defined in (0.1) satisfies *the usual conditions*:

$$\begin{aligned} s(r) = -s(l) = \infty\,, \\ |m| < \infty\,. \end{aligned} \tag{1.5}$$

3.2 Extremal behaviour of diffusions

The following formulation can be found in Davis (1982).

Proposition 3.2.1 *Let (X_t) satisfy the usual conditions (1.5). Then for any initial value $X_0 = y \in (l, r)$ and any $u_t \uparrow r$,*

$$\lim_{t\to\infty} |P^y(M_t^X \le u_t) - F^t(u_t)| = 0\,, \tag{2.1}$$

where F is a distribution function, defined by

$$F(x) = e^{-1/(|m|s(x))} 1_{(z,r)}(x)\,, \quad x \in \mathbb{R}\,, \tag{2.2}$$

for any $z \in (l, r)$. (1_A denotes the indicator function of A.) The function s and the quantity $|m|$ also depend on the choice of z. □

Various proofs of this result exist and we refer to Davis (1982) for further references. Davis' proof is based on a representation of such a diffusion as an Ornstein-Uhlenbeck process after a random time-change . Standard techniques for extremes of Gaussian processes apply leading to the above result.

It is not difficult to show that Proposition 3.2.1 is true for arbitrary initial probability measure H. For the special choice of $H = m/|m|$ the diffusion (X_t) is stationary.

As a consequence of Proposition 3.2.1, the maxima M_t^X have, properly normalised, a non-degenerate limit distribution Q if and only if F belongs to the maximum domain of attraction of Q (we write MDA(Q)) for some extreme value distribution Q . In Proposition 3.2.1, any function $u_t \uparrow r$ is possible, but as usual in classical extreme value theory we restrict ourselves to positive affine functions, i.e.

$$u_t = a_t x + b_t \,. \tag{2.3}$$

The norming constants $a_t > 0$ and $b_t \in \mathbb{R}$ have to be chosen appropriately to ensure convergence to a non-degenerate limit.

The extremal behaviour (in particular the behaviour of the maximum) of an i.i.d. sequence with common distribution function F is determined by the far end of the right tail $\overline{F} = 1 - F$. In our situation the asymptotic behaviour of the maxima M_t^X is determined by the tail of F as in (2.2): If $F \in$ MDA(Q) with norming constants $a_t > 0$ and $b_t \in \mathbb{R}$, then

$$a_t^{-1}\left(M_t^X - b_t\right) \xrightarrow{d} Q\,, \quad t \to \infty\,. \tag{2.4}$$

As already noted the scale and speed measure of a diffusion (X_t) are not unique. Different scale and speed measures (and therefore different z) lead to different distribution function's F in Proposition 3.2.1. They are however all tail-equivalent.

Corollary 3.2.2 *Under the conditions of Proposition 3.2.1 the tail of the distribution function F in (2.2) satisfies*

$$\overline{F}(x) \sim \left(|m| \int_z^x s'(y)dy\right)^{-1} \sim \left(|m| s(x)\right)^{-1}\,, \quad x \uparrow r\,, \tag{2.5}$$

where $\sim$ means that the quotient of lhs and rhs converges to 1.

Proof. The representation of (2.5) follows immediately by Taylor expansion from (2.2) and the fact that $s(x) \to \infty$ as $x \uparrow r$.

We show that the rhs is for different z asymptotically equivalent, and thus independent of $z \in (l,r)$. Let $z_1, z_2 \in (l,r)$ and $z_1 \neq z_2$. Denote s_i, m_i' and $|m_i|$ the functions and constants corresponding to z_i for $i = 1,2$. Then from (1.1) we obtain

$$s_1(x) = \alpha + \beta s_2(x)\,,$$

where α and β are constants depending on z_1 and z_2. Furthermore, from (1.3) we obtain

$$m_1'(x) = \frac{2}{\sigma^2(x)s_1'(x)} = \frac{2}{\sigma^2(x)\beta s_2'(x)} = \frac{1}{\beta}m_2'(x)\,.$$

Hence $|m_1| = |m_2|/\beta$ and

$$\overline{F}(x) \sim (|m_1|s_1(x))^{-1} = \left(|m_2|(\frac{\alpha}{\beta} + s_2(x))\right)^{-1} \sim (|m_2|s_2(x))^{-1}\,, \quad x \uparrow r\,,$$

since $\lim_{x\uparrow r} s_2(x) = \infty$. □

Proposition 3.2.1 reduces the asymptotic behaviour of the maximum of (X_t) to that of the maximum of i.i.d. random variables with distribution function F having tail (2.5). It would be interesting to know more about the extremal behaviour of the corresponding diffusion (X_t) than just the behaviour of its maxima. From classical extreme value theory it is well-known that the point process of exceedances of an i.i.d. sequence of a level u_t, plotted at points i/t, converges to a homogeneous Poisson process for $u_t \uparrow r$ as $t \uparrow \infty$ in an appropriate way. Extremes of a continuous time stochastic process over a high threshold u_t typically occur on intervals and form excursions over this level. However, an analogous discrete skeleton which describes the behaviour of the extremes of a continuous time stochastic process is provided by a point process of the upcrossings (i.e. the events where excursions above a level begin). This is quite natural and upcrossings are well–defined if the sample paths of the corresponding process are regular (i.e. differentiable in the L^2–sense). In cases with irregular sample paths there can be infinitely many upcrossings on a finite interval.

To avoid such problems special upcrossings, namely ε-uprossings, are considered. We use the definition given by Pickands (1969) for continuous processes. We also refer to Leadbetter, Lindgren and Rootzén (1983), Chapter 12, for more mathematical background.

Definition 3.2.3 *Let (X_t) be a diffusion satisfying the usual conditions (1.5). Take $\varepsilon > 0$.*
*(a) The process (X_t) is said to have an ε–*upcrossing of the level u at t_0 *if $X_t < u$ for $t \in (t_0-\epsilon, t_0)$ and $X_{t_0} = u$.*
(b) Let $N_{\varepsilon,u}(t)$ denote the number of ε-upcrossings of u by $(X_s)_{0\le s\le t}$. Then for any $t > 0$,

$$N_t^*(B) = N_{\varepsilon,u_t}(tB) = \#\{\varepsilon\text{-upcrossings of } u_t \text{ by } (X_s)_{0\le s\le t} : \frac{s}{t} \in B\}$$

is the time–normalised point process of ε-upcrossings *on the Borel sets B of $(0,1]$.* □

The point process (N_t^*) has a point at t_0 if $(X_s)_{0\le s\le t}$ has an ε–upcrossing at $t_0 t$. ε–upcrossings of a continuous time process correspond to exceedances of an i.i.d. sequence. It is well-known

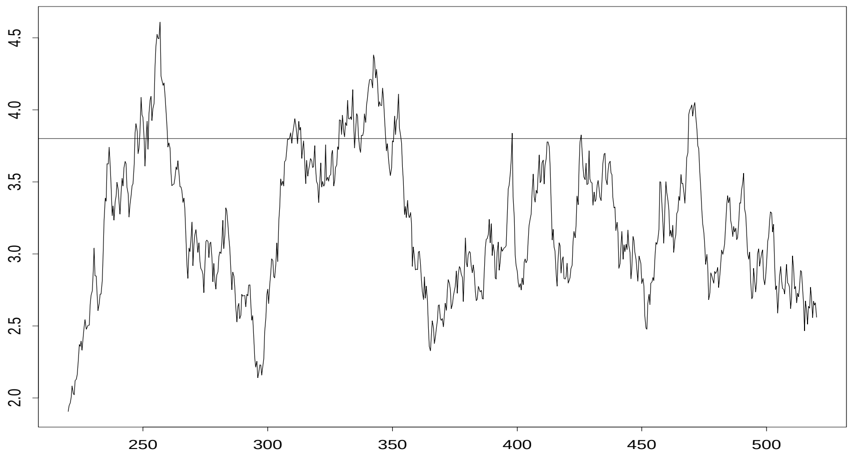

Figure 3.1: Sample path of a diffusion with threshold $u = 3.8$. For the values of $\varepsilon = 3.2, 1.2, 0.8, 0.4$ we get 6, 7, 10, 14 ε–upcrossings, respectively. The number of ε–upcrossings depends crucially on ε. The dependence only disappears in the limit.

that for a sequence (X_t) of i.i.d. random variables, all with distribution function F, the point processes (N_t^*) of exceedances converge to a homogeneous Poisson process with intensity τ, provided the u_t are appropriately chosen, namely such that

$$t\overline{F}(u_t) \to \tau \in (0,\infty)\,, \quad t \to \infty\,. \tag{2.6}$$

Recall from (2.4) that for the choice of $u_t = a_t x + b_t$:

$$P\left(M_t^X \leq a_t x + b_t\right) = F^t(a_t x + b_t) \to Q(x) = e^{-\tau}\,, \quad x \in \mathbb{R}\,. \tag{2.7}$$

Taking logarithms in (2.7) shows that (2.6) is equivalent to (2.7). Convergence of the point processes of exceedances to a Poisson process also holds for more general sequences (X_t) if the dependence structure is nice enough to prevent clustering of the extremes in the limit. For diffusions (0.1) the dependence structure of the extremes is such that the point processes of ε-upcrossings converge to a homogeneous Poisson process, however, the intensity is not determined by the stationary distribution function H , but by the distribution function F from Proposition 3.2.1. This means that the ε-upcrossings of (X_t) are likely to behave as the exceedances of i.i.d. random variables with distribution function F.

The extra condition (2.10) of the following theorem relates the scale function s and speed measure m of (X_t) to the corresponding quantities s_{ou} and m_{ou} of the standard Ornstein-

Uhlenbeck process, defined by

$$s_{ou}(x) = \sqrt{2\pi}\int_0^x e^{t^2/2}dt \quad \text{and} \quad m'_{ou}(x) = 1/s'_{ou}(x)\,, \quad x\in\mathbb{R}\,. \tag{2.8}$$

Theorem 3.2.4 *Let (X_t) satisfy the usual conditions (1.5) and $u_t \uparrow r$ such that*

$$\lim_{t\to\infty}\frac{t}{|m|s(u_t)} = \tau \in (0,\infty)\,. \tag{2.9}$$

Assume there exists some positive constant c such that

$$\frac{m'_{ou}(s_{ou}^{-1}(s(z)))}{s'_{ou}(s_{ou}^{-1}(s(z)))}\frac{s'(z)}{m'(z)} \ge c\,, \quad \forall z\in(l,r)\,. \tag{2.10}$$

Then for all starting points $y\in(l,r)$ of (X_t) and $\varepsilon>0$ the time-normalised point processes (N_t^*) of ε-upcrossings of the level u_t converge in distribution to N as $t\uparrow\infty$, where N is a homogeneous Poisson process with intensity τ on $(0,1]$.

Remark 3.2.5 (a) Notice from Corollary 3.2.2 that $t\overline{F}(u_t) \sim t/(|m|s(u_t))$. Hence, if $u_t = a_t x + b_t$ and $\tau = -\ln Q(x)$, then condition (2.9) guarantees that F belongs to some maximum domain of attraction.

(b) Pickands (1969) proved that the point processes of ε-upcrossings converge to a homogeneous Poisson process in the case when (X_t) is a Gaussian process. Notice that the assumptions of Theorem 3.2.4 are particular satisfied for the Ornstein-Uhlenbeck process with $c=1$.

(c) Examples which satisfy condition (2.10) are the Vasicek model, the Cox-Ingersoll-Ross model or the generalised Cox-Ingersoll-Ross model for $\gamma\neq 1$. All these models are presented in Section 4. Nevertheless not every diffusion satisfies the assumptions in Theorem 3.2.4. Lemma 3.2.6 indicates that for the generalised inverse Gaussian diffusion with $\chi>0, \psi>0$ and $\gamma>1.5$ or $\gamma<0.5$ the assertion of Theorem 3.2.4 may not hold.

Proof. The proof invokes a random time change argument. An application of Theorem 12.4.2 of Leadbetter et al. (1983) shows that the theorem holds for the standard Ornstein-Uhlenbeck (O_t) process. Denote by

$$Z_t = s_{ou}(O_t)\,, \quad t\ge 0\,,$$

the Ornstein-Uhlenbeck process in natural scale. Now define

$$Y_t = s(X_t)\,, \quad t\ge 0\,,$$

which is again a diffusion process in natural scale. (Y_t) can then be considered as a random time change of the process (Z_t), i.e.

$$Y_t = Z_{\tau_t} \quad a.s. \tag{2.11}$$

The random time τ_t has a representation via the local time of the process Y. This is a consequence of the Dambis-Dubins-Schwarz Theorem (Revuz and Yor (1991), Theorem 1.6, p.173), Theorem 47.1 of Rogers and Williams (1987), p.277 and Exercise 1.27 of Revuz and Yor (1991), p.226. For $z \in (l,r)$ denote $L_t(z)$ the local time of $(Y_s)_{0\le s\le t}$ in z. Then

$$\begin{aligned}
\tau_t &= \int_{-\infty}^{\infty} L_t(z)dm_{ou}(s_{ou}^{-1}(z)) \\
&= \int_{-\infty}^{\infty} L_t(z)\frac{m'_{ou}(s_{ou}^{-1}(z))}{s'_{ou}(s_{ou}^{-1}(z))}\frac{s'(s^{-1}(z))}{m'(s^{-1}(z))}dm(s^{-1}(z)) \\
&= \int_0^t \frac{m'_{ou}(s_{ou}^{-1}(Y_s))}{s'_{ou}(s_{ou}^{-1}(Y_s))}\frac{s'(s^{-1}(Y_s))}{m'(s^{-1}(Y_s))}ds \\
&= \int_0^t \frac{m'_{ou}(s_{ou}^{-1}(s(X_s)))}{s'_{ou}(s_{ou}^{-1}(s(X_s)))}\frac{s'(X_s)}{m'(X_s)}ds\,, \quad t\ge 0\,,
\end{aligned}$$

where we used the occupation time formula (cf. Revuz and Yor (1991), p.215) . Notice also that τ_t is continuous and strictly increasing. Under condition (2.10) we obtain

$$\tau_t - \tau_{t-\varepsilon} \ge c\varepsilon\,, \quad t\ge 0\,. \tag{2.12}$$

Moreover, Itô and McKean (1974), p. 228 proved the following ergodic theorem

$$\frac{\tau_t}{t} \overset{\text{a.s.}}{\to} \frac{1}{|\widetilde{m}|} = \frac{1}{|m|}\,. \tag{2.13}$$

Wlog we assume $|m| = 1$ in the following.

According to Theorem 4.7 of Kallenberg (1983) it suffices to show for any $y \in (l,r)$

$$\lim_{t\to\infty} P^y(N^X_{\varepsilon,u_t}(tU) = 0) = P(N(U) = 0)\,, \tag{2.14}$$

where U is an arbitrary union of semi-open intervals, and

$$\limsup_{t\to\infty} E^y(N^X_{\varepsilon,u_t}(t(a,b])) \le E(N((a,b])) < \infty, \quad \text{for every } (a,b] \subset (0,1]\,. \tag{2.15}$$

By definition of the processes O, Z, X and Y, setting $v_t = s(u_t)$, $z = s(y)$, $w_t = s_{ou}^{-1}(v_t)$ and $x = s_{ou}^{-1}(z)$, we have for $k \ge 1$,

$P^y(N^X_{\varepsilon,u_t}(t(a,b]) \ge k)$

$$
\begin{aligned}
&= P(\#\{\varepsilon\text{-upcrossings of } u_t \text{ by} X_\nu\,, \nu\in t(a,b]\}\geq k\,|\,X_0=y)\\
&= P(\#\{\varepsilon\text{-upcrossings of } v_t \text{ by } Y_\nu\,, \nu\in t(a,b]\}\geq k\,|\,Y_0=z)\\
&= P(\{\exists\nu_1,\dots,\nu_k\in t(a,b] : \forall i=1,\dots,k\,, Y_\nu<v_t\,\forall\nu\in(\nu_i-\varepsilon,\nu_i) \text{ and } Y_{\nu_i}=v_t\}\,|\,Y_0=z)\\
&= P(\{\exists\tau_{\nu_1},\dots,\tau_{\nu_k}\in(\tau_{ta},\tau_{tb}] : \forall i=1,\dots,k\,, Z_u<v_t\,\forall u\in(\tau_{\nu_i-\varepsilon},\tau_{\nu_i}) \text{ and } Z_{\tau_{\nu_i}}=v_t\,\}\,|\,Z_0=z)\\
&\leq P(\{\exists\tau_{\nu_1},\dots,\tau_{\nu_k}\in(\tau_{ta},\tau_{tb}] : \forall i=1,\dots,k\,, Z_u<v_t\;\;\forall u\in(\tau_{\nu_i}-c\varepsilon,\tau_{\nu_i}) \text{ and } Z_{\tau_{\nu_i}}=v_t\,\}\,|\,Z_0=z)\\
&= P(\#\{c\varepsilon\text{-upcrossings of } v_t \text{ by } Z_u\,, u\in(\tau_{ta},\tau_{tb}]\}\geq k\,|\,Z_0=z) \qquad (2.16)\\
&= P(\#\{c\varepsilon\text{-upcrossings of } s_{ou}^{-1}(v_t) \text{ by } s_{ou}^{-1}(Z_u)\,, u\in(\tau_{ta},\tau_{tb}]\}\geq k\,|\,s_{ou}^{-1}(Z_0)=s_{ou}^{-1}(z))\\
&= P^x(N^O_{c\varepsilon,w_t}((\tau_{ta},\tau_{tb}])\geq k)\,.
\end{aligned}
$$

The inequality is a consequence of (2.12). Note, since all transformations are strictly monotone and continuous, when we start with $(a,b]$, then we get again an interval $(\tau_a,\tau_b]$. Furthermore, we know already that the theorem holds for the OU-process O. We show that for all $k\geq 0$,

$$\limsup_{t\to\infty}\;\;\left|P^x(N^O_{c\varepsilon,w_t}((\tau_{ta},\tau_{tb}])\geq k)-P^x(N^O_{c\varepsilon,w_t}((ta,tb]\geq k)\right|=0\,,\quad x\in\mathbb{R}\,,$$

equivalently, for all $k\geq 0$,

$$\limsup_{t\to\infty}\;\;\left|P^x(N^O_{c\varepsilon,w_t}((\tau_{ta},\tau_{tb}])\leq k)-P^x(N^O_{c\varepsilon,w_t}((ta,tb]\leq k)\right|=0\,,\quad x\in\mathbb{R}\,, \qquad (2.17)$$

For any $0<\delta<1$, define

$$A_t=\{|\tau_{ta}-ta|\leq\delta ta\,,|\tau_{tb}-tb|\leq\delta tb\}\,,\quad t\geq 0\,.$$

By the triangular inequality , the lhs of (2.17) is bounded by

$$
\begin{aligned}
&\limsup_{t\to\infty}\left|P^x(N^O_{c\varepsilon,v_t}((\tau_{ta},\tau_{tb}])\leq k,A_t)-P^x(N^O_{c\varepsilon,w_t}((ta,tb])\leq k,A_t)\right|\\
&\quad+2\limsup_{t\to\infty}\{P^x(|\tau_{ta}-ta|\geq\delta ta)+P^x(|\tau_{tb}-tb|\geq\delta tb)\}\quad=:\quad I_1+I_2\,.
\end{aligned}
$$

Merely observed that I_2=0 by (2.13).

Again by the triangular inequality and the fact that $(\tau_{ta},\tau_{tb}]\subset((1-\delta)ta,(1+\delta)tb]$ in A_t,

$$
\begin{aligned}
I_1 &\leq \limsup_{t\to\infty}\big(P^x(N^O_{c\varepsilon,w_t}((\tau_{ta},\tau_{tb}])\leq k,A_t)-P^x(N^O_{c\varepsilon,w_t}(((1-\delta)ta,(1+\delta)tb])\leq k,A_t)\big)\\
&\quad+\limsup_{t\to\infty}\big(P^x(N^O_{c\varepsilon,w_t}((ta,tb])\leq k,A_t)-P^x(N^O_{c\varepsilon,w_t}(((1-\delta)ta,(1+\delta)tb])\leq k,A_t)\big)\\
&=: J_1+J_2\,.
\end{aligned}
$$

Furthermore,

$$\begin{aligned}
J_1 &\le \limsup_{t\to\infty}\big(P^x(N^O_{c\varepsilon,w_t}((\tau_{ta},\tau_{tb}]) \le k, A_t, N^O_{c\varepsilon,w_t}(((1-\delta)ta,(1+\delta)tb]) > k)\big)\\
&\le \limsup_{t\to\infty}\big(P^x(N^O_{c\varepsilon,w_t}(((1-\delta)ta,(1+\delta)ta]) > 0) + P^x(N^O_{c\varepsilon,w_t}(((1-\delta)tb,(1+\delta)tb]) > 0)\big)\\
&= \limsup_{t\to\infty}\big(P^H(N^*_t((0,2\delta a] > 0) + P^H(N^*_t((0,2\delta b] > 0)\big)\\
&= P(N((0,2\delta a] > 0) + P(N((0,2\delta b] > 0) \le 2(1-e^{-\tau 2\delta b}),
\end{aligned}$$

where H is the stationary distribution and N^*_t is the time-normalised point process of ε-upcrossings of the process O. We used that the Ornstein-Uhlenbeck process O has the strong Markov property and is ergodic, and that the result holds for O.

Similar considerations yield the same upper bound for J_2 and hence the lhs of (2.17) is bounded by $4(1-e^{-\tau 2\delta b})$. Letting $\delta \downarrow 0$ we have proved (2.17) for all $k \ge 0$, which yields together with (2.16) and $x = s^{-1}_{ou}(s(y))$,

$$\begin{aligned}
\limsup_{t\to\infty} E^y(N^X_{\varepsilon,u_t}(t(a,b])) &= \limsup_{t\to\infty}\sum_{k=1}^{\infty} P^y(N^X_{\varepsilon,u_t}(t(a,b]) \ge k)\\
&\le \sum_{k=1}^{\infty}\limsup_{t\to\infty} P^x(N^O_{c\varepsilon,w_t}(t(a,b]) \ge k)\\
&= \sum_{k=1}^{\infty} P(N((a,b]) \ge k) \;=\; E(N((a,b])),
\end{aligned}$$

and therefore (2.15) holds. Now we check (2.14):

W.l.o.g. choose an arbitrary U of the form $U = \bigcup_{i=1}^d (a_i,b_i]$ with disjoint intervals and $a_1 \le a_2 \le \ldots \le a_d$. Then, by definition of the ε-upcrossings,

$$\begin{aligned}
\lim_{t\to\infty} P^y(N^X_{\varepsilon,u_t}(tU) = 0) &= \lim_{t\to\infty} P^y(\{N^X_{\varepsilon,u_t}(tU) = 0\} \cap \bigcap_{i=1}^d \{M^X_{[ta_i,ta_i+\varepsilon]} < u_t\})\\
&\quad + \lim_{t\to\infty} P^y(\{N^X_{\varepsilon,u_t}(tU) = 0\} \cap \bigcup_{i=1}^d \{M^X_{[ta_i,ta_i+\varepsilon]} \ge u_t\})\\
&= \lim_{t\to\infty} P^y(\bigcap_{i=1}^d \{M^X_{t[a_i,b_i]} < u_t\})\\
&\quad + \lim_{t\to\infty} P^y(\{N^X_{\varepsilon,u_t}(tU) = 0\} \cap \bigcup_{i=1}^d \{M^X_{[ta_i,ta_i+\varepsilon]} \ge u_t\})\\
&=: K_1 + K_2 .
\end{aligned}$$

We show by induction that the rhs equals $P(\bigcap_{i=1}^d \{N((a_i,b_i]) = 0)$. Because of Proposition 3.2.1 and the fact that $K_2 = 0$ (see below) this is true for $d = 1$. Now we may assume that

$$1_{\{\bigcap_{i=1}^{d-1}\{M^X_{t[a_i,b_i]} < u_t\}\}} \quad \stackrel{d}{\to} \quad 1_{\{\bigcap_{i=1}^{d-1}\{N((a_i,b_i])=0\}\}}, \quad t\to\infty, \tag{2.18}$$

and by the Markov property,

$$P(M^X_{t[a_d,b_d]} < u_t | X_{ta_d}) \quad \stackrel{P}{\to} \quad e^{-\tau(b_d-a_d)} \quad t\to\infty\,. \tag{2.19}$$

By Slutzki's theorem, the product of the lhss of (2.18) and (2.19) converges in distribution to the product of their rhss. Applying Theorem 5.2 of Billingsley (1968) we obtain

$$\begin{aligned}
K_1 &= \lim_{t\to\infty} E^y(1_{\{\bigcap_{i=1}^{d-1}\{M^X_{t[a_i,b_i]}<u_t\}\}} P(M^X_{t[a_d,b_d]} < u_t|X_{ta_d})) \\
&= E(1_{\{\bigcap_{i=1}^{d-1}\{N((a_i,b_i])=0\}\}} e^{-\tau(b_d-a_d)}) \\
&= P(\bigcap_{i=1}^{d-1}\{N((a_i,b_i])=0\}) e^{-\tau(b_d-a_d)} \\
&= P(\bigcap_{i=1}^{d}\{N((a_i,b_i])=0\}) = P(N(U)=0)\,.
\end{aligned}$$

In the last step we used that a homogeneous Poisson process has independent increments. It remains to show $K_2 = 0$. With the same notation as before we have

$$\begin{aligned}
K_2 &\le \lim_{t\to\infty} P^y(\bigcup_{i=1}^{d}\{M^X_{[ta_i,ta_i+\varepsilon]} \ge u_t\}) \\
&\le \sum_{i=1}^{d} \lim_{t\to\infty} P^y(M^X_{[ta_i,ta_i+\varepsilon]} \ge u_t) \\
&\le \sum_{i=1}^{d} \lim_{t\to\infty} P^z(M^Z_{[\tau_{ta_i},\tau_{ta_i+\varepsilon}]} \ge v_t) \\
&\le \sum_{i=1}^{d} \lim_{t\to\infty} P^z(M^Z_{[\tau_{ta_i},\tau_{ta_i+\varepsilon}]} \ge v_t, |\tau_{ta_i} - ta_i| < \delta ta_i, |\tau_{ta_i+\varepsilon} - ta_i| < \delta ta_i) \\
&\quad + \sum_{i=1}^{d} \lim_{t\to\infty} (P^z(|\tau_{ta_i} - ta_i| \ge \delta ta_i) + P^z(|\tau_{ta_i+\varepsilon} - ta_i| \ge \delta ta_i))\ .
\end{aligned}$$

Because of (2.13), the second and third term vanish. Again by ergodicy and Proposition 3.2.1, $K_2 \le \sum_{i=1}^d (1 - e^{-2\tau\delta a_i})$. Letting $\delta \downarrow 0$, $K_2 = 0$ and we proved

$$\lim_{t\to\infty} P^y(N^X_{\varepsilon,u_t}(tU) = 0) = P(N(U)=0)\,. \tag{2.20}$$

and hence (2.14). □

Theorem 3.2.4 describes the asymptotic behaviour of the number of ε-upcrossings of a suitably increasing level. In particular, on average there are τ ε-upcrossings of u_t by $(X_s)_{0\le s\le t}$ for large t. Notice furthermore, that we get a Poisson process in the limit which is independent of the choice of $\varepsilon > 0$. A visualisation of the Poisson approximation of Theorem 3.2.4 is shown in Figure 3.13 for the generalised inverse Gaussian diffusion.

The next lemma provides a simple sufficient condition, only on scale function and speed measure of (X_t), for (2.10). By positivity and continuity, (2.10) holds automatically on compact intervals. It remains to check this condition for z in a neighbourhood of r and l.

Lemma 3.2.6 *Let (X_t) satisfy the usual conditions (1.5). Assume furthermore that (2.9) holds and that there exist $c_1, c_2 \in (0, \infty]$ such that*

$$\frac{1}{4\ln(|s(z)|)s(z)}\left(\frac{s''(z)}{s'(z)m'(z)} - \frac{m''(z)}{(m'(z))^2}\right) \quad \to \quad c_1 \text{ or } c_2 \tag{2.21}$$

according as $z \uparrow r$ or $z \downarrow l$, then the assertion of Theorem 3.2.5 holds.

Proof. By l'Hospital ,

$$s_{ou}(x) \sim g(x) = \sqrt{2\pi}e^{x^2/2}/x\,, \quad x \to \infty\,, \tag{2.22}$$

and s_{ou} and g are unbounded and non-decreasing for all x large enough. Moreover, s_{ou} and g are inversely asymptotic , i.e. for all $\lambda > 1$, there exists some $x_0(\lambda)$ such that

$$s_{ou}(x/\lambda) \le g(x) \le s_{ou}(\lambda x), \quad \forall x \ge x_0(\lambda)\,.$$

This implies by Exercise 14 of Bingham, Goldie and Teugels (1987), Section 3.13, that $s_{ou}^{-1}(x) \sim g^{-1}(x) \sim \sqrt{2\ln x}$ as $x \to \infty$. Thus, by l'Hospital,

$$\begin{aligned}
\frac{m'_{ou}(s_{ou}^{-1}(s(z)))}{s'_{ou}(s_{ou}^{-1}(s(z)))}\frac{s'(z)}{m'(z)} \quad &\sim \quad \frac{s''(z)/m'(z) - s'(z)m''(z)/(m'(z))^2}{2s''_{ou}(s_{ou}^{-1}(s(z)))s'(z)} \\
&\sim \quad \frac{1}{2(s_{ou}^{-1}(s(z)))^2 s(z)}\left(\frac{s''(z)}{s'(z)m'(z)} - \frac{m''(z)}{(m'(z))^2}\right) \\
&\sim \quad \frac{1}{4\ln(|s(z)|)s(z)}\left(\frac{s''(z)}{s'(z)m'(z)} - \frac{m''(z)}{(m'(z))^2}\right), \quad z \uparrow r \text{ or } z \downarrow l\,.
\end{aligned}$$

The second line is a consequence of (2.22) for $x = s_{ou}^{-1}(s(z))$ which tends to $\pm\infty$ as $z \downarrow l$ or $z \uparrow r$. In the last line we have used that $s_{ou}^{-1}(x) \sim \pm\sqrt{2\ln|x|}$ as $x \to \pm\infty$. $\square$

In the following situations we work out conditions on μ and σ such that the tail behaviour of F can easily be described. We apply these results to the examples in Sections 3.3 and 3.4.

Theorem 3.2.7 *Assume that the usual conditions hold.*

(a) Assume that $\mu \equiv 0$. Then $(l, r) = (-\infty, \infty)$ and

$$\overline{F}(x) \sim \left(\int_{-\infty}^{\infty}(2/\sigma^2(t))dt\right)^{-1} x^{-1}, \quad x \to \infty\,.$$

(b) Assume that $r = \infty$ and $-\infty < \rho = \int_z^r \mu(t)/\sigma^2(t)dt < \infty$ for some $z \in (l, \infty)$. Then

$$\overline{F}(x) \sim e^{2\rho}|m|^{-1}x^{-1}\,, \quad x \to \infty\,. \tag{2.23}$$

(c) Let μ and σ^2 be differentiable functions on (x_0, r) for some $x_0 < r$ such that

$$\lim_{x\uparrow r} \frac{d}{dx}\left\{\frac{\sigma^2(x)}{\mu(x)}\right\} = 0 \quad \text{and} \quad \lim_{x\uparrow r} \frac{\sigma^2(x)}{\mu(x)} \exp\left\{-2\int_z^x \frac{\mu(t)}{\sigma^2(t)}dt\right\} = -\infty\,. \tag{2.24}$$

Then

$$\overline{F}(x) \sim |\mu(x)|h(x)\,, \quad x \uparrow r\,, \tag{2.25}$$

where h is the stationary density of (X_t).

Proof. We first prove (b). By l'Hospital and (1.2),

$$\lim_{x\to\infty} \frac{s(x)}{x} = \lim_{x\to\infty} s'(x) \to e^{-2\rho}\,, \quad x \to \infty\,.$$

This implies that $s(x) \sim e^{-2\rho}x$ as $x \to \infty$. Now Corollary 3.2.2 applies.

(a) Immediately from (1.2) we have $s'(x) = 1$ for all $x \in (l, r)$. Hence by (1.1) $s(x) = x - z$ for $z \in (l, r)$. Since $\lim_{x\uparrow r} s(x) = \infty$ and $\lim_{x\downarrow l} s(x) = -\infty$, we must have $l = -\infty$ and $r = \infty$. Then part (b) applies with $\rho = 0$ and $|m| = \int_{-\infty}^{\infty} (2/\sigma^2(t))dt$.

(c) s' is an exponential function , hence

$$s''(x) = -2s'(x)\frac{\mu(x)}{\sigma^2(x)}\,, \quad x \in (l, r)\,.$$

Then by l'Hospital (which can be applied because of (2.24)),

$$\lim_{x\uparrow r} \frac{2\int_z^x s'(y)dy}{-s'(x)\sigma^2(x)/\mu(x)} = \lim_{x\uparrow r} \frac{2s'(x)}{-s'(x)o(1) - s''(x)\sigma^2(x)/\mu(x)} = 1\,, \quad x \uparrow r\,.$$

Inserting this in (2.5) yields $\overline{F}(x) \sim -2\mu(x)/(|m|s'(x)\sigma^2(x))$ as $x \uparrow r$, and the result follows from (1.3) and (1.4). □

From equations (1.1)-(1.4) it is clear that (X_t) is also uniquely determined by its stationary density $h(x)$ and the diffusion coefficient $\sigma(x)$. They determine the drift term which is for differentiable volatility σ

$$\mu(x) = \frac{\sigma^2(x)}{2}\frac{d}{dx}\ln(\sigma^2(x)h(x))\,, \quad x \in (l, r)\,. \tag{2.26}$$

Theorem 3.2.8 *Assume that the usual conditions hold with $r = \infty$. Let h be the stationary density, h positive on (x_0, ∞) for some $x_0 > 0$.*
(a) If $\sigma^2(x) \sim x^{1-\delta}\ell(x)/h(x)$ as $x \to \infty$ for some $\delta > 0$, where ℓ is a slowly varying function such that $1/\ell$ is locally bounded. Then

$$\overline{F}(x) \sim \frac{\delta}{2} x^{-\delta} \ell(x) \,, \quad x \to \infty \,.$$

(b) If $\sigma^2(x) \sim c x^{\delta-1} e^{-\alpha x^\beta}/h(x)$ as $x \to \infty$ for $\alpha, \beta, c > 0, \delta \in \mathbb{R}$, then

$$\overline{F}(x) \sim \frac{c}{2} x^{\delta-2} e^{-\alpha x^\beta} \,, \quad x \to \infty \,. \tag{2.27}$$

Proof. (a) By (1.3) and (1.4) $s'(x) \sim 2x^{-(1-\delta)}/(|m|\ell(x))$ as $x \to \infty$. Hence s' is regularly varying with index $\delta - 1$ and is locally bounded. From Corollary 3.2.2 it follows with Karamata's theorem (Theorem 1.5.11 of Bingham, Goldie and Teugels (1987)) that

$$\overline{F}(x) \sim \frac{\delta}{2} x^{-\delta} \ell(x) \,, \quad x \to \infty \,.$$

(b) By (1.3) and (1.4) we obtain $s'(x) \sim 2x^{-(\delta-1)} e^{\alpha x^\beta}/(c|m|)$ as $x \to \infty$. Then by l'Hospital

$$s(x) \sim \frac{2}{|m|c} x^{-\delta+2} e^{\alpha x^\beta} \,, \quad x \to \infty \,.$$

giving (2.27) by Corollary 3.2.2. □

This result provides a method to construct diffusions with any arbitrary stationary density (with right endpoint $r = \infty$) and any extremal behaviour.

3.3 Extremes of stochastic models in finance

Diffusion processes given by the SDE (0.1)

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \,, \quad t > 0 \,,$$

with properties as described in Section 3.1 are common models in finance; see e.g. Lamberton and Lapeyre (1991), Duffie (1992), Merton (1994) or Baxter and Rennie (1996). Examples 3.3.1, 3.3.2, and 3.3.3 are standard models for the term structure of interest rates; diffusions as Example 3.3.4 have been successfully fitted to share prices (Küchler et al. (1994), Eberlein and Keller (1995)).

The state space (l, r) and the range of parameters of all models below is such that $\lim_{x\uparrow r} s(x) = \infty$ and $\lim_{x\downarrow l} s(x) = -\infty$, hence the boundaries are inaccessible. This can easily be checked by standard calculations and (1.1). Furthermore, the speed measure m is finite for all models, the processes are ergodic with stationary distribution which is absolutely continuous with density h given by (1.4). Hence all these models satisfy the usual conditions (1.5).

Once F is determined for any of these models, classical extreme value theory takes over. Recall that there are three extreme value distribution functions (up to affine transformations). Since all the examples we treat in Section 3.3 are diffusions with state space unbounded above, we only consider the Fréchet distribution function and the Gumbel distribution function given by

$$\begin{aligned}\Phi_\alpha(x) &= \exp\{-x^{-\alpha}\}1_{(0,\infty)}(x)\,, \quad \alpha > 0\,,\\ \Lambda(x) &= \exp\{-e^{-x}\}\,, \quad x \in \mathbb{R}\,.\end{aligned}$$

If $F \in \mathrm{MDA}(\Phi_\alpha)$, then the norming constants a_t and b_t can be chosen such that

$$\overline{F}(a_t) \sim t^{-1} \quad \text{and} \quad b_t = 0\,. \tag{3.28}$$

If $F \in \mathrm{MDA}(\Lambda)$, then the norming constants a_t and b_t can be chosen such that

$$\overline{F}(b_t) = t^{-1} \quad \text{and} \quad a_t \sim a(b_t)\,, \tag{3.29}$$

where a is the so-called auxiliary function; see e.g. Theorem 3.3.26 of Embrechts et al. (1997). Calculating the norming constants explicitly is then a standard, though often tedious task. For b_t a Taylor expansion leads to the necessary accurracy required by the convergence to types theorem. We refer to the monographs by Leadbetter, Lindgren and Rootzén (1983), Resnick (1987) or Embrechts et al. (1997) for some tutorial examples (see also Appendix A2).

Then (2.4) implies that

$$\frac{M_t^X}{a_t} \xrightarrow{d} \Phi_\alpha \quad \text{if} \quad F \in \mathrm{MDA}(\Phi_\alpha) \tag{3.30}$$

and

$$\frac{M_t^X - b_t}{a_t} \xrightarrow{d} \Lambda \quad \text{and} \quad \frac{M_t^X}{b_t} \xrightarrow{P} 1 \quad \text{if} \quad F \in \mathrm{MDA}(\Lambda)\,. \tag{3.31}$$

Furthermore, all the models in this section except the generalised Cox-Ingersoll-Ross model $\gamma = 1$ satisfy condition (2.21) of Lemma 3.2.6, hence the Poisson approximation of the ε–upcrossings is also explicitly given for $u_t = a_t x + b_t$ and $\tau = -\ln Q(x)$, where Q is either Φ_α or Λ.

Figures 3.2, 3.4, 3.5, 3.7, 3.9, 3.11 and 3.12 show simulated sample paths (of length $t =$ 1000, 10 000 and 25 000, respectively) of the different models. The solid line indicates those norming constants which describe the increase of M_t^X for large t, i.e. in MDA(Φ_α) we plot a_t (see (3.30)) and in MDA(Λ) we plot b_t (see (3.31)). Figures 3.3, 3.6, 3.8, 3.10 display the empirical distribution function of the normalised maximum of the different models and the corresponding limit distributions (based on 50 simulated sample paths of length $n = 20\,000$). Figure 3.15 and 3.16 show the empirical distribution function, the empirical density and the QQ-plot (based on 350 simulated maxima, each taken from a sample path with $t = 25\,000$) of the normalised maxima of the generalised inverse Gaussian model for certain parameter values together with the corresponding limit distribution function and density. The models were simulated by means of the Milstein Scheme (strong Taylor approximation of convergence order 1) and we refer to Kloeden and Platen (1992) for details. The quality of the simulation depends on the stepsize Δ of the discretisation: a too big stepsize Δ can actually have disastrous effects on the precision of the simulation in the extremes. For our simulations we obtained acceptable results for $\Delta = 10^{-4}$, but this may also depend on the parameters chosen.

Example 3.3.1 (The Vasicek model)

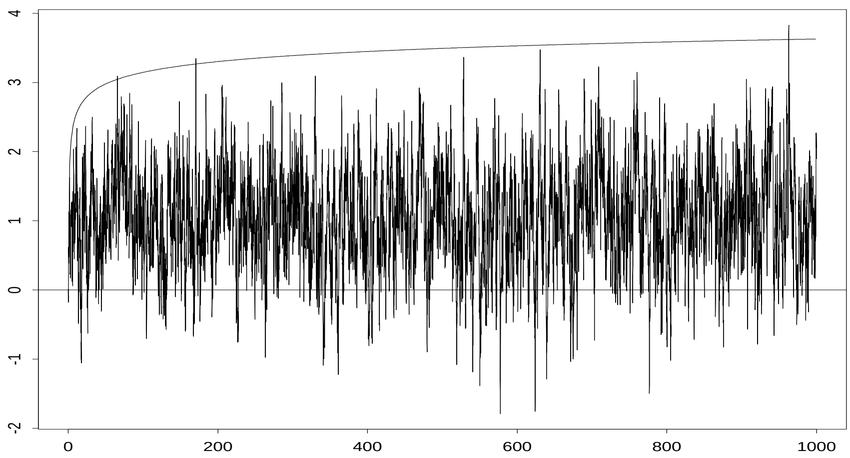

Figure 3.2: Simulated sample path of the Vasicek model (with parameters $c = d = \sigma = 1$) and corresponding normalising constants b_t.

In this model the drift term is $\mu(x) = c - dx$ for $c \in \mathbb{R}$, $d > 0$ and the diffusion coefficient $\sigma(x) \equiv \sigma > 0$. The solution of the SDE (0.1) with $X_0 = x$ is given by

$$X_t = \frac{c}{d} + (x - \frac{c}{d})e^{-dt} + \sigma \int_0^t e^{-d(t-s)} dW_s\,, \quad t \geq 0\,.$$

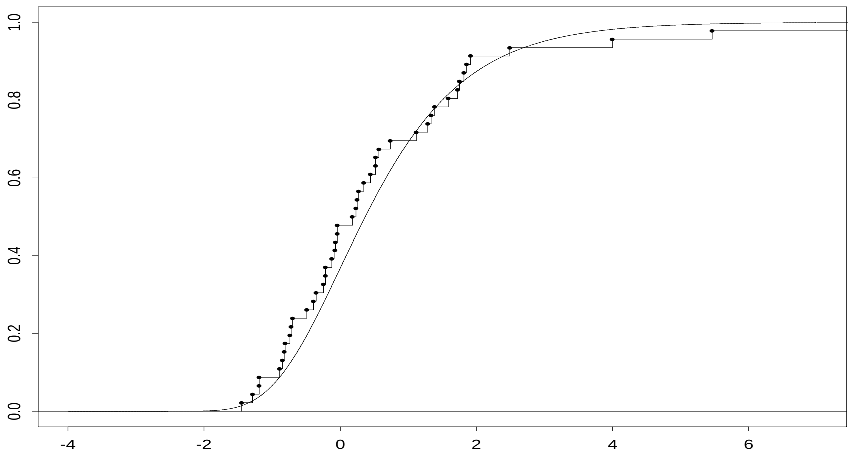

Figure 3.3: Empirical distribution function of the normalised maxima of the Vasicek model and the Gumbel distribution function with the same parameters as in Figure 3.2. We used 50 simulated sample paths of length $n = 20000$.

(X_t) is a Gaussian process and has state space $\mathbb{R}$ (which is not very satisfactory from a practical point of view), mean value function

$$EX_t = \frac{c}{d} + (x - \frac{c}{d})e^{-dt} \to \frac{c}{d} \quad \text{and} \quad \text{var}X_t = \frac{\sigma^2}{2d}\left(1 - e^{-2dt}\right) \to \frac{\sigma^2}{2d}, \quad t \to \infty\,.$$

The scale function s has density according to (1.2) (with $z = 0$)

$$s'(x) = \exp\left\{-\frac{c^2}{\sigma^2 d}\right\} \exp\left\{\frac{d}{\sigma^2}\left(x - \frac{c}{d}\right)^2\right\}, \quad x \in \mathbb{R}, \tag{3.32}$$

and speed measure m with density according to (1.3)

$$m'(x) = \frac{2}{\sigma^2} \exp\left\{\frac{c^2}{\sigma^2 d}\right\} \exp\left\{-\frac{d}{\sigma^2}\left(x - \frac{c}{d}\right)^2\right\}, \quad x \in \mathbb{R}.$$

By standard calculation,

$$|m| = \int_{-\infty}^{\infty} m'(x)dx = \frac{2}{\sigma}\sqrt{\frac{\pi}{d}}\, e^{c^2/(d\sigma^2)} < \infty\,.$$

The assumptions of Theorem 3.2.7(c) are satisfied and hence

$$\begin{aligned} \overline{F}(x) \quad &\sim \quad \frac{d}{\sqrt{2\pi}} \frac{x - c/d}{\sqrt{\sigma^2/(2d)}} \exp\left\{-\frac{1}{2}\frac{(x - c/d)^2}{\sigma^2/(2d)}\right\} \\ &= \quad \frac{d}{\sqrt{2\pi}} \overline{G}\left(\frac{x - c/d}{\sqrt{\sigma^2/(2d)}}\right), \quad x \to \infty\,, \end{aligned}$$

where $\overline{G}(x) = x\exp\{-x^2/2\}$. Note that $F(x) \sim A\,(x - c/d)^2\,\overline{H}(x)$, where $\overline{H}(x)$ is the tail of the stationary normal distribution function and A is some constant; hence F has heavier tail than H. In order to investigate the extremal behaviour of F we first consider the one of G. One can readilly see that G is a von Mises function (see Proposition A1.2) with auxiliary function $a(x) \sim 1/x$. Hence $G \in \mathrm{MDA}(\Lambda)$ and the norming constants can be calculated as in Proposition A1.3(b): the centering constants $\widetilde{b}_t$ as asymptotic solution to $G(\widetilde{b}_t) = 1 - t^{-1}$ and the normalising constants $\widetilde{a}_t \sim a(\widetilde{b}_t)$. Standard calculation (see Appendix A2) yields $G^{\leftarrow}(1-t^{-1}) = \sqrt{2\ln t} + \frac{\ln(2\ln t)}{2\sqrt{2\ln t}} + o(\frac{1}{\sqrt{2\ln t}})$. Hence by Proposition A1.3(c), we choose

$$\widetilde{a}_t = \frac{1}{\sqrt{2\ln t}} \quad \text{and} \quad \widetilde{b}_t = \sqrt{2\ln t} + \frac{\ln(2\ln t)}{2\sqrt{2\ln t}} .$$

From Proposition A1.7 we obtain thus the norming constants to the tail $\overline{G}\left((x - c/d)/\sqrt{\sigma^2/(2d)}\right)$ as

$$\widehat{a}_t = \frac{\sigma}{\sqrt{2d}}\widetilde{a}_t \quad \text{and} \quad \widehat{b}_t = \frac{c}{d} + \frac{\sigma}{\sqrt{2d}}\widetilde{b}_t ,$$

Finally, by tail equivalence and Proposition A1.6 we obtain the norming constants of F as

$$a_t = \frac{\sigma}{2\sqrt{d\ln t}} ,$$

$$b_t = \frac{\sigma}{\sqrt{d}}\sqrt{\ln t} + \frac{c}{d} + \frac{\sigma}{4\sqrt{d}} \frac{\ln\ln t + \ln(d^2/\pi)}{\sqrt{\ln t}} .$$

By Proposition 3.2.1 this implies that

$$\frac{2}{\sigma}\sqrt{d\ln t}\left(M_t^X - \frac{\sigma}{\sqrt{d}}\sqrt{\ln t} - \frac{c}{d} - \frac{\sigma}{4\sqrt{d}} \frac{\ln\ln t + \ln(d^2/\pi)}{\sqrt{\ln t}}\right) \xrightarrow{d} \Lambda .$$

Example 3.3.2 (The Cox-Ingersoll-Ross model)
In this model (X_t) satisfies the SDE (0.1) with $\mu(x) = c - dx$ for $d > 0$, $\sigma(x) = \sigma\sqrt{x}$ for $\sigma > 0$ and $2c \geq \sigma^2$. It has state space $(0, \infty)$, mean value function

$$EX_t = \frac{c}{d} + \left(x - \frac{c}{d}\right) e^{-dt} \to \frac{c}{d}, \quad t \to \infty$$

and

$$\mathrm{var}X_t = \frac{c\sigma^2}{2d^2}\left(1 - \left(1 + \left(x - \frac{c}{d}\right)\frac{2d}{c}\right)e^{-2dt} + \left(x - \frac{c}{d}\right)\frac{2d}{c}e^{-3dt}\right) \to \frac{c\sigma^2}{2d^2}, \quad t \to \infty ,$$

where $X_0 = x$. For the exact marginal distribution of X_t we refer to Lamberton, Lapeyre (1991), Chapter 6. The scale function s has density according to (3.32) (with $z = 1$),

$$s'(x) = x^{-2c/\sigma^2} \exp\left\{\frac{2d}{\sigma^2}(x - 1)\right\} , \quad x \in (0, \infty) .$$

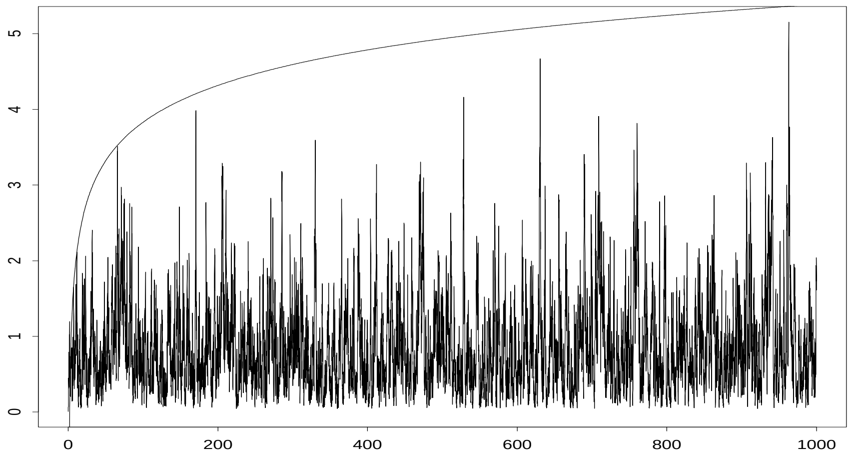

Figure 3.4: Simulated sample path of the Cox-Ingersoll-Ross model (with parameters $c = d = \sigma = 1$) and the corresponding norming constants b_t.

The speed measure m has density

$$m'(x) = \frac{2}{\sigma^2} x^{2c/\sigma^2 - 1} \exp\left\{ -\frac{2}{\sigma^2} d(x-1) \right\}, \quad x \in (0, \infty).$$

and hence

$$\begin{aligned} |m| &= \frac{2}{\sigma^2} e^{2d/\sigma^2} \int_0^\infty x^{2c/\sigma^2 - 1} e^{-(2d/\sigma^2)x} dx \\ &= \frac{2}{\sigma^2} e^{2d/\sigma^2} \left(\frac{\sigma^2}{2d} \right)^{2c/\sigma^2} \Gamma\left(\frac{2c}{\sigma^2} \right) < \infty . \end{aligned}$$

We conclude that the stationary distribution is $\Gamma(\frac{2c}{\sigma^2}, \frac{2d}{\sigma^2})$.

For the asymptotic distribution of the maximum M_t^X of (X_t) we calculate the tail of F. Theorem 3.2.7(c) applies and we obtain

$$\begin{aligned} \overline{F}(x) &\sim (dx - c)h(x) \\ &\sim dxh(x) \\ &\sim d \left(\frac{2d}{\sigma^2} \right)^{2c/\sigma^2} \left(\Gamma\left(\frac{2c}{\sigma^2} \right) \right)^{-1} x^{(2c/\sigma^2 + 1) - 1} e^{-(2d/\sigma^2)x} \\ &\sim cg(x) \\ &\sim \frac{2cd}{\sigma^2} \overline{G}(x), \quad x \to \infty, \end{aligned}$$

where g is the density and $\overline{G}(x)$ is the tail of the $\Gamma(\frac{2c}{\sigma^2} + 1, \frac{2d}{\sigma^2})$ distribution. It can be seen that $\overline{F}(x) \sim B\, x\, \overline{H}(x)$ for some $B > 0$. It is well-known (see e.g. Embrechts at al. (1997), Section 3.3, p.156) that the gamma distributions are in MDA(Λ) and the norming constants for G are

$$\widetilde{a}_t = \sigma^2/(2d) \quad \text{and} \quad \widetilde{b}_t = \frac{\sigma^2}{2d}\left(\ln t + \frac{2c}{\sigma^2}\ln\ln t + \ln\left(\frac{1}{\Gamma(2c/\sigma^2+1)}\right)\right).$$

By Proposition **??** we obtain as norming constants for F

$$a_t = \sigma^2/(2d) \tag{3.33}$$

$$b_t = \frac{\sigma^2}{2d}\left(\ln t + \frac{2c}{\sigma^2}\ln\ln t + \ln\left(\frac{d}{\Gamma(2c/\sigma^2)}\right)\right). \tag{3.34}$$

This implies that

$$\frac{2d}{\sigma^2}\left(M_t^X - \frac{\sigma^2}{2d}\left(\ln t + \frac{2c}{\sigma^2}\ln\ln t + \ln\left(\frac{d}{\Gamma(2c/\sigma^2)}\right)\right)\right) \xrightarrow{d} \Lambda\,.$$

Notice that for $\sigma^2 << c$ the constant $\Gamma(2c/\sigma^2)$ is very large and consequently b_t may become negative for small t. In extreme cases b_t becomes positive only for very large t.

Example 3.3.3 (Generalised Cox-Ingersoll-Ross model)
In this model the drift term is given by $\mu(x) = c - dx$ and the diffusion coefficient has the form $\sigma(x) = \sigma x^\gamma$ for $\gamma \in [\frac{1}{2}, \infty)$. For $\gamma < \frac{1}{2}$ we have $|m| = \infty$ and hence by Theorem 7 of Mandl (1968) , p.90, the process is not ergodic. For $\gamma \geq \frac{1}{2}$ the process is ergodic with state space $(0,\infty)$.

We distinguish the following four cases:

$$\begin{array}{llll} \gamma = 1/2 & : & 2c \geq \sigma^2\,, & d > 0 \quad \text{(see Example 3.3.2)} \\ 1/2 < \gamma < 1 & : & c > 0\,, & d \geq 0 \\ \gamma = 1 & : & c > 0\,, & d > -\sigma^2/2 \\ \gamma > 1 & : & c > 0\,, d \in \mathbb{R} & \text{or} \quad c = 0\,, d < 0\,. \end{array} \tag{3.35}$$

For $\frac{1}{2} \leq \gamma \leq 1$ the mean value function of (X_t) is given by

$$EX_t = \begin{cases} \dfrac{c}{d} + \left(x - \dfrac{c}{d}\right)e^{-dt} & \to \dfrac{c}{d} \quad \text{if} \quad d > 0 \\ \dfrac{c}{d} + \left(x - \dfrac{c}{d}\right)e^{-dt} & \to \infty \quad \text{if} \quad d < 0 \\ x + ct & \to \infty \quad \text{if} \quad d = 0 \end{cases} \tag{3.36}$$

as $t \to \infty$ where $X_0 = x$. This indicates already that for certain parameter values the model can capture large fluctuations in data, which will reflect also in the behaviour of the maxima.

In all three cases we calculate the respective quantities with $z = 1$.

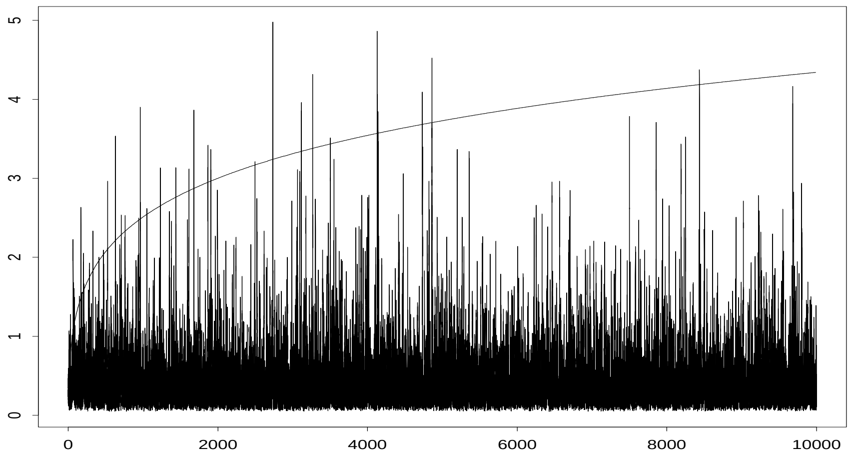

Figure 3.5: Simulated sample path of the generalised Cox-Ingersoll-Ross model for $\gamma = 0.75$ (with parameters $c = d = \sigma = 1$) and the corresponding norming constants b_t. A sample sample path of length $t = 10\,000$ has been simulated in order to show that at least for large t the approximation by b_t is reasonable.

- $\frac{1}{2} < \gamma < 1$

We calculate (1.2):

$$s'(x) = \exp\left\{-\frac{2}{\sigma^2}\left(\frac{c}{2\gamma-1} + \frac{d}{2-2\gamma}\right)\right\} \exp\left\{\frac{2}{\sigma^2}\left(\frac{c}{2\gamma-1}x^{-2\gamma+1} + \frac{d}{2-2\gamma}x^{2-2\gamma}\right)\right\}.$$

With $m'(x)$ as in (1.3) we obtain

$$\begin{aligned} |m| &= \frac{2}{\sigma^2}\exp\left\{\frac{2}{\sigma^2}\left(\frac{c}{2\gamma-1} + \frac{d}{2-2\gamma}\right)\right\} \\ &\qquad \int_0^\infty t^{-2\gamma}\exp\left\{-\frac{2}{\sigma^2}\left(\frac{c}{2\gamma-1}t^{-(2\gamma-1)} + \frac{d}{2-2\gamma}t^{2-2\gamma}\right)\right\} dt \end{aligned} \tag{3.37}$$

The stationary density as in (1.4) is

$$h(x) = A\frac{\sigma^2}{2}x^{-2\gamma}\exp\left\{-\frac{2}{\sigma^2}\left(\frac{c}{2\gamma-1}x^{-(2\gamma-1)} + \frac{d}{2-2\gamma}x^{2-2\gamma}\right)\right\}, \quad x > 0,$$

where

$$A = |m|\exp\left\{-\frac{2}{\sigma^2}\left(\frac{c}{2\gamma-1} + \frac{d}{2-2\gamma}\right)\right\}.$$

The assumptions of Theorem 3.2.7(c) are satisfied and hence

$$\begin{aligned} \overline{F}(x) &\sim dxh(x) \\ &\sim \frac{d}{A}\frac{2}{\sigma^2}x^{-2\gamma+1}\exp\left\{-\frac{d}{\sigma^2(1-\gamma)}x^{2(1-\gamma)}\right\} \\ &\sim \frac{d}{A}\frac{2}{\sigma^2}\overline{G}(x), \quad x \to \infty. \end{aligned}$$

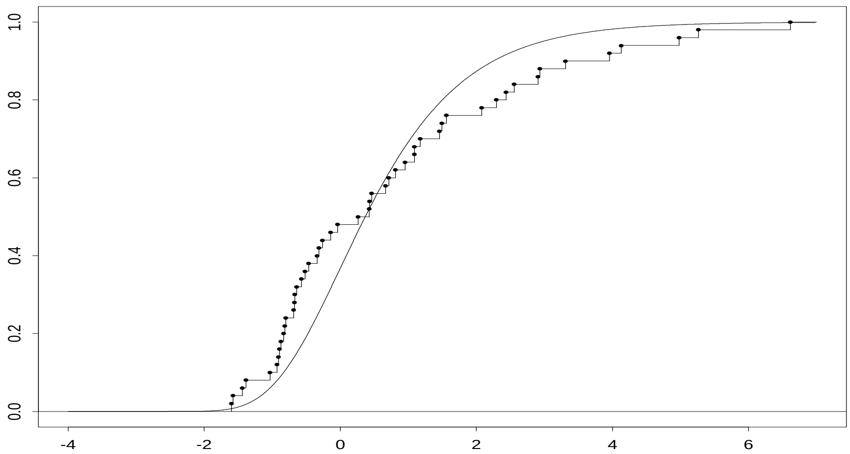

Figure 3.6: Empirical distribution function of the normalised maxima of the generalised Cox-Ingersoll-Ross model and the Gumbel distribution function for the same parameters as in Figure 3.5.

Notice that $\overline{F}(x) \sim C\, x^{2(1-\gamma)}\, \overline{H}(x)$ for some $C > 0$.

The distribution function G is a von Mises function with representation as in Proposition A1.2, hence $G \in \mathrm{MDA}(\Lambda)$. Rather lengthy, but standard calculations (see Appendix A2) yield the norming constants for G and then for F by Proposition A1.6, giving

$$a_t = \frac{\sigma^2}{2d}\left(\frac{\sigma^2(1-\gamma)}{d}\ln t\right)^{\frac{2\gamma-1}{2-2\gamma}} \tag{3.38}$$

$$b_t = \left(\frac{\sigma^2(1-\gamma)}{d}\ln t\right)^{\frac{1}{2-2\gamma}}\left(1-\frac{2\gamma-1}{2(1-\gamma)}\frac{\ln\left(\frac{\sigma^2(1-\gamma)}{d}\ln t\right)}{\ln t}\right)^{\frac{1}{2-2\gamma}} + a_t \ln\left(\frac{2d}{A\sigma^2}\right), \tag{3.39}$$

and hence

$$a_t^{-1}(M_t^X - b_t) \xrightarrow{d} \Lambda\,.$$

Note that a_t is continuous in the point $\gamma = 1/2$, i.e. a_t as above converges to $\sigma^2/(2d)$ as $\gamma \downarrow 1/2$, which is the same as (3.33). For the norming constants b_t we obtain

$$b_t \sim \frac{\sigma^2}{2d}\ln t + \frac{\sigma^2}{2d}\ln\left(\frac{2d}{A\sigma^2}\right), \quad \gamma \downarrow 1/2\,,$$

hence its first term coincides with the first term of (3.34).

The behaviour of a_t and b_t as $\gamma \uparrow 1$ is much more dramatic. It indicates already that at $\gamma = 1$ there must be a qualitative change in the extremal behaviour. This is confirmed in the following.

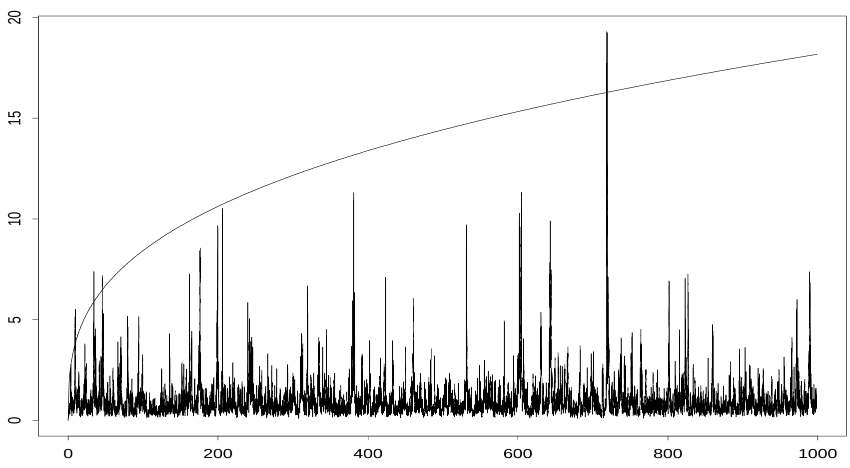

Figure 3.7: Simulated sample path of the generalised Cox-Ingersoll-Ross model for $\gamma = 1$ (with parameters $c = d = \sigma = 1$) and the corresponding norming constants b_t.

- $\gamma = 1$

In this case the solution of the SDE (0.1) with $X_0 = x$ is explicitly given by

$$X_t = e^{-(d+\frac{\sigma^2}{2})t+\sigma W_t}\left(x + c\int_0^t e^{(d+\frac{\sigma^2}{2})s-\sigma W_s}ds\right), \quad t \geq 0\,.$$

We obtain from (1.2)

$$s'(x) = \exp\left\{-\frac{2c}{\sigma^2}\right\} x^{2d/\sigma^2} \exp\left\{\frac{2c}{\sigma^2}x^{-1}\right\}.$$

With $m'(x)$ as in (1.3) we obtain

$$|m| = \frac{2}{\sigma^2}\exp\left\{\frac{2c}{\sigma^2}\right\}\Gamma\left(\frac{2d}{\sigma^2}+1\right)\left(\frac{\sigma^2}{2c}\right)^{\frac{2d}{\sigma^2}+1}.$$

The density of the stationary distribution is

$$h(x) = \left(\frac{\sigma^2}{2c}\right)^{-\frac{2d}{\sigma^2}-1}\left(\Gamma\left(\frac{2d}{\sigma^2}+1\right)\right)^{-1} x^{-2d/\sigma^2-2}\exp\left\{-\frac{2c}{\sigma^2}x^{-1}\right\}, \quad x > 0\,.$$

h is regularly varying with index $-2d/\sigma^2 - 2$ and hence by Karamata's theorem (Theorem 1.5.11 of Bingham, Goldie and Teugels (1987)) the tail $\overline{H}$ of the stationary distribution is also regularly varying. This implies that certain moments are infinite:

$$\lim_{t\to\infty} EX_t^\delta = \begin{cases} \left(\frac{2c}{\sigma^2}\right)^\delta \dfrac{\Gamma\left(\frac{2d}{\sigma^2}+1-\delta\right)}{\Gamma\left(\frac{2d}{\sigma^2}+1\right)} & \text{if} \quad \delta < \dfrac{2d}{\sigma^2}+1\,, \\ \infty & \text{if} \quad \delta \geq \dfrac{2d}{\sigma^2}+1\,. \end{cases}$$

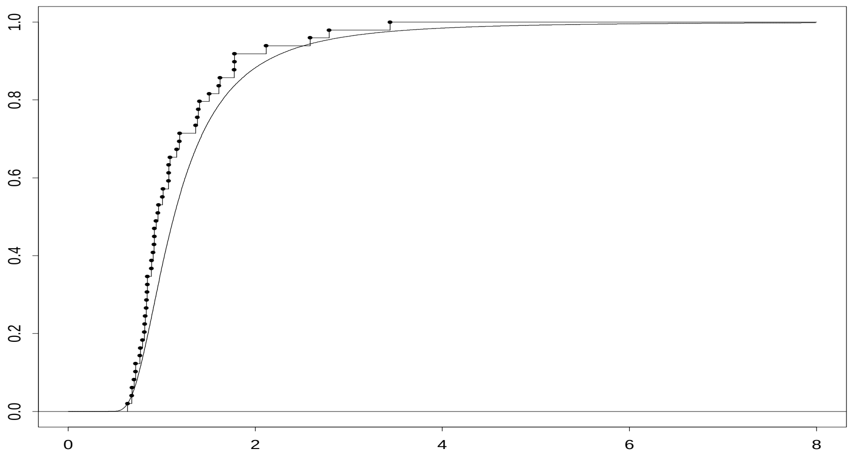

Figure 3.8: Empirical distribution function of the normalised maxima of the generalised Cox-Ingersoll-Ross model for $\gamma = 1$ and the Fréchet distribution function for the same paramters as in Figure 3.4.

In particular,

$$\lim_{t\to\infty} \mathrm{var} X_t = \begin{cases} \dfrac{2c^2}{d(2d-\sigma^2)} < \infty & \text{if } \dfrac{2d}{\sigma^2} > 1\,, \\ \infty & \text{if } -1 < \dfrac{2d}{\sigma^2} \le 1\,. \end{cases}$$

For the tail of F we obtain by Theorem 3.2.7(c)

$$\begin{aligned} \overline{F}(x) &\sim \frac{\sigma^2}{2}\left(\frac{\sigma^2}{2c}\right)^{-\frac{2d}{\sigma^2}-1}\left(\Gamma\left(\frac{2d}{\sigma^2}+1\right)\right)^{-1}\left(\frac{2d}{\sigma^2}+1\right) x^{-2d/\sigma^2-1}\exp\left\{-\frac{2c}{\sigma^2}x^{-1}\right\} \\ &\sim \frac{\sigma^2}{2}\left(\frac{\sigma^2}{2c}\right)^{-\frac{2d}{\sigma^2}-1}\left(\Gamma\left(\frac{2d}{\sigma^2}+1\right)\right)^{-1}\left(\frac{2d}{\sigma^2}+1\right) x^{-2d/\sigma^2-1}\,, \quad x\to\infty\,. \end{aligned}$$

Hence $\overline{F}$ is regularly varying and by Proposition A1.1 $F \in \mathrm{MDA}(\Phi_{1+2d/\sigma^2})$, with norming constants and the normalising constants a_t chosen according to Proposition A1.3 as

$$a_t \sim \left(\frac{\sigma^2}{2}\left(\frac{\sigma^2}{2c}\right)^{-\frac{2d}{\sigma^2}-1}\left(\Gamma\left(\frac{2d}{\sigma^2}+1\right)\right)^{-1}\left(\frac{2d}{\sigma^2}+1\right) t\right)^{1/(1+2d/\sigma^2)}\,, \quad b_t = 0\,.$$

Notice that $a_t \sim D\, t^{1/(1+2d/\sigma^2)}$ for a constant D, i.e. a_t is a decreasing function of d/σ^2. Hence the maxima M_t^X are likely do increase slower, when d/σ^2 gets larger. In particular,

$$M_t^X / \left(Dt^{1/(1+2d/\sigma^2)}\right) \xrightarrow{d} \Phi_{1+2d/\sigma^2}\,, \quad t\to\infty\,.$$

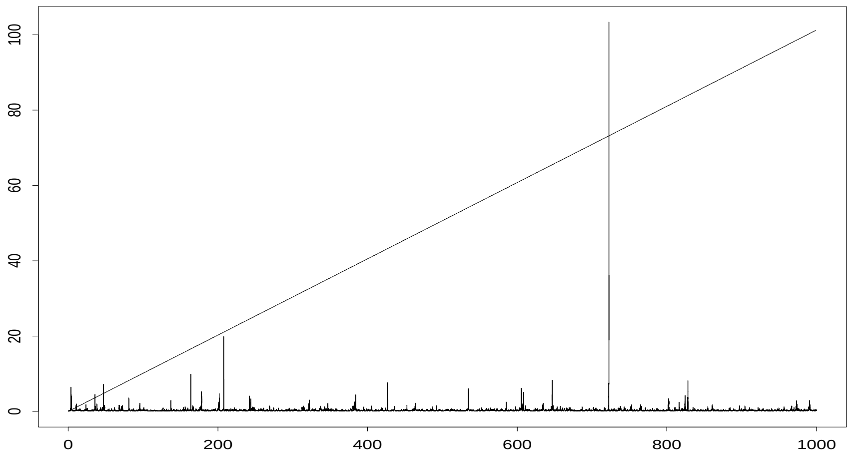

Figure 3.9: Simulated sample path of the generalised Cox-Ingersoll-Ross model for $\gamma = 1.5$ (with parameters $c = d = \sigma = 1$) and the corresponding norming constants a_t.

- $\gamma > 1$

Notice first that the functions s', m', h and the constant A are of the same form as in the case $\frac{1}{2} < \gamma < 1$. We apply Theorem 3.2.7(b) and obtain

$$\overline{F}(x) \sim \frac{e^{2\rho}}{|m|} x^{-1} = (Ax)^{-1}, \quad x \to \infty,$$

where

$$\rho = \frac{1}{\sigma^2}\left(\frac{c}{2\gamma - 1} + \frac{d}{2 - 2\gamma}\right).$$

Hence $F \in \text{MDA}(\Phi_1)$ with norming constants $a_t \sim t/A$. One can observe that the order of increase of a_t is always linear. The constant A decides about the slope. We obtain in particular

$$AM_t^X/t \xrightarrow{d} \Phi_1 .$$

For $\gamma = 3/2$ it is possible to calculate A explicitly. We obtain

$$\begin{aligned} A &= \frac{2}{\sigma^2}\int_0^\infty x^{-3}\exp(-2/\sigma^2(x^{-2}c/2 - x^{-1}d))dx \\ &= \frac{2}{\sigma^2}\int_0^\infty y\exp(-\frac{1}{2}\frac{(y^2 - 2yd/c)}{\sigma^2/2c})dy \\ &= \frac{2}{\sigma^2}e^{d^2/\sigma^2 c}\int_0^\infty y\exp(-\frac{1}{2}\frac{(y - d/c)^2}{\sigma^2/2c})dy \\ &= \frac{1}{c}\left(1 + 2\sqrt{\frac{d^2\pi}{\sigma^2 c}}e^{d^2/(c\sigma^2)}\Phi\left(\sqrt{\frac{2d^2}{c\sigma^2}}\right)\right), \end{aligned}$$

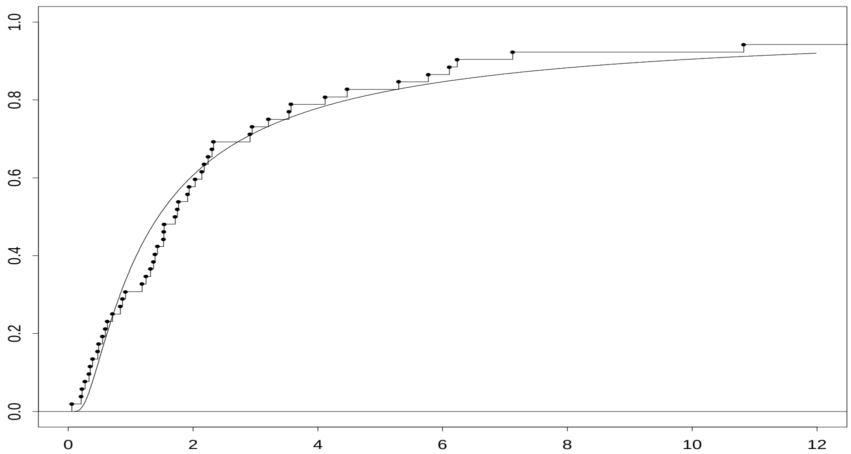

Figure 3.10: Empirical distribution function of the normalised maxima of the generalised Cox-Ingersoll-Ross model for $\gamma = 1.5$ and the Fréchet distribution function Φ_1 with the same parameters as in Figure 3.9.

where Φ denotes the standard normal distribution function . Note that (if we ignore c for the moment) A is increasing in the quotient $d^2/(c\sigma^2)$.

Example 3.3.4 (Generalised hyperbolic diffusion)
Diffusions with given stationary distribution have been considered as appropriate models for asset prices. Models considered assume that the price process follows the SDE (0.1) with drift term zero; i.e.

$$dX_t = \sigma(X_t)dW_t\,, \quad t > 0\,,$$

with diffusion coefficient $\sigma(x)$ and state space $\mathbb{R}$. Choose

$$\sigma^2(x) = \sigma^2/h(x)\,,$$

where h is an arbitrary density, then (X_t) has exactly this stationary density h. These diffusion models have been considered as alternatives to Gaussian processes for asset prices. In their most general form, as introduced by Rydberg (1996) the stationary distribution is a generalised hyperbolic distribution. A generalised hyperbolic random variable is $N(a + bZ, Z)$, where Z is a generalised inverse Gaussian random variable; hence it is a normal variance-mean mixture. The model includes the hyperbolic diffusion and the normal inverse Gaussian diffusion.

The hyperbolic diffusion has been considered as a model for asset prices by Bibby and Sørensen (1995). Eberlein and Keller (1995) and Küchler et al. (1994) fitted the hyperbolic distribution to the marginals of the price process of certain assets. Statistical modelling by means of the hyperbolic distribution has been effective in a number of contexts (see Barndorff-Nielsen (1995) for further references).

The normal inverse Gaussian diffusion has been considered by Rydberg (1996). The stationary distribution is the normal inverse Gaussian distribution as defined in Rydberg (1996) or Barndorff-Nielsen (1995). A state space/stochastic volatility model based on the normal inverse Gaussian distribution has been introduced in the latter paper.

These models have in common that the tails of their stationary distribution are log-linear, hence the stationary distribution belongs to MDA(Λ).

Since all these diffusions have been constructed with drift term $\mu = 0$ and $\sigma^2(x) = \sigma^2/h(x)$ for $\sigma > 0$ and pre-determined stationary density h, Theorem 3.2.7(a) applies, yielding

$$2\,M_t^X/(\sigma^2 t) \xrightarrow{d} \Phi_1\,, \quad t \to \infty\,,$$

regardless of their stationary distribution. $F \in \mathrm{MDA}(\Phi_1)$ means that the maximum of the process is likely to behave as the maximum of i.i.d. random variables with distribution tail $\overline{F}(x) \sim \frac{\sigma^2}{2} x^{-1}$, so the process is likely to show more extreme fluctuation than one expects from its stationary distribution.

3.4 Generalised inverse Gaussian diffusion

In Example 3.3.4 we have seen how an ergodic diffusion with drift term $\mu \equiv 0$ and arbitrary stationary distribution can be constructed. This construction has the drawback that all these diffusions show the same behaviour in their maxima M_t^X represented by $\overline{F}(x) \sim Ax^{-1}$ for some $A > 0$. Guided by Theorem 3.2.8 we choose another method of construction. We choose a density $h(x)$ and a diffusion coefficient $\sigma(x)$. By equation (2.26) this defines a drift term μ, giving an SDE (0.1).

We shall present this method by introducing a new class of diffusions with generalised inverse Gaussian stationary distribution and state space $(0, \infty)$. Its stationary distribution has (like the generalised hyperbolic distribution) tails with asymptotic behaviour reaching from exponential to regularly varying. Moreover, this model can be viewed as a further generalisation of the Cox-Ross-Ingersoll model (Example 3.3.2). It also includes Example 3.3.3 for $\gamma = 1$.

The density of the generalised inverse Gaussian distribution is given by

$$h(x) = \frac{(\chi/\psi)^{\lambda/2}}{2K_\lambda\left(\sqrt{\chi\psi}\right)} x^{\lambda-1} \exp\left\{-\frac{1}{2}(\chi x + \psi x^{-1})\right\}\,, \quad x > 0\,,$$

where K_λ is the modified Bessel function of the third kind and index λ. The following parameter sets are possible

$$\begin{aligned}
&\{\chi > 0\,, \quad \psi > 0\,, \quad \lambda \in \mathbb{R}\}\\
&\{\chi = 0\,, \quad \psi > 0\,, \quad \lambda < 0\}\\
&\{\chi > 0\,, \quad \psi = 0\,, \quad \lambda > 0\}\,.
\end{aligned}$$

The norming constant simplifies for $\chi = 0$ and $\psi = 0$. For further details concerning the generalised inverse Gaussian distribution and its properties we refer to Joergensen (1982).

Now we consider the special case of $\sigma(x) = \sigma x^\gamma$ for $\sigma > 0$ and $\gamma \geq 0$. For the sake of comparison we choose the diffusion coefficient to be the same as in the Cox-Ingersoll-Ross model (Example 3.3.2) and its generalised version (Example 3.3.3). Of course, any other diffusion coefficient is possible, leading to different classes of models (with the appropriate restriction of the parameter space).

By equation (2.26), (1.2) and (1.3),

$$\begin{aligned}
\mu(x) &= \frac{1}{4}\sigma^2 x^{2\gamma-2}\left(\psi + 2(2\gamma+\lambda-1)x - \chi x^2\right)\,,\\
s'(x) &= \exp\left\{-\frac{1}{2}(\chi+\psi)\right\} x^{-(2\gamma+\lambda-1)} \exp\left\{\frac{1}{2}(\chi x + \psi x^{-1})\right\} \text{ and}\\
m'(x) &= \frac{2}{\sigma^2}\exp\left\{\frac{1}{2}(\chi+\psi)\right\} x^{\lambda-1}\exp\left\{-\frac{1}{2}(\chi x + \psi x^{-1})\right\}\,.
\end{aligned}$$

As formulated in Section 3.1 we require the process to be recurrent and to have inaccessible boundaries 0 and ∞, i.e.

$$\int_1^\infty x^{1-\lambda-2\gamma}\exp\left\{\frac{\chi}{2}x\right\}dx = \infty \quad \text{and} \quad \int_0^1 x^{1-\lambda-2\gamma}\exp\left\{\frac{\psi}{2}x^{-1}\right\}dx = \infty\,.$$

This puts further restrictions on the parameter space and we consider

$$\begin{aligned}
&\{\chi > 0\,, \quad \psi > 0\,, \quad \lambda \in \mathbb{R}\}\\
&\{\chi = 0\,, \quad \psi > 0\,, \quad \lambda < 0 \quad \text{and} \quad \lambda \leq 2(1-\gamma)\}\\
&\{\chi > 0\,, \quad \psi = 0\,, \quad \lambda > 0 \quad \text{and} \quad \lambda \geq 2(1-\gamma)\}\,.
\end{aligned} \tag{4.1}$$

The SDE (0.1) with μ and σ as above with this restricted parameter space has a unique solution (X_t). This can be shown for instance by an application of Theorem 5.13 of Karatzas and Shreve (1987) to $\ln(X_t)$. We call this solution (X_t) *generalised inverse Gaussian diffusion* (GIG diffusion). For all parameters in (4.1), $|m| < \infty$ and hence by Theorem 7 of Mandl (1968), p. 90, (X_t) is ergodic.

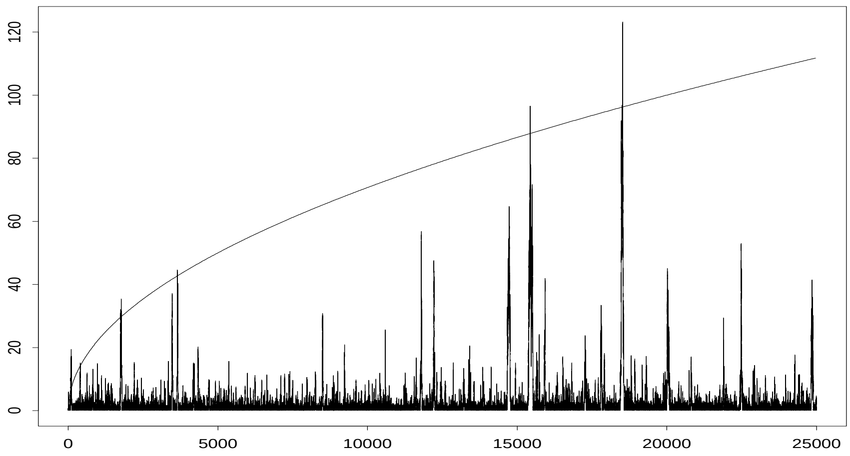

Figure 3.11: Simulated sample path of the GIG model (with parameters $\gamma = 0.5$, $\chi = 0$, $\psi = \sigma = 1$, $\lambda = -1$) and the corresponding norming constants a_t.

Comparison of the drift terms show that the GIG diffusion for $\gamma = 1/2$ and $\psi = 0$ (which implies $\chi > 0$ and $\lambda \geq 1$) is just the CIR model with parameters $c = \sigma^2\lambda/2$ and $d = \sigma^2\chi/4$. If we choose $\gamma = 1$ and $\chi = 0$ (which implies $\psi > 0$ and $\lambda < 0$), then we obtain the generalised CIR model with parameters $c = \sigma^2\psi/4$ and $d = -\sigma^2(1+\lambda)/2$.

For the study of the extremal behaviour of (X_t) we distinguish three different cases:

- $\chi = 0$

Then $\psi > 0$ and $\lambda < 0$ and $\lambda \leq 2(1-\gamma)$. The stationary density is inverse gamma of the form

$$\begin{aligned} h(x) &= \frac{(\psi/2)^{-\lambda}}{\Gamma(-\lambda)} x^{\lambda-1} \exp\left\{-\frac{1}{2}\psi x^{-1}\right\} \\ &\sim \frac{(\psi/2)^{-\lambda}}{\Gamma(-\lambda)} x^{\lambda-1}\,, \quad x \to \infty\,, \end{aligned}$$

i.e. it is regularly varying. Now Theorem 3.2.8(b) applies giving

$$\overline{F}(x) \sim \frac{\sigma^2(\psi/2)^{-\lambda}(2-2\gamma-\lambda)}{2\Gamma(-\lambda)} x^{-(2-2\gamma-\lambda)}\,, \quad x \to \infty\,.$$

By Proposition A1.1 $F \in \mathrm{MDA}(\Phi_{2-2\gamma-\lambda})$ with norming constants chosen according to Proposition A1.3(a) yielding

$$a_t \sim \left(\frac{\sigma^2(\psi/2)^{-\lambda}(2-2\gamma-\lambda)}{\Gamma(-\lambda)}\, t\right)^{1/(2-2\gamma-\lambda)} \quad \text{and} \quad b_t = 0\,.$$

By Karamata's theorem (Theorem 1.5.8 of Bingham, Goldie and Teugels (1987))

$$\overline{H}(x) \sim \frac{(\psi/2)^{-\gamma}}{\Gamma(-\lambda)(-\lambda)} x^{\lambda}, \quad x \to \infty,$$

and hence

$$\overline{H}(x) \sim \frac{2}{\sigma^2(-\lambda)(2-2\gamma-\lambda)} x^{2-2\gamma}\overline{F}(x), \quad x \to \infty.$$

Hence, depending on the choice of γ, the tail $\overline{H}$ of the stationary distribution can be heavier or lighter than $\overline{F}$, the tail which describes the asymptotic behaviour of M_t^X.

- $\psi = 0$

Then $\chi > 0$ and $\lambda > 0$ and $\lambda \geq 2(1-\gamma)$. The stationary density simplifies

$$h(x) = \frac{(\chi/2)^{-\lambda}}{\Gamma(\lambda)} x^{\lambda-1} \exp\left\{-\frac{1}{2}\chi x\right\}, \quad x > 0,$$

which is a $\Gamma(\lambda, \chi/2)$ density. Now Theorem 3.2.8(b) applies giving

$$\begin{aligned}
\overline{F}(x) &\sim \frac{\sigma^2(\chi/2)^{\lambda+1}}{2\Gamma(\lambda)} x^{\lambda+2\gamma-1} \exp\{-\frac{\chi}{2}x\} \\
&\sim \frac{\sigma^2(\chi/2)^{2-2\gamma}\Gamma(2\gamma+\lambda)}{2\Gamma(\lambda)} \overline{G}(x), \quad x \to \infty,
\end{aligned}$$

where G is the $\Gamma(2\gamma+\lambda, \chi/2)$ distribution function. Then as in the Cox-Ingersoll-Ross model, F is of gamma-type. Hence $F \in \mathrm{MDA}(\Lambda)$ with norming constants chosen according to Proposition A1.3(b). By tail-equivalence and Proposition A1.6 we obtain

$$a_t = 2/\chi \quad \text{and} \quad b_t = \frac{2}{\chi}\left(\ln t + (2\gamma+\lambda-1)\ln\ln t + \ln\left(\frac{\sigma^2(\chi/2)^{2-2\gamma}}{2\Gamma(\lambda)}\right)\right).$$

Similar calculations as above yield

$$\overline{H}(x) \sim \frac{2}{\sigma^2}\left(\frac{\chi}{2}\right)^{-2\lambda-3} x^{-2\gamma}\overline{F}(x), \quad x \to \infty.$$

Since $\gamma > 0$ this implies that F has heavier-tail than the stationary distribution function H. Hence the extremal behaviour of (X_t) shows larger fluctuations than an i.i.d. family of random variables with distribution function H was likely to show.

- $\psi > 0, \chi > 0$

Then λ is arbitrary in $\mathbb{R}$. Theorem 3.2.8(b) applies giving

$$\begin{aligned}
\overline{F}(x) &\sim \frac{\sigma^2(\chi/2)}{4K_\lambda(\sqrt{\chi\psi})} x^{2\gamma+\lambda-1} \exp\{-\frac{\chi}{2}x\} \\
&\sim \frac{\sigma^2(\chi/2)^{2-2\gamma-\lambda}}{4}\left(\frac{\chi}{\psi}\right)^{\lambda/2} \frac{\Gamma(2\gamma+\lambda)}{K_\lambda(\sqrt{\chi\psi})} \overline{G}(x), \quad x \to \infty,
\end{aligned}$$

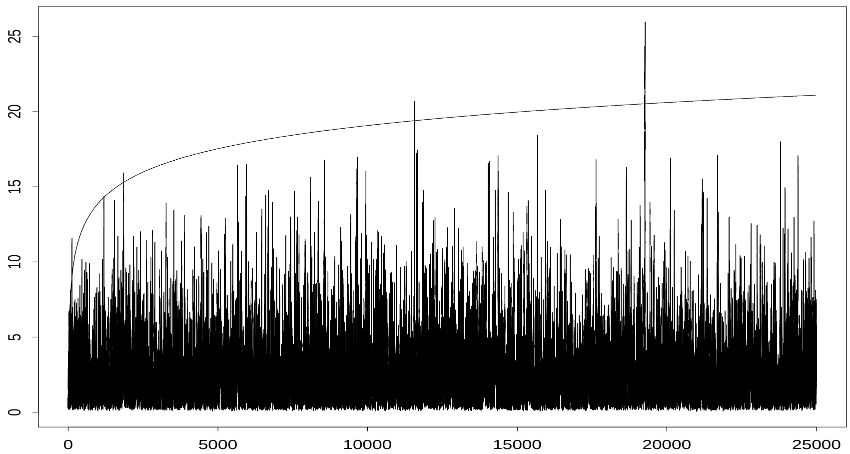

Figure 3.12: Simulated sample path of the GIG model (with parameters $\gamma = 0.5$, $\chi = \psi = \sigma = \lambda = 1$) and the corresponding norming constants b_t.

where G is the $\Gamma(2\gamma+\lambda, \chi/2)$ distribution function. Notice that G is exactly the same distribution function as in the previous case. By tail-equivalence and Proposition A1.6 we obtain the norming constants

$$a_t = 2/\chi \quad \textit{and} \quad b_t = \frac{2}{\chi}\left(\ln t + (2\gamma+\lambda-1)\ln\ln t + \ln\left(\frac{\sigma^2(\chi/2)^{2-2\gamma-\lambda}}{4K_\lambda(\sqrt{\chi\psi})}\left(\frac{\chi}{\psi}\right)^{\lambda/2}\right)\right).$$

As above

$$\overline{H}(x) \sim \frac{2}{\sigma^2}\left(\frac{2}{\chi}\right)^2 x^{-2\gamma}\overline{F}(x)\,, \quad x \to \infty\,.$$

The remark at the end of the case $\psi = 0$ applies.

Finally we investigate the assumptions in Theorem 3.2.4 for this case in detail. First notice that $s'(x) \to \infty$ for $x \downarrow 0$ or $x \uparrow \infty$. Thus by l'Hospital

$$\frac{s'(x)}{s(x)} \sim \frac{\chi}{2}, \quad x \uparrow \infty \quad \text{and} \quad \frac{s'(x)}{s(x)} \sim -\frac{\psi}{2}x^{-2}, \quad x \downarrow 0.$$

By Lemma 3.2.6 and the fact that $m'(x)s'(x) = 2x^{-2\gamma}/\sigma^2$,

$$\frac{m'_{ou}(s_{ou}^{-1}(s(x)))}{s'_{ou}(s_{ou}^{-1}(s(x)))}\frac{s'(x)}{m'(x)} \sim \frac{\sigma^2}{4}\frac{s'(x)}{s(x)}\frac{x^{2\gamma-1}}{\ln(|s(x)|)}\left(\frac{s''(x)}{s'(x)}x+\gamma\right), \quad x \uparrow \infty \text{ or } x \downarrow 0.$$

If we further distinguish between left and right endpoint we derive

$$\frac{m'_{ou}(s_{ou}^{-1}(s(x)))}{s'_{ou}(s_{ou}^{-1}(s(x)))}\frac{s'(x)}{m'(x)} \to \begin{cases} 0 & \gamma < 0.5 \\ \frac{\sigma^2}{4}\frac{\chi}{2} & \gamma = 0.5 \\ \infty & \gamma > 0.5 \end{cases}, \quad x \uparrow \infty,$$

and

$$\frac{m'_{ou}(s_{ou}^{-1}(s(x)))}{s'_{ou}(s_{ou}^{-1}(s(x)))}\frac{s'(x)}{m'(x)} \to \begin{cases} \infty & \gamma < 1.5 \\ \dfrac{\sigma^2}{4}\dfrac{\psi}{2} & \gamma = 1.5 \\ 0 & \gamma > 1.5 \end{cases}, \quad x \downarrow 0.$$

Hence by Remark 3.2.5(c), we may conclude that in the case $0.5 \leq \gamma \leq 1.5$ the assumptions of the Theorem 3.2.4 are fulfilled while in the other cases condition (2.21) of Lemma 3.2.6 does not hold.

3.5 Concluding remarks

As was demonstrated in this chapter extreme value theory for a large class of diffusion processes is strongly connected with classical extreme value theory for i.i.d. random variables. This connection is not only valid for the behaviour of the maximum of a diffusion process but also, under some additional restrictions, for the point processes of ε-upcrossings. It appeared that ε-upcrossings of a diffusion process are likely to behave as the exceedances of i.i.d. random variables with a well-specified distribution function, i.e. the associated point processes converge to the same homogeneous Poisson process. It would be interesting to see whether this result even holds if the technical assumption (2.10) is not satisfied (e.g. in the case of the generalised CIR model with $\gamma = 1$).

The results of this chapter may be applied to study risk measures of financial products as for instance the value at risk or related quantile risk measures; see Embrechts, Klüppelberg and Mikosch (1997), Example 6.1.6. Nevertheless, one has to consider the quality of the approximation in the extremes of the diffusion process and the associated sequence of i.i.d. random variables. The speed of convergence might be quite slow. Indeed, not much is known about the rate of convergence for extremes of continuous parameter processes. Konakov and Piterbarg (1982) give bounds for the maximum of general stationary normal processes. However, the bounds do not seem to be very sharp (logarithmical decrease). In Kratz and Rootzén (1998) they were improved considerably in the case of mean square differentiable stationary normal processes. There is no related result for diffusion processes. An obvious idea is to first investigate again the order of convergence for the Ornstein-Uhlenbeck process and then transform the result to arbitrary ergodic diffusion processes with inaccessible boundaries by means of our random time change argument as in section 3.2. This work is currently under way and will be presented in a forthcoming paper.

Another problem concerning application of the results in this chapter to financial problems is that the assumption of ergodicity is crucial. Unfortunately, diffusion models for stocks and other securities clearly do not satisfy this assumption. In these cases, the main problem is the random time change transformation which makes no sense for t tending to infinity, since the total speed measure is not finite. Investigating how to overcome these difficulties would be a suitable topic for future research.

Finally, another practical question is the behaviour of the maximum of a bond or swap price for a very large maturity T. Bond and swap prices are functions of interest rates which are usually modelled by diffusion processes that fit in our framework. We are optimistic that our results might help in solving this problem. However, a detailed analysis of this question is again left for future research.

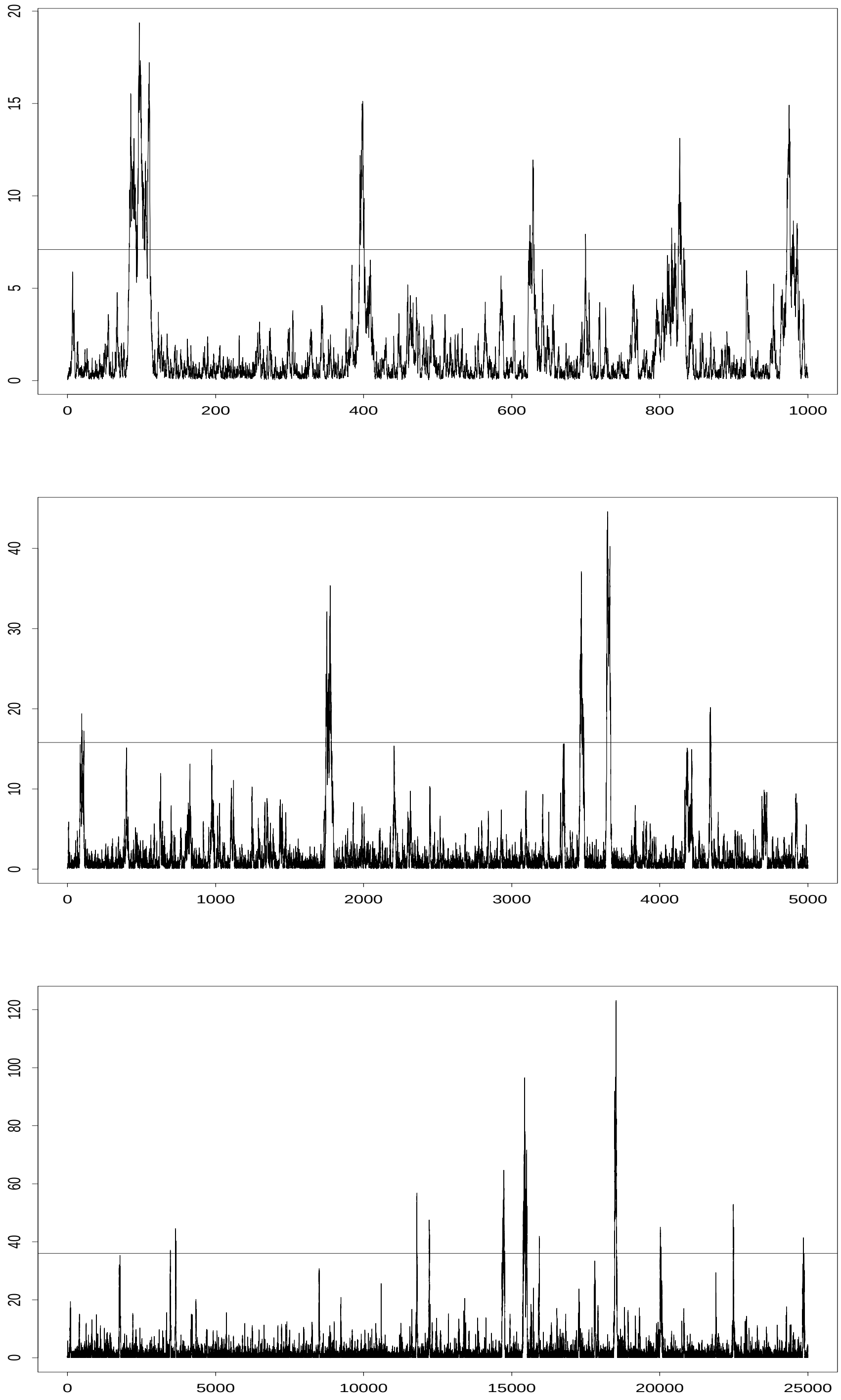

Figure 3.13: The Poisson approximation for ε-upcrossings of the GIG diffusion with parameters $\gamma = 0.5, \chi = 0, \psi = 1, \sigma = 1, \lambda = -1$ as in Figure 3.11. The threshold increases with the sample size. For the calculation of the thresholds we used $\tau = 10$, i.e. on average there are 10 ε–upcrossings for large t and fixed small $\varepsilon > 0$. The first figure shows a realisation of the process X_t for $0 \le t \le 1\,000$, the last two figures represent continuations of this realisation to $t = 5\,000$ and $t = 25\,000$, respectively.

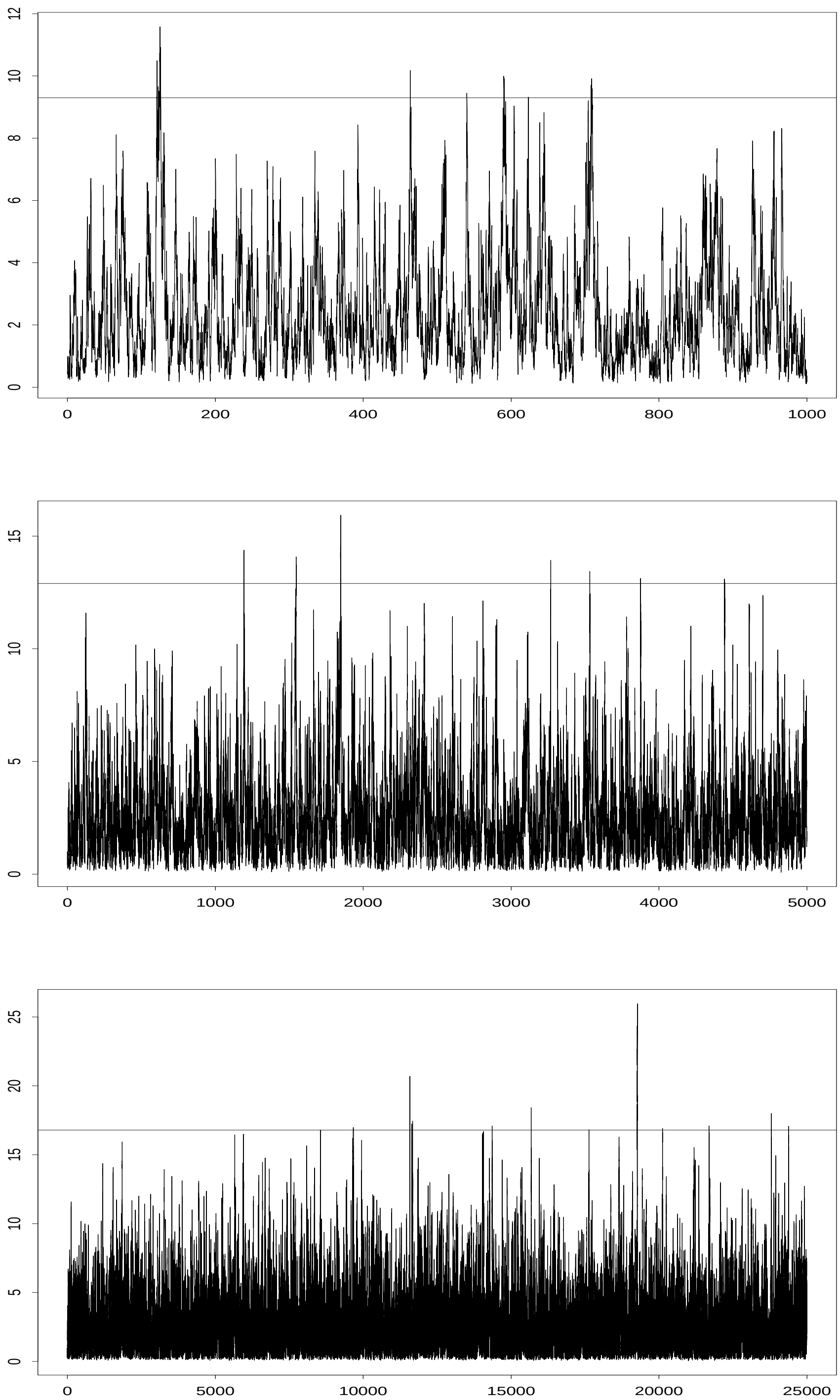

Figure 3.14: The Poisson approximation for ε-upcrossings of the GIG diffusion with parameters $\gamma = 0.5, \chi = \psi = \sigma = \lambda = 1$ as in Figures 3.12. The threshold increases with the sample size. For the calculation of the thresholds we used $\tau = 10$, i.e. on average there are 10 ε–upcrossings for large t and fixed small $\varepsilon > 0$. The first figure shows a realisation of the process X_t for $0 \leq t \leq 1\,000$, the last two figures represent continuations of this realisation to $t = 5\,000$ and $t = 25\,000$, respectively.

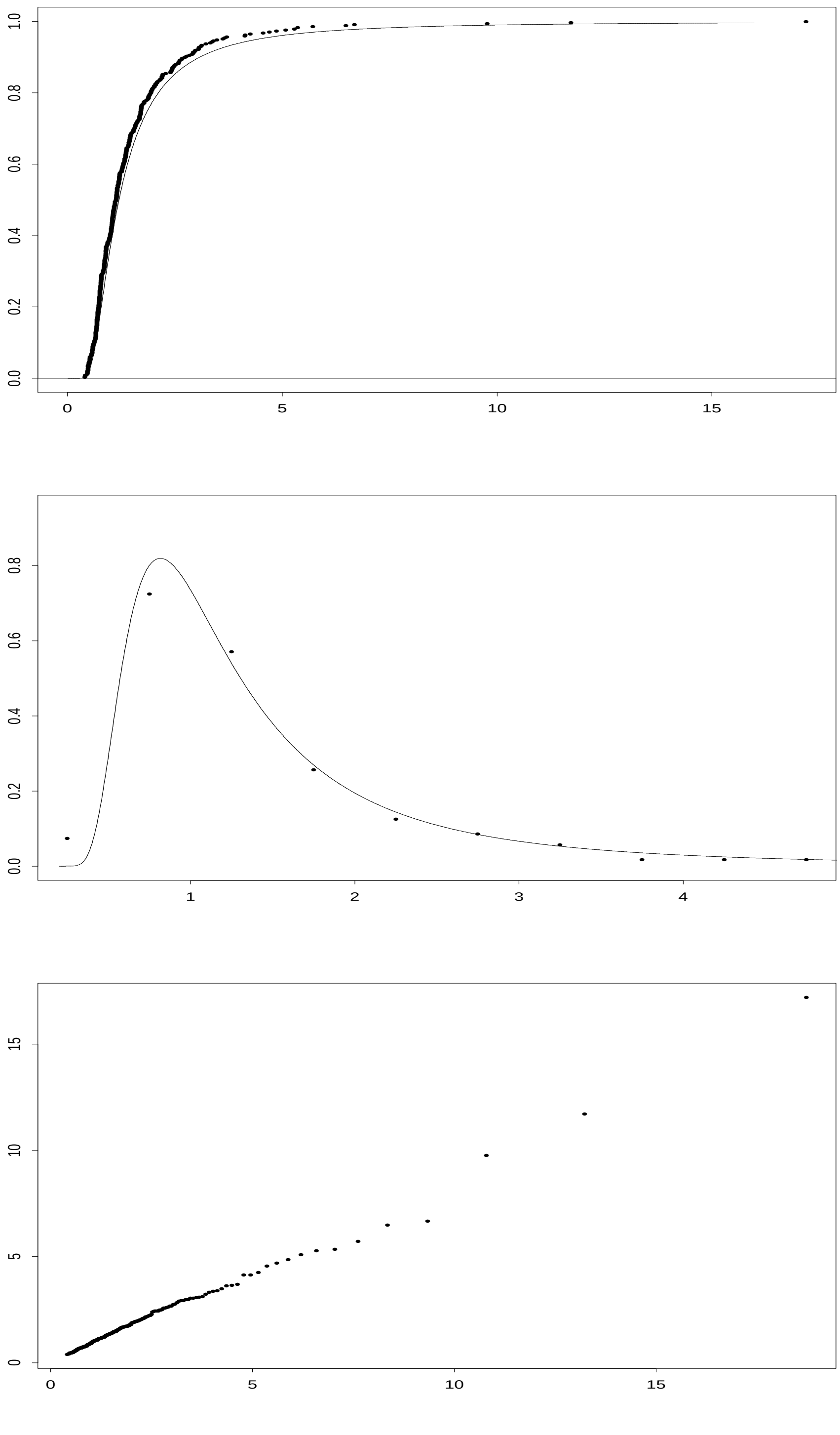

Figure 3.15: The empirical distribution function (top), the empirical density (middle) and the QQ-plot (bottom) of the normalised maxima of the GIG model and the Frechet distribution function and density (solid line), based on 350 simulations with $t = 25\,000$ and parameters $\gamma = 0.5, \chi = 0, \psi = 1, \sigma = 1, \lambda = -1$ as in Figure 3.11.

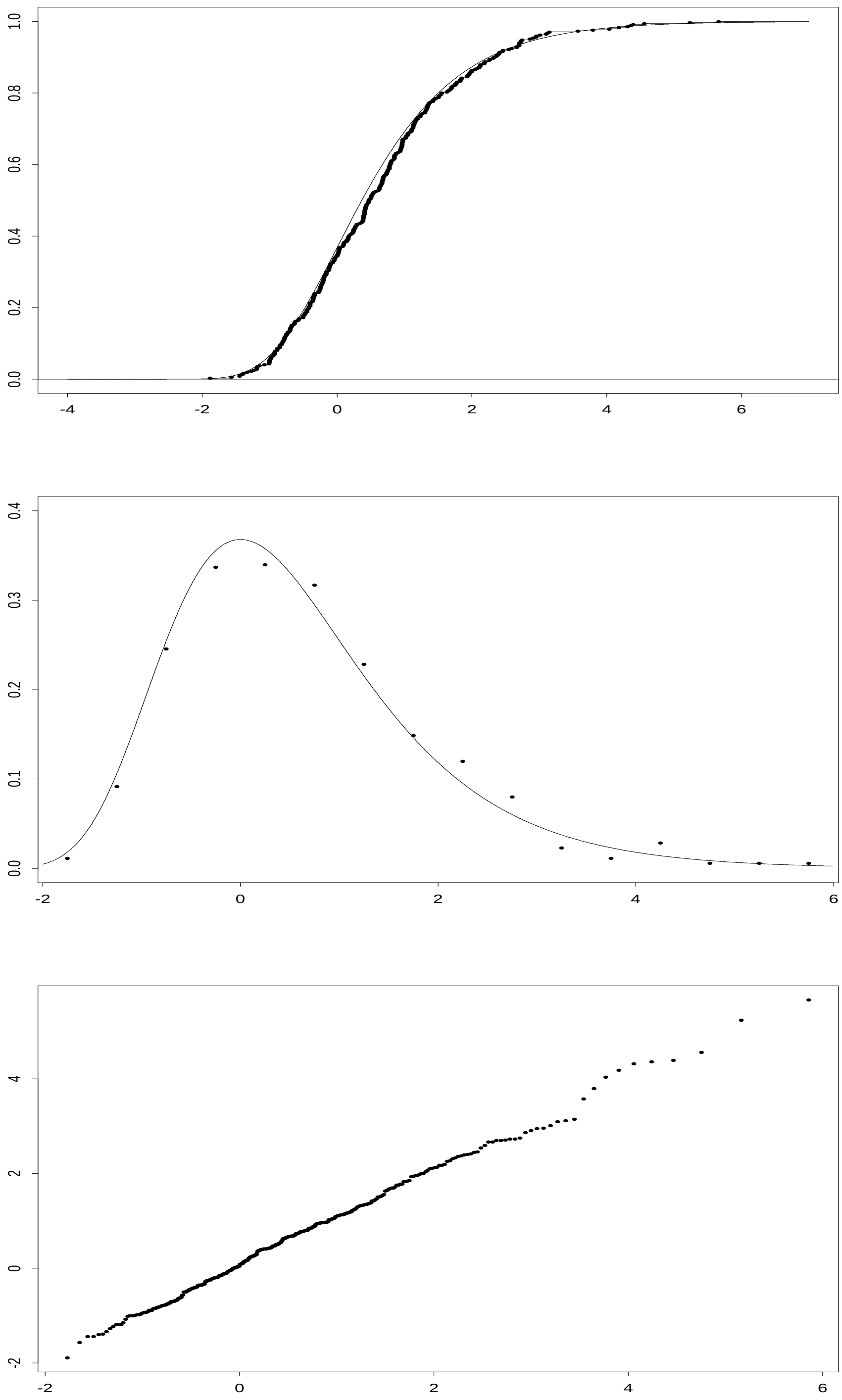

Figure 3.16: The empirical distribution function (top), the empirical density (middle) and the QQ-plot (bottom) of the normalised maxima of the GIG model and the Gumbel distribution function and density (solid line), based on 350 simulations with $t = 25\,000$ and parameters $\gamma = 0.5, \chi = \psi = \sigma = \lambda = 1$ as in Figure 3.12.

The Vasicek model: $dX_t = (c - d\,X_t)\,dt \,+\, \sigma\,dW_t, \quad \sigma > 0,\ d > 0,\ c \in \mathbb{R}$

$$\left(X_t = \tfrac{c}{d} + (x - \tfrac{c}{d})e^{-dt} + \sigma \int_0^t e^{-d(t-s)} dW_s\,,\ t \geq 0 \right)$$

$$h(x) = \frac{1}{\sqrt{2\pi\sigma^2/2}} \exp\{-\frac{1}{2}\frac{(x-c/d)^2}{\sigma^2/2d}\},\ x \in \mathbb{R}$$
$$\overline{F}(x) \sim d\frac{(x-c/d)^2}{\sigma^2/2d}\,\overline{H}(x), \quad x \to \infty$$
$$a_t = \frac{\sigma}{2\sqrt{d\ln t}}$$
$$b_t = \frac{\sigma}{\sqrt{d}}\sqrt{\ln t} + \frac{c}{d} + \frac{\sigma}{4\sqrt{d}}\frac{\ln\ln t + \ln(d^2/\pi)}{\sqrt{\ln t}}$$
$$a_t^{-1}(M_t^X - b_t) \xrightarrow{d} \Lambda$$

The Cox-Ingersoll-Ross model:

$$dX_t = (c - d\,X_t)\,dt \,+\, \sigma\,\sqrt{X_t}\,dW_t, \quad \sigma > 0,\ d > 0,\ c \geq \sigma^2/2$$

$$h(x) = \frac{(2d/\sigma^2)^{2c/\sigma^2}}{\Gamma(2c/\sigma^2)} x^{-1+2c/\sigma^2}\, e^{-2dx/\sigma^2},\ x > 0$$
$$\overline{F}(x) \sim \frac{2d^2}{\sigma^2}\,x\,\overline{H}(x), \quad x \to \infty$$
$$a_t = \frac{\sigma^2}{2d}$$
$$b_t = \frac{\sigma^2}{2d}\left(\ln t + \frac{2c}{\sigma^2}\ln\ln t + \ln\left(\frac{d}{\Gamma(2c/\sigma^2)}\right)\right)$$
$$a_t^{-1}(M_t^X - b_t) \xrightarrow{d} \Lambda$$

The generalised Cox-Ingersoll-Ross model:

$$dX_t = (c - d\,X_t)\,dt \,+\, \sigma\,X_t^{\gamma}\,dW_t, \quad \gamma > 1/2$$

$(a)\quad 1/2 < \gamma < 1: \quad \sigma > 0,\ d \geq 0,\ c > 0$

$$h(x) = const\,x^{-2\gamma}e^{-\frac{2}{\sigma^2}(\frac{c}{2\gamma-1}x^{-(2\gamma-1)} + \frac{d}{2-2\gamma}x^{2-2\gamma})},\ x > 0$$
$$\overline{F}(x) \sim const\,x^{2(1-\gamma)}\overline{H}(x), \quad x \to \infty$$
$$a_t = \tfrac{\sigma^2}{2d}\left(\tfrac{\sigma^2(1-\gamma)}{d}\ln t\right)^{\frac{2\gamma-1}{2-2\gamma}}$$
$$b_t = \left(\frac{\sigma^2(1-\gamma)}{d}\ln t\right)^{\frac{1}{2-2\gamma}}\left(1 - \frac{2\gamma-1}{2(1-\gamma)}\ln(\frac{\sigma^2(1-\gamma)}{d}\ln t)/\ln t\right)$$
$$+\frac{\sigma^2}{2d}\left(\frac{\sigma^2(1-\gamma)}{d}\ln t\right)^{\frac{2\gamma-1}{2-2\gamma}}\ln\left(\frac{2d}{A\sigma^2}\right)$$
$$a_t^{-1}(M_t^X - b_t) \xrightarrow{d} \Lambda$$

$(b)\quad \gamma = 1: \quad \sigma > 0,\ d > -\sigma^2/2,\ c > 0$

$$\left(X_t = e^{-(d+\frac{\sigma^2}{2})t+\sigma W_t}(x + c\int_0^t e^{(d+\frac{\sigma^2}{2})s-\sigma W_s}ds),\ t \geq 0\right)$$

$$h(x) = \left(\frac{\sigma^2}{2c}\right)^{-\frac{2d}{\sigma^2}-1}\Gamma\left(\frac{2d}{\sigma^2}+1\right)^{-1} x^{-2d/\sigma^2-2}e^{-\frac{2c}{\sigma^2}x^{-1}}$$

$$\overline{F}(x) \sim \frac{(2d+\sigma^2)^2}{2\sigma^2}\overline{H}(x), \quad x \to \infty$$

$$a_t \sim \left(c\left(\frac{\sigma^2}{2c}\right)^{-\frac{2d}{\sigma^2}}\Gamma\left(\frac{2d}{\sigma^2}+1\right)^{-1}\left(\frac{2d}{\sigma^2}+1\right)t\right)^{1/(1+2d/\sigma^2)}$$

$$b_t = 0$$

$$a_t^{-1}(M_t^X - b_t) \xrightarrow{d} \Phi_{1+2d/\sigma^2}$$

$(c)\ \gamma > 1: \quad \sigma > 0,\ d \in \mathbb{R},\ c > 0$

$$h(x) = const\, x^{-2\gamma}e^{-\frac{2}{\sigma^2}(\frac{c}{2\gamma-1}x^{-(2\gamma-1)}+\frac{d}{2-2\gamma}x^{2-2\gamma})},\ x > 0$$

$$\overline{F}(x) \sim const\, x^{2(1-\gamma)}\overline{H}(x), \quad x \to \infty$$

$$a_t \sim t/A,$$

$$\text{where } A = \frac{2}{\sigma^2}\int_0^\infty t^{-2\gamma}\exp\left\{-\frac{2}{\sigma^2}\left(\frac{c}{2\gamma-1}t^{-(2\gamma-1)}+\frac{d}{2-2\gamma}t^{2-2\gamma}\right)\right\}dt$$

$$b_t = 0$$

$$a_t^{-1}(M_t^X - b_t) \xrightarrow{d} \Phi_1$$

The Normal Inverse Gaussian Diffusion:

$$dX_t = \sigma\frac{(1+(X_t-\mu)^2/\delta^2)^{1/4}}{\sqrt{K_1(\alpha\delta\sqrt{1+(X_t-\mu)^2/\delta^2})}}\exp\{-\frac{1}{2}\beta X_t\},\ \ \alpha,\beta,\sigma,\delta,\mu$$

where $\mu \in \mathbb{R},\ \delta > 0,\ 0 \leq |\beta| < \alpha,\ \sigma > 0$ and

$$K_1(\gamma) = \frac{1}{2}\int_0^\infty \exp\{-\frac{1}{2}\gamma(t+t^{-1})\}dt$$

$$h(x) = \frac{\alpha}{\pi}\exp\{\delta\sqrt{\alpha^2-\beta^2}-\beta\mu\}\frac{K_1(\alpha\delta\sqrt{1+(x-\mu)^2/\delta^2})}{\sqrt{1+(x-\mu)^2/\delta^2}}\exp\{\beta z\}$$

$$a_t = \frac{\alpha\sigma^2}{2\pi}\exp\{\delta\sqrt{\alpha^2-\beta^2}-\beta\mu\}t$$

$$b_t = 0$$

$$a_t^{-1}(M_t^X - b_t) \xrightarrow{d} \Phi_1$$

The Generalised Inverse Gaussian Diffusion:

$$dX_t = \tfrac{1}{4}\sigma^2 X_t^{2\gamma-2}\left(\psi + 2(2\gamma+\lambda-1)X_t - \chi X_t^2\right) + \sigma\, X_t^{\gamma} dW_t,\ \ \sigma>0,\ \ \gamma>0$$

(a) $\chi = 0,\ \psi > 0,\ \lambda < \min\{0, 2(1-\gamma)\}$:

$$h(x) = \frac{(\psi/2)^{-\lambda}}{\Gamma(-\lambda)} x^{\lambda-1} \exp\left\{-\frac{1}{2}\psi x^{-1}\right\},\ x>0$$

$$\overline{F}(x) \sim \frac{\sigma^2(-\lambda)(2-2\gamma-\lambda)}{2} x^{2\gamma-2}\overline{H}(x), \quad x\to\infty$$

$$a_t \sim \left(\frac{\sigma^2(\psi/2)^{-\lambda}(2-2\gamma-\lambda)}{\Gamma(-\lambda)}\, t\right)^{1/(2-2\gamma-\lambda)}$$

$$b_t = 0$$

$$a_t^{-1}(M_t^X - b_t) \xrightarrow{d} \Phi_{2-2\gamma-\lambda}$$

(b) $\psi = 0,\ \chi > 0,\ \lambda > \max\{0, 2(1-\gamma)\}$:

$$h(x) = \frac{(\chi/2)^{-\lambda}}{\Gamma(\lambda)} x^{\lambda-1} \exp\left\{-\frac{1}{2}\chi x\right\},\ x>0$$

$$\overline{F}(x) \sim \frac{\sigma^2}{2}\left(\frac{\chi}{2}\right)^{2\lambda+3} x^{2\gamma}\overline{H}(x), \quad x\to\infty$$

$$a_t = 2/\chi$$

$$b_t = \frac{2}{\chi}\left(\ln t + (2\gamma+\lambda-1)\ln\ln t + \ln\left(\frac{\sigma^2(\chi/2)^{2-2\gamma}}{2\Gamma(\lambda)}\right)\right)$$

$$a_t^{-1}(M_t^X - b_t) \xrightarrow{d} \Lambda$$

(c) $\chi > 0,\ \psi > 0,\ \lambda\in\mathbb{R}$:

$$h(x) = \frac{(\chi/\psi)^{\lambda/2}}{2K_\lambda\left(\sqrt{\chi\psi}\right)} x^{\lambda-1} \exp\left\{-\frac{1}{2}(\chi x + \psi x^{-1})\right\},\ x>0\,,$$

$$\overline{F}(x) \sim \frac{\sigma^2}{2}\left(\frac{2}{\chi}\right)^{-2} x^{2\gamma}\overline{H}(x), \quad x\to\infty$$

$$a_t = 2/\chi$$

$$b_t = \frac{2}{\chi}(\ln t + (2\gamma+\lambda-1)\ln\ln t + \ln(\frac{\sigma^2(\chi/2)^{2-2\gamma-\lambda}}{4K_\lambda(\sqrt{\chi\psi})}(\frac{\chi}{\psi})^{\lambda/2}))$$

$$a_t^{-1}(M_t^X - b_t) \xrightarrow{d} \Lambda$$

Appendix

A1 Classical Extreme Value Theory

There are three extreme value distribution functions (up to affine transformations):

$$\begin{aligned} &\text{Fréchet:} && \Phi_\alpha(x) = \exp\{-x^{-\alpha}\} I_{(0,\infty)}(x)\,, \quad \alpha > 0\,. \\ &\text{Weibull:} && \psi_\alpha(x) = \exp\{-(-x)^{\alpha}\} I_{(-\infty,0)}(x)\,, \quad \alpha > 0\,. \\ &\text{Gumbel:} && \Lambda(x) = \exp\{-e^{-x}\}\,, \quad x \in \mathbb{R}\,. \end{aligned} \tag{A.1}$$

The distribution function F belongs to the maximum domain of attraction of an extreme value distribution function Q ($F \in \mathrm{MDA}(Q)$) if

$$\lim_{t\to\infty} F^t(a_t x + b_t) = Q(x)\,, \quad x \in \mathbb{R}\,.$$

$\mathrm{MDA}(\psi_\alpha)$ contains only distribution functions with finite right endpoint. Since all the examples we treat in this work are stochastic processes with state space unbounded above, we only present results on $\mathrm{MDA}(\Phi_\alpha)$ and $\mathrm{MDA}(\Lambda)$.

Proposition A1.1 *$F \in \mathrm{MDA}(\Phi_\alpha)$ if and only if $\overline{F}(x) = x^{-\alpha}\ell(x)$, where ℓ is a slowly varying function (i.e. $\lim_{x\to\infty} \ell(xt)/\ell(x) = 1 \quad \forall\, t > 0$).*

Proposition A1.2 *$F \in \mathrm{MDA}(\Lambda)$ if $\overline{F}$ has the representation*

$$\overline{F}(x) = c \exp\left\{-\int_z^x \frac{1}{a(t)} dt\right\}\,, \quad x > z\,, \tag{A.2}$$

where $c > 0$ and $a(\cdot)$ is an absolutely continuous function with Lebesgue density a' such that $\lim_{x\uparrow r} a'(x) = 0$. □

A distribution function with representation (A.2) is called a *von Mises function.* The function $a(\cdot)$ is called *auxiliary function.*

Proposition A1.3 *(a) If $F \in \mathrm{MDA}(\Phi_\alpha)$, then the norming constants a_t can be chosen such that*

$$\overline{F}(a_t) = t^{-1} \quad \text{and} \quad b_t = 0\,.$$

(b) If $F \in \mathrm{MDA}(\Lambda)$, then the norming constants a_t and b_t can be chosen such that

$$\overline{F}(b_t) = t^{-1} \quad \text{and} \quad a_t = a(b_t).$$

(c) The norming constants are not unique. If

$$\lim_{t\to\infty} F^t(a_t x + b_t) = N(x)\,, \quad x \in \mathbb{R}\,,$$

and $\widetilde{a}_t \sim a_t$ and $\widetilde{b}_t - b_t = o(a_t)$, then

$$\lim_{t\to\infty} F^t(\widetilde{a}_t x + \widetilde{b}_t) = N(x)\,, \quad x \in \mathbb{R}\,. \qquad \square$$

Proposition A1.4 describes the increase of M_t^X.

Corollary A1.4 *(a) If $F \in \mathrm{MDA}(\Phi_\alpha)$, then*

$$\frac{M_t^X}{a_t} \quad \stackrel{d}{\to} \quad \Phi_\alpha\,, \quad t \uparrow r\,.$$

(b) If $F \in \mathrm{MDA}(\Lambda)$, then

$$\frac{M_t^X - b_t}{a_t} \quad \stackrel{d}{\to} \quad \Lambda\,, \quad t \uparrow r\,.$$

In particular,

$$\frac{M_t^X}{b_t} \quad \stackrel{P}{\to} \quad 1\,, \quad t \uparrow r\,. \qquad \square$$

Definition A1.5 *Let F and G be distribution functions with right endpoint $r \le \infty$. If*

$$\lim_{x\uparrow r} \frac{\overline{F}(x)}{\overline{G}(x)} = c \in (0,\infty)\,,$$

then F and G are called tail-equivalent.

Proposition A1.6 *Let F and G be tail-equivalent distribution functions with right endpoint $r \le \infty$ and $\overline{F}(x) \sim c\overline{G}(x)$ as $x \uparrow r$. Assume that $G \in \mathrm{MDA}(Q)$ with norming functions $\widetilde{a}_t > 0$ and $\widetilde{b}_t \in \mathbb{R}$ such that*

$$\lim_{t\to\infty} G^t(\widetilde{a}_t x + \widetilde{b}_t) = Q(x)\,, \quad x \in \mathbb{R}\,,$$

for an extreme value distribution function Q. Then $F \in \mathrm{MDA}(Q)$ and

$$\lim_{t\to\infty} F^t(a_t x + b_t) = Q(x)\,, \quad x \in \mathbb{R}\,,$$

where

$$\begin{aligned} a_t &= \widetilde{a}_t \qquad \textit{and} \quad b_t = \widetilde{b}_t + \widetilde{a}_t \ln c \quad \textit{if} \quad Q = \Lambda\,, \\ a_t &= c^{1/\alpha}\widetilde{a}_t \quad \textit{and} \quad b_t = b_t = 0 \qquad\quad \textit{if} \quad Q = \Phi_\alpha\,. \end{aligned}$$

□

Proposition A1.7 *Let* $F(x) = G(cx + d)$ *for any* $c > 0$ *and* $d \in \mathbb{R}$. *Assume that* $G \in \mathrm{MDA}(Q)$ *with norming functions* $\widetilde{a}_t > 0$ *and* $\widetilde{b}_t \in \mathbb{R}$ *such that*

$$\lim_{t\to\infty} G^t(\widetilde{a}_t x + \widetilde{b}_t) = Q(x)\,, \quad x \in \mathbb{R},$$

for an extreme value distribution function Q. *Then* $F \in \mathrm{MDA}(Q)$ *and*

$$\lim_{t\to\infty} F^t(a_t x + b_t) = Q(x)\,, \quad x \in \mathbb{R},$$

where $a_t = \widetilde{a}_t/c$ *and* $b_t = (\widetilde{b}_t - d)/c\,.$ □

A2 Computation of Normalising Constants

In this section we compute explicitly the normalising constants a_n and b_n which we presented in Chapter 3 in Examples 3.3.1 and 3.3.3. Recall that the constants a_n and b_n have to be chosen such that the underlying distribution function F belongs to the maximum domain of attraction of one of the three extreme value distribution functions, i.e.

$$\lim_{n\to\infty} F^n(a_n x + b_n) = Q(x)\,,$$

where Q is either a Fréchet, Weibull or Gumbel distribution function. We will proceed in a similar way as in Embrechts et al. (1997, Chapter 3.3) or Resnick (1988, Chapter 1.5).

Concerning Example 3.3.1:
Suppose the tail of F is given by

$$\overline{F}(x) \sim x\,\exp\{-x^2/2\}\,, \quad x \to \infty\,.$$

Due to Proposition A1.6 it is sufficient to compute the normalising constants a_n and b_n of the distribution function $G(x) = 1 - x\,\exp\{-x^2/2\}$ for x large enough. Notice that G is a von Mises function with auxiliary function $a(x) = (x + 1/x)^{-1}$. Thus $G \in MDA(\Lambda)$ and according to

Proposition A1.3(b) $b_n = G^{\leftarrow}(1-1/n)$ and $a_n = a(b_n)$. Hence look for a solution of $-\ln \overline{G}(b_n) = \ln n$, i.e.

$$\frac{1}{2}b_n^2 - \ln b_n = \ln n \,. \tag{A.1}$$

Since $b_n \to \infty$ we see by dividing through that

$$b_n \sim \sqrt{2\ln n}$$

and consequently

$$b_n = \sqrt{2\ln n} + r_n \,, \tag{A.2}$$

where $r_n = o(\sqrt{\ln n})$. By Proposition A1.3(b) we may hence choose $a_n = 1/\sqrt{2\ln n}$. Substituting (A.2) in (A.1) we obtain

$$\ln n + \sqrt{2\ln n} r_n + \frac{1}{2} r_n^2 - \ln(\sqrt{2\ln n} + r_n) = \ln n \,,$$

i.e.

$$\sqrt{2\ln n} r_n + \frac{1}{2} r_n^2 - \ln\left(\sqrt{2\ln n}(1 + r_n/\sqrt{2\ln n})\right) = 0 \,,$$

i.e.

$$\sqrt{2\ln n} r_n + \frac{1}{2} r_n^2 - \frac{1}{2}\ln(2\ln n) - \ln(1 + r_n/\sqrt{2\ln n}) = 0 \,. \tag{A.3}$$

Divide through by $\sqrt{2\ln n}\, r_n$ and we get

$$1 + \frac{1}{2}\frac{r_n}{\sqrt{2\ln n}} - \frac{1}{2}\frac{\ln(2\ln n)}{\sqrt{2\ln n} r_n} - \frac{\ln(1 + r_n/\sqrt{2\ln n})}{\sqrt{2\ln n} r_n} = 0 \,.$$

Because $r_n = o(\sqrt{lnn})$ and since the last term is asymptotic to $\dfrac{r_n/\sqrt{2\ln n})}{\sqrt{2\ln n} r_n} = \dfrac{1}{2\ln n} \to 0$ as $n \to \infty$ we obtain that

$$r_n = \frac{1}{2}\frac{\ln(2\ln n)}{\sqrt{2\ln n}} + s_n \,, \tag{A.4}$$

where $s_n = o(\ln\ln n/\sqrt{\ln n})$. In fact $s_n = o((\ln n)^{-1/2})$. To see this observe that (A.3) implies

$$\sqrt{2\ln n} r_n - \frac{1}{2}\ln(2\ln n) = \ln(1 + r_n/\sqrt{2\ln n}) - \frac{1}{2} r_n^2 \to 0 \,, \quad n \to \infty \,. \tag{A.5}$$

Plugging now (A.4) in the lhs of (A.5) we get

$$\sqrt{2\ln n}\, s_n \to 0 \,, \quad n \to \infty \,.$$

Therefore

$$\frac{b_n - (\sqrt{2\ln n} + \ln(2\ln n)/2\sqrt{2\ln n})}{a_n} = \frac{s_n}{a_n} \to 0\,, \quad n \to \infty$$

and hence, because of Proposition A1.3(c),

$$b_n = \sqrt{2\ln n} + \frac{\ln(2\ln n)}{\sqrt{2\ln n}}$$

is an acceptable choice.

Concerning Example 3.3.3:

Consider the tail

$$\overline{F}(x) \sim x^{-2\gamma+1} \exp\{-\frac{d}{\sigma^2(1-\gamma)} x^{2(1-\gamma)}\} \tag{A.6}$$

of a distribution function F, where $1/2 < \gamma < 1, c > 0, d \geq 0$ and $\sigma > 0$. Again with the same arguments as in the last case we may assume w.l.o.g. that the tail of the distribution function F equals the rhs of (A.6). It is straightforward that F is a von Mises function, hence $F \in MDA(\Lambda)$ and the auxiliary function $a(x)$ satisfies

$$a(x) = \left((2\gamma-1)\frac{1}{x} + \frac{2d}{\sigma^2}x^{1-2\gamma}\right)^{-1} \sim \frac{\sigma^2}{2d}x^{2\gamma-1}\,, \quad x \to \infty\,.$$

We now show that

$$a_n = \frac{\sigma^2}{2d}\left(\frac{\sigma^2(1-\gamma)}{d}\ln n\right)^{\frac{2\gamma-1}{2-2\gamma}} \tag{A.7}$$

and

$$b_n = \left(\frac{\sigma^2(1-\gamma)}{d}\ln t\right)^{\frac{1}{2-2\gamma}} \left(1 - \frac{2\gamma-1}{2(1-\gamma)} \frac{\ln\left(\frac{\sigma^2(1-\gamma)}{d}\ln t\right)}{\ln t}\right)^{\frac{1}{2-2\gamma}} \tag{A.8}$$

are acceptable choices of the norming constants. By Proposition A1.3(b) we have to solve

$$\overline{F}(b_n) = \frac{1}{n}\,,$$

i.e.

$$b_n^{1-2\gamma} \exp\{-\frac{d}{\sigma^2(1-\gamma)} b_n^{2(1-\gamma)}\} = \frac{1}{n}$$

and logarithming both sides gives

$$\frac{d}{\sigma^2(1-\gamma)} b_n^{2(1-\gamma)} - (1-2\gamma)\ln b_n = \ln n\,. \tag{A.9}$$

We will construct an expansion of b_n and indicate how many terms are necessary. Since $b_n \to \infty$ we see by dividing left and right sides of (A.9) by $b_n^{2(1-\gamma)}$ that as $n \to \infty$

$$b_n \sim \left(\frac{\sigma^2(1-\gamma)}{d} \ln n \right)^{1/2(1-\gamma)} . \tag{A.10}$$

Since $a_n = a(b_n) \sim \frac{\sigma^2}{2d} b_n^{2\gamma-1} = \frac{\sigma^2}{2d} \left(\frac{\sigma^2(1-\gamma)}{d} \ln n \right)^{2\gamma-1/2(1-\gamma)}$ we see that an acceptable choice for a_n is

$$a_n = \frac{\sigma^2}{2d} \left(\frac{\sigma^2(1-\gamma)}{d} \ln n \right)^{2\gamma-1/2(1-\gamma)} .$$

Next we have to study b_n in more detail. In order to do this we define the auxiliary sequence

$$v_n = b_n^{2(1-\gamma)} . \tag{A.11}$$

From (A.9) we get thus

$$\frac{d}{\sigma^2(1-\gamma)} v_n - \frac{(1-2\gamma)}{2(1-\gamma)} \ln v_n = \ln n . \tag{A.12}$$

From (A.10) we see that

$$v_n = \frac{\sigma^2(1-\gamma)}{d} \ln n + r_n , \tag{A.13}$$

where $r_n = o(\ln n)$. Substituting (A.13) into (A.12) we find

$$\frac{d}{\sigma^2(1-\gamma)} r_n = \frac{(1-2\gamma)}{2(1-\gamma)} \ln \left(\frac{\sigma^2(1-\gamma)}{d} \ln n \right) + \frac{(1-2\gamma)}{2(1-\gamma)} \ln \left(1 + \frac{d\, r_n}{\sigma^2(1-\gamma) \ln n} \right)$$

and therefore

$$r_n = \frac{\sigma^2(1-2\gamma)}{2d} \ln \left(\frac{\sigma^2(1-\gamma)}{d} \ln n \right) + s_n , \tag{A.14}$$

where $s_n = \frac{\sigma^2}{2d}(1-2\gamma) \ln(1 + \frac{d\, r_n}{\sigma^2(1-\gamma)\ln n}) \sim \frac{\sigma^2}{2d}(1-2\gamma) \frac{d\, r_n}{\sigma^2(1-\gamma)\ln n} \to 0$ as $n \to \infty$, i.e. $s_n = o(1)$. Plugging all this together we obtain

$$\begin{aligned} b_n &= v_n^{1/2(1-\gamma)} = \left(\frac{\sigma^2(1-\gamma)}{d} \ln n + \frac{\sigma^2}{2d}(1-2\gamma) \ln(\frac{\sigma^2(1-\gamma)}{d} \ln n) + s_n \right)^{1/2(1-\gamma)} \\ &= \left(\frac{\sigma^2(1-\gamma)}{d} \ln n + \frac{\sigma^2}{2d}(1-2\gamma) \ln(\frac{\sigma^2(1-\gamma)}{d} \ln n) \right)^{1/2(1-\gamma)} \\ &\qquad \cdot \left(1 + \frac{s_n}{\frac{\sigma^2(1-\gamma)}{d} \ln n + \frac{\sigma^2}{2d}(1-2\gamma) \ln(\frac{\sigma^2(1-\gamma)}{d} \ln n)} \right)^{1/2(1-\gamma)} . \end{aligned}$$

A Taylor expansion and the fact that $1/2 < \gamma < 1$ yields that

$$\frac{b_n - (\frac{\sigma^2(1-\gamma)}{d}\ln n + \frac{\sigma^2}{2d}(1-2\gamma)\ln(\frac{\sigma^2(1-\gamma)}{d}\ln n))^{1/2(1-\gamma)}}{a_n} \sim \frac{2d}{\sigma^2}s_n \to 0\,, \quad n\to\infty.$$

By Proposition A1.3(c)

$$b_n = \left(\frac{\sigma^2(1-\gamma)}{d}\ln n + \frac{\sigma^2}{2d}(1-2\gamma)\ln(\frac{\sigma^2(1-\gamma)}{d}\ln n)\right)^{1/2(1-\gamma)}$$

is an acceptable choice.

A3 Some Extreme Value Theory for Markov Chains

The theorem below gives the extremal properties of a fairly large class of stationary Markov chains. The original version can be found in Perfekt (1994, Theorem 3.2, p. 538). We present a simplified version of Perfekt's result which can be directly applied to our situation in Chapter 2.

Theorem A3.1 *Suppose $(X_n)_{n\in\mathbb{N}}$ is a stationary Markov chain which satisfies for some $\gamma \in (-\infty,\infty)$ the following properties*

(i)

$$\lim_{u\uparrow x_F}\frac{1-F(u+g(u)x)}{1-F(u)} = (1-\gamma\,x)_+^{1/\gamma},\ x\in(-\infty,\infty),$$

where F is the stationary distribution function, $x_F := \sup\{x\,;\,F(x)<1\}$, $y_+ := \max\{0,y\}$ and

$$\begin{aligned} x_F = \infty \quad &\text{and} \quad g(u) = -\gamma\,u \qquad \text{if } \gamma<0\\ x_F < \infty \quad &\text{and} \quad g(u) = \gamma(x_F-u) \quad \text{if } \gamma>0 \end{aligned}$$

If $\gamma = 0$, then the auxiliary function g is unique up to asymptotic equivalence and strictly positive on (x_0, x_F) for some $x_0 < x_F$.

(ii)

$$\lim_{u\to x_F} P\Big((1-\gamma\,\frac{(X_1-u)}{g(u)})_+^{-1/\gamma} \le x \,|\, X_0 = u\Big) = H(x)$$

for some distribution function H on $[0,\infty)$.

Let furthermore $(A_n)_{n\in\mathbb{N}}$ be an i.i.d. sequence with marginal distribution function H and let Y_0 be a random variable independent of $(A_n)_{n\in\mathbb{N}}$. Define the tail chain $(Y_n)_{n\in\mathbb{N}}$ by $Y_n = A_n\,Y_{n-1}$

for $n \geq 1$ *and denote by* P^μ *the law of* $(Y_n)_{n\in\mathbb{N}}$ *when* Y_0 *has distribution* μ*. Assume* $\mu(dx) = x^{-2}dx$, $x > 1$ *and let* $(u_n(\tau))$ *be a sequence which satisfies*

$$\lim_{n\to\infty} n(1 - F(u_n(\tau))) \to \tau\,.$$

(a) Assume $D(u_n(\tau))$ *holds for each* $\tau > 0$*. If for some* τ_0 *there is a* $D(u_n(\tau_0))$*-separating sequence* $(p_n)_{n\in\mathbb{N}}$ *such that*

$$\lim_{p\to\infty} \limsup_{n\to\infty} P(\max_{p\leq j\leq p_n} X_j > u_n \,|\, X_0 > u_n) = 0 \tag{A.1}$$

holds with $u_n = u_n(\tau)$ *then* $(X_n)_{n\in\mathbb{N}}$ *has extremal index* θ *given by*

$$\theta = P^\mu(\#\{n \geq 1 \,|\, Y_n > 1\} = 0)\,.$$

(b) Suppose $(X_n)_{n\in\mathbb{N}}$ *has extremal index* $\theta > 0$ *and, for some* $\tau_1 > 0$ *satisfies* $\Delta(u_n(\sigma\tau_1))$ *for each* $\sigma > 0$*. Suppose further there is a* $\Delta(u_n(\tau_1))$*-separating sequence* $(p_n)_{n\in\mathbb{N}}$ *such that (A.1) holds with* $u_n = u_n(\tau_1)$*. Then, for each* $\sigma > 0$, $N_n^{\sigma\tau_1} := \#\{k \in \{1,...,n\} \,|\, k/n \in \cdot, X_k > u_n(\sigma\tau_1)\}$ *converges in distribution to a compound Poisson process* N *with intensity* $\theta\sigma\tau_1$ *and jump probabilities* π_i *given by*

$$\pi_i = \frac{1}{\theta}\Big(P^\mu(\#\{n \geq 1 \,|\, Y_n > 1\} \;=\; i-1) - P^\mu(\#\{n \geq 1 \,|\, Y_n > 1\} \;=\; i)\Big)\,, \quad i \in \mathbb{N}.$$

The next theorem is an extension of Theorem A3.1. In some cases it is easier to apply then the last one.

Theorem A3.2 *(Extension of Theorem 3.2 of Perfekt (1994), p. 543) Suppose* $(X_n)_{n\in\mathbb{N}}$ *is a stationary Markov chain which satisfies*

$$\lim_{u\uparrow x_F} \frac{1 - F(u + g(u)x)}{1 - F(u)} = (1 - \gamma\, x)_+^{1/\gamma}\,, \quad x \in (-\infty, \infty),$$

where F *is the stationary distribution function,* $x_F := \sup\{x\,;\, F(x) < 1\} = \infty$, $y_+ := \max\{0, y\}$ *and*

$$g(u) = -\gamma\, u \text{ for some } \gamma < 0\,.$$

Suppose furthermore that $\inf\{x\,;\, F(x) > 0\} = -\infty$ *and that*

$$\lim_{u\to\infty} P(\frac{X_1}{u} \leq x \,|\, X_0 = u) = H(x)\,,$$

for some distribution function H on $(-\infty,\infty)$. Let $(A_n)_{n\in\mathbb{N}}$ be an i.i.d. sequence with distribution H and define the tail chain through $Y_n = A_nY_{n-1}$, $n \geq 1$, Y_0 being independent of $(A_n)_{n\in\mathbb{N}}$. Then, if $(X_n)_{n\in\mathbb{N}}$ satisfies the conditions in (a) and (b) of Theorem A3.1, the result of the theorem holds with the initial distribution μ given by $\mu(dx) := |\gamma|^{-1}x^{1/\gamma-1}dx$, for $x > 1$.

A4 Classical Markov Chain Theory

Here we give a short introduction to the Markov chain theory we need in this paper. For details we refer to Tweedie (1976) or Meyn and Tweedie (1993) .

Consider a homogeneous Markov chain $(X_n)_{n\in\mathbb{N}}$ on a locally compact complete separable metric space E and $\mathcal{E}$ is the countably generated σ-algebra on E. Let $P = \{P(x,A) := P(X_1 \in A \,|\, X_0 = x)\,;\, x \in E, A \in \mathcal{E}\}$ denote the transition probability kernel and μ the initial distribution. The n-step transition probabilities $P^n(x,A) := P(X_n \in A|X_0 = x)$ can be iterated as

$$P^n(x,A) = \int_E P^{n-1}(y,A)P(x,dy)\,, \quad x \in E, A \in \mathcal{E},\ n = 1,2,\ldots,$$

where $P^0(y,A) = 1_A(y)$. Let P_μ be the corresponding probability measure which makes $(X_n)_{n\in\mathbb{N}}$ to a Markov chain with transition probability P and initial distribution μ. If $\mu = \delta_x$, $x \in E$, write P_x instead of P_μ.

Suppose ϕ is a σ-finite non-trivial measure on $\mathcal{E}$. Then $(X_n)_{n\in\mathbb{N}}$ is called *ϕ-irreducible* if $\sum_{n=1}^\infty P^n(x,A) > 0$ for every $x \in E$ whenever $\phi(A) > 0$. Note that if $(X_n)_{n\in\mathbb{N}}$ is ϕ-irreducible and if ψ is another σ-finite non-trivial measure on $\mathcal{E}$ which is absolutely continuous with respect to ϕ then the process $(X_n)_{n\in\mathbb{N}}$ is ψ-irreducible as well. $(X_n)_{n\in\mathbb{N}}$ is called a ϕ-irreducible *Feller chain* if for each bounded continuous g on E, the function $x \mapsto E(g(X_n)|X_{n-1} = x)$ is continuous in x.

For any $A \in \mathcal{E}$, define $\tau(A) = \inf\{n \geq 1;\, X_n \in A\}$. We call $A \in \mathcal{E}$ *recurrent* if $\sum_{n=1}^\infty P^n(x,A) = \infty$, and *transient* otherwise. A is *Harris recurrent* if $P_x(\tau(A) < \infty) = 1$ for all $x \in E$. By the strong Markov property, this is equivalent to $\{n;\, X_n \in A\}$ being unbounded with probability 1, independent of the initial distribution. Note that if a set is Harris recurrent then it is recurrent. We call A a *regeneration set* if A is Harris recurrent and for some $r > 0$ there exist some $\varepsilon \in (0,1]$ and some probability measure ν on E such that

$$P^r(x,B) \geq \varepsilon\nu(B), \quad x \in A$$

for all $B \in \mathcal{E}$. We call a Markov chain with a regeneration set *Harris (ν-)recurrent* or just a *Harris chain.* There are mainly two situations when a regeneration set exists:

(a) When there is a Harris recurrent one-point set x_0 (one can then take $r = 1, \varepsilon = 1, A = x_0$ and $\nu(B) = P(x_0, B)$).

(b) When, for some $r > 0$, a transition density $f^r(\cdot, \cdot)$ exists (i.e. when $P^r(x, dy) = f^r(x, y)\lambda(dy)$ for some measure λ) together with a Harris recurrent set A and a set S with $0 < \lambda(S) < \infty$ such that $f^r(x, y) \geq \varepsilon > 0$ for any $x \in A, y \in S$.

The process $(X_n)_{n\in\mathbb{N}}$ is called *regenerative* if there exist integer-valued random variables $0 < T_0 < T_1 < T_2 < ...$ which split the sequence up into independent cycles $B_0, B_1, ...$, i.e. if

$$B_0 = \{X_n;\, 0 \leq n < T_0\}, \quad B_1 = \{X_n;\, T_0 \leq n < T_1\},$$
$$B_2 = \{X_n;\, T_1 \leq n < T_2\}, ...$$

are independent and if in addition $B_1, B_2, ...$ have the same distribution. Note that $(T_n)_{n\in\mathbb{N}_0}$ is a renewal process. The process $(X_n)_{n\in\mathbb{N}}$ is called *1-regenerative* if there exists a renewal process $(T_n)_{n\in\mathbb{N}_0}$ which splits $(X_n)_{n\in\mathbb{N}}$ up into 1-dependent cycles $B_0, B_1, ...$ Hence adjacent cycles might be dependent, while cycles seperated by at least one cycle are independent. If $(X_n)_{n\in\mathbb{N}}$ has a regeneration set then a renewal process $(T_n)_{n\in\mathbb{N}_0}$ can be constructed from $(X_n)_{n\in\mathbb{N}}$ which makes $(X_n)_{n\in\mathbb{N}}$ either regenerative or 1-regenerative (see for example Asmussen (1987), p.151) .

The chain $(X_n)_{n\in\mathbb{N}}$ is called *recurrent* if it is ϕ-irreducible and every set in $\mathcal{E}^+ = \{A \in \mathcal{E};\, \phi(A) > 0\}$ is recurrent. The chain $(X_n)_{n\in\mathbb{N}}$ is Harris recurrent if and only if it is ϕ-irreducible and every set in $\mathcal{E}^+ = \{A \in \mathcal{E};\, \phi(A) > 0\}$ is Harris recurrent. If $P_x(\tau(A) < \infty) = 1$ and if $E(\tau(A)|X_0 = x) < \infty$ then A is called *positive.* We call $(X_n)_{n\in\mathbb{N}}$ *positive* if $(X_n)_{n\in\mathbb{N}}$ is ϕ-irreducible and if every set in $\mathcal{E}^+$ is positive; otherwise we call $(X_n)_{n\in\mathbb{N}}$ *null* . If the chain $(X_n)_{n\in\mathbb{N}}$ is Harris recurrent then there exists an invariant measure π, i.e. a σ-finite measure π on $\mathcal{E}$ with the property

$$\pi(A) = \int_E P(x, A)\pi(dx)\,, \quad A \in \mathcal{E}\,.$$

A set $A \in \mathcal{E}$ is called a *small set* if for every $B \in \mathcal{E}$ such that $\phi(B) > 0$, there exists an integer $n \geq 1$ such that $\inf_{x\in A} \sum_{i=1}^n P^i(x, B) > 0$. If $(X_n)_{n\in\mathbb{N}}$ is a ϕ-irreducible Feller chain, then the topological conditions on our space imply that a set $A \in \mathcal{E}$ is small if A is relatively compact and $\phi(A) > 0$.

Let Θ be a class of chains $(Y_n)_{n\in\mathbb{N}}$ on $(E,\mathcal{E})$ which are ϕ-irreducible for some ϕ. A set $A \in \mathcal{E}$ is called a *status set* for Θ if, for each $(Y_n)_{n\in\mathbb{N}} \in \Theta$ with transition law $\{P(x,A)\}$,

(a) $\sum_{n=1}^{\infty} P^n(x,A) < \infty\,, \quad \phi\text{-a.a. } x \in E\,, \quad \text{if } (Y_n)_{n\in\mathbb{N}} \text{ is transient}\,,$

(b) $\lim_{n\to\infty} P^n(x,A) = 0\,, \quad \phi\text{-a.a. } x \in E\,, \quad \text{if } (Y_n)_{n\in\mathbb{N}} \text{ is null}\,.$

This notation is often somewhat abused by calling A a status set for $(X_n)_{n\in\mathbb{N}}$ if A is a status set for Θ and $(X_n)_{n\in\mathbb{N}} \in \Theta$. Note furthermore that a status set A is characterized by the fact that its status is always the same as that of the underlying chain. In general it is difficult to show that a set is status set of Θ. But if we consider $\Theta_s = \{(Y_n)_{n\in\mathbb{N}};\ (Y_n)$ is ϕ-irreducible for some ϕ and $P(x,A)$ is a continuous function of x for every $A \in \mathcal{E}\}$ then every relatively compact set in $\mathcal{E}^+$ is a status set for Θ_s.

Finally the ϕ-irreducible Markov chain $(X_n)_{n\in\mathbb{N}}$ is called *aperiodic* if there does not exist an integer $d \geq 2$ and disjoint cycles $C_1, C_2, ..., C_d \in \mathcal{E}$ such that for any $j \in \{1,...,d\}$

$$P(x, C_{j+1}) = 1\,, \quad x \in C_j\,,$$

with $C_{d+1} = C_1$ and

$$\phi(E \setminus \bigcup_{i=1}^{d} C_i) = 0\,.$$

With these notions we are now ready to state some criteria for finiteness of hitting time moments, positivity, geometric ergodicity and existence of moments which we apply in this paper.

Lemma A4.1 (Theorem 3 of Tweedie (1983)(a))

Suppose that g is a non-negative measurable function on E. If, for some $\varepsilon > 0$ and some $A \in \mathcal{E}$,

$$\int_E g(y)P(x,dy) \leq g(x) - \varepsilon\,, \quad x \in A^c\,, \tag{A.1}$$

then

$$E(\tau_A | X_0 = x) \leq \frac{g(x)}{\varepsilon}\,, \quad x \in A^c\,.$$

Theorem A4.2 (Theorem 9.1 of Tweedie (1976))

Suppose $(X_n)_{n\in\mathbb{N}}$ is ϕ-irreducible, and let g be a non-negative measurable function on E. The chain $(X_n)_{n\in\mathbb{N}}$ is positive if there exists $\varepsilon > 0, \theta < \infty$ and a status set A for $(X_n)_{n\in\mathbb{N}}$ such that (A.1) holds and

$$\int_{A^c} g(y)P(x,dy) \leq \theta\,, \quad x \in A\,. \tag{A.2}$$

Theorem A4.3 (Theorem 4 of Tweedie (1983)(a))
Suppose $(X_n)_{n\in\mathbb{N}}$ is aperiodic positive Harris recurrent. Moreover, g is a non-negative measurable function on E and A is a small set, and that

$$\sup_{x\in A}\int_E g(y)P(x,dy) < \infty\,.$$

If g and A satisfy (A.1) then $(X_n)_{n\in\mathbb{N}}$ is geometric ergodic with stationary distribution π, i.e. there exists a $\rho < 1$ such that

$$\rho^{-n}||P^n(x,\cdot)-\pi(\cdot)|| \to 0\,, \quad n\to\infty\,,$$

for every $x \in E$, where $||\cdot||$ denotes total variation of signed measures on $\mathcal{E}$.

Theorem A4.4 (Theorem 3 of Tweedie (1983)(b))
Suppose that $(X_n)_{n\in\mathbb{N}}$ is aperiodic and Harris recurrent with $\pi(E)=1$. If A is small and if

$$\sup_{x\in A}\int_{A^c} g(y)P(x,dy) < \infty$$

holds for some non-negative measurable g for which there exists $\varepsilon > 0$ such that

$$\int_{A^c} g(y)P(x,dy) \le (1-\varepsilon)g(x)\,, \quad x\in A^c \tag{A.3}$$

and which is bounded away from 0 and ∞ on A, then

$$\int_E g(x)\pi(dx) < \infty\,,$$

and further, for some $\rho < 1$,

$$\int_E ||P^n(x,\cdot)-\pi(\cdot)||_g\pi(dx) = O(\rho^n)\,, \quad n\to\infty\,,$$

where $||\mu||_g := sup_{|h|\le g}\int_E h(y)\mu(dy)$ for any signed measure μ.

Bibliography

[1] Adler, R.J. (1990)
An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes.
Hayward, California: Institute of Mathematical Statistics.

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

[3] Asmussen, S. (1998) Extreme value theory for queues via cycle maxima. *Extremes* **1**(2).

[4] Barndorff-Nielsen, O.E. (1995)
Normal inverse Gaussian processes and the modelling of stock returns. Research Report No. 300.
Department of Theoretical Statistics, University of Aarhus.

[5] Baxter, M. and Rennie, A. (1996) Financial Calculus. Cambridge University Press.

[6] Berman, S.M. (1969) Limiting distribution of the maximum of a diffusion process. *Ann. Math. Stat.* **35**, 319-329.

[7] Berman, S.M. (1992)
Sojourns and Extremes of Stochastic Processes.
Wadsworth, Belmond, California.

[8] Bibby and Sørensen, M. (1995)
A hyperbolic diffusion model for stock prices.
Research Report No. 331. Department of Theoretical Statistics, Institute of Mathematics, University of Aarhus.

[9] Billingsley, P. (1968) *Convergence of Probability Measures.* Wiley, New York.

[10] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987) *Regular Variation.* Cambridge University Press, Cambridge.

[11] Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedastic. *Journ. Econometrics* **31**, 307–327.

[12] Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992) ARCH modelling in finance: a review of the theory and empirical evidence. *Journ. Econometrics* **52**, 5-59.

[13] Borovkov, A. A. (1976) *Stochastic Processes in Queueing Theory.* Springer, New York

[14] Cox, J.C., Ingersoll, J.E. and Ross, S.A. (1985)
A theory of the term structure of interest rates.
Econometrica **53**, 385-408.

[15] Daley, D. and Vere–Jones, D. (1988) *An Introduction to the Theory of Point Processes.* Springer, Berlin.

[16] Davis, R.A. (1982)
Maximum and minimum of one-dimensional diffusions.
Stoch. Proc. Appl. **13**, 1-9.

[17] Davis, R.A. and Mikosch, T. (1998) The sample autocorrelations of heavy-tailed processes with applications to ARCH.
*Annals of Statistics***26**.

[18] Davis, R.A. and Resnick, S.I. (1985) More limit theory for the sample correlation function of moving averages. *Stoch. Proc. Appl.* **20**, 257–279.

[19] Davis, R.A. and Resnick, S.I. (1986) Limit theory for the sample covariance and correlation functions of moving averages. *Ann. Statist. 14*, 533-558.

[20] Diebolt, J. and Guégan, D. (1990) Probabilistic properties of the general nonlinear Markovian process of order one and applications to time series modelling. *Rapp. Tech.* **125**, L.S.T.A., Paris 6.

[21] Duan, J.-C. (1996) Augmented GARCH(p,q) process and its diffusion limit. *Journ. of Econometrics*

[22] Duffie, D. (1992) *Dynamic Asset Pricing Theory.* Princeton University Press, Princeton.

[23] Eberlein, E. and Keller, U. (1995) Hyperbolic distributions in finance.
Bernoulli **1**, 281-299.

[24] Embrechts, P. and Goldie, C.M. (1994) Perpetuities and random equations. In: Mandl, P., Huskova, M. (Eds.) Asymptotic Statistics. *Proceedings of the 5th Prague Symposium.* September 4–9, 75–86, Physica-Verlag, Heidelberg.

[25] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997)
Modelling Extremal Events for Insurance and Finance.
Springer, Heidelberg.

[26] Embrechts, P., Resnick, S.I. and Samorodnitsky, G. (1997)
Extreme value theory as a risk management tool. *ETH Zürich.*

[27] Emmer, S., Klüppelberg, C., Trüstedt, M. (1998) VaR - ein Mass fr das extreme Risiko. *Solutions 1*, 53-63.

[28] Engle, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation. *Econometrica* **50**, 987–1007.

[29] Fama, E.F. (1965) The behavior of stock market prices. *Journ. Business* **38**, 34–105.

[30] Feigin, P.D. and Tweedie, R.L. (1985) Random coefficient autoregressive processes: A Markov chain analysis of stationarity and finiteness of moments. *Journ. of Time Series Anal.***6**, 1–14.

[31] Fisher, R.A. and Tippett, L.H.C. (1928)
Limiting forms of the frequency distribution of the largest or smallest member of a sample.
*Proc. Cambridge Phil. Soc.***24**, 180–190.

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

[33] Gumbel, E.J. (1958) *Statistics of Extremes.* Columbia University Press, New York.

[34] Haan, L. de, Resnick, S.I., Rootzén, H. and Vries, C.G. de (1989) Extremal behavior of solutions to a stochastic difference equation with applications to ARCH processes. *Stoch. Proc. Appl.* **32**, 213–224.

[35] Hooghiemstra, G. and Meester L.E. (1995) Computing the extremal index of special Markov chains and queues. *Stoch. Proc. Appl.* **65**, 171–185.

[36] Hsing, T. (1984) Point processes associated with extreme value theory. *Dissertation, Technical Report 83, Center for Stochastic Processes, Univ. North Carolina.*

[37] Hsing, T., Hüsler, J. and Leadbetter, M.R. (1988) On the exceedance of point process for a stationary sequence. *Prob. Theory and Rel. Fields* **78**, 97–112.

[38] Itô, K. and McKean, H.P.. (1974) *Diffusion Processes and their Sample Paths.* Springer, Berlin.

[39] Joergensen, B. (1982) *Statistical Properties of the Generalised Inverse Gaussian Diffusion.* Springer, Berlin.

[40] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995) *Continuous Univariate Distributions, Vol.2.* Wiley, New York.

[41] Kallenberg, O. (1983) *Random measures.* Akademie Verlag, Berlin.

[42] Karatzas, I. and Shreve, S.E. (1988)
Brownian Motion and Stochastic Calculus. Springer, New York.

[43] Karlin, S. and Taylor, H.M. (1981)
A Second Course in Stochastic Processes.
Academic Press, New York.

[44] Kesten, H. (1973) Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131**, 207–248.

[45] Kloeden, P.E. and Platen, E. (1992)
Numerical Solutions of Stochastic Differential Equations.
Springer, Berlin.

[46] Klüppelberg, C. and Mikosch, T. (1996) Self–normalised and randomly centred spectral estimates. In: Heyde, C.C., Prokhorov, Yu.V., Pyke, R. and Rachev, S.T. (Eds.) *Proceedings of the Athens International Conference on Applied Probability and Time Series,* vol. 2: Time Series, pp. 259–271. Springer, Berlin.

[47] Konakov, V.D. and Piterbarg, V.I. (1982) Convergence rate of maximum deviation distributions of Gaussian processes and empirical densities.
*Theory of Prob. and Appl.***27**, 759–779.

[48] Kratz, M.F. and Rootzén, H. (1997) On the rate of convergence for extremes of mean square differentiable stationary normal processes.
*Journ. Appl. Probab.***34**, 908–923.

[49] Küchler, U. Neumann, K. Sorensen, M. and Streller, A. (1994)
Stock returns and hyperbolic distributions. Research Report, Sonderforschungsbereich 373, Humboldt University, Berlin.

[50] Lamberton, D. and Lapeyre, B. (1991) *Introduction au calcul stochastique appliqué á la finance.* Ellipses–Edition Marketing, SMAI, Paris.

[51] Leadbetter, M.R. (1983) Extremes and local dependence of stationary sequences. *Z. Wahrsch.theorie verw. Gebiete* **65**, 291–306.

[52] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983)
Extremes and Related Properties of Random Sequences and Processes.
Springer, Berlin.

[53] Leadbetter, M.R. and Rootzén, H. (1982) Extreme value theory for continuous stationary processes. *Z. Wahrsch. verw. Gebiete***60**, 1–20.

[54] Leadbetter, M.R. and Rootzén, H. (1988) Extremal theory for stochastic processes. *Ann. Prob.***16**, 431–478.

[55] Lindvall, T. (1992) *Lectures on the Coupling Method.* Wiley, New York.

[56] Loynes, R.M. (1965) Extreme values in uniformly mixing stationary stochastic processes. *Ann. Math. Statist.* **36**, 993–999.

[57] Maercker, G. (1997) *Statistical Inference in Conditional Heteroskedastic Autoregressive Models.* Shaker, Aachen. (Dissertation, University of Braunschweig)

[58] Mandelbrot, B. (1963) The variation of certain speculative prices. *Journ. Business* **36**, 394–419.

[59] Mandl, P. (1968) *Analytical Treatment of One-Dimensional Markov Processes.* Springer, New York.

[60] McNeil, A., J. (1997) Calculating quantile risk measures for financial return series using extreme value theory. Departement of Mathematics, ETH Zürich, *Preprint.*

[61] Merton, R.C. (1994) *Continuous–time Finance.* Revised Ed. Blackwell, Cambridge, MA.

[62] Meyn, S.P. and Tweedie, R.L. (1993) *Markov Chains and Stochastic Stability.* Springer, London.

[63] Mittnik, S. and Rachev, S.T. (1997) *Financial Modeling and Option Pricing with Alternative Stable Models.* Series in Financial Economics and Quantitative Analysis. Wiley, New York. To appear.

[64] Newell, G.F. (1962) Asymptotic extreme value distributions for one-dimensional diffusion processes. *J. Math. Mech.* **11**, 481-496.

[65] Newell, G.F. (1964) Asymptotic extremes for m-dependent random variables. *Ann. Math. Statist.* **35**, 1322-1325.

[66] O'Brien, G.L. (1974) Limit theorems for the maximum term of a stationary process. *Ann. Probab.* **2**, 540–545.

[67] O'Brien, G. L. (1987) Extreme values for stationary and Markov sequences. *Ann. Prob.* **15**, 281–291.

[68] Perfekt, R. (1994) Extremal behavior of stationary Markov chains with applications. *Ann. Appl. Prob.* **4**, 529–548.

[69] Pickands III, J. (1969) Upcrossing probabilities for stationary Gaussian processes. *Transactions American Math. Soc.* **145**, 51-73.

[70] Priestley, M.B. (1988) *Non-linear and Non-stationary Time Series.* Academic Press, New York.

[71] Protter, P. (1992) *Stochastic Integration and Differential Equations.* Springer, Berlin.

[72] Resnick, S.I. (1987)
Extreme Values, Regular Variation and Point Processes.
Springer, New York.

[73] Revuz, D. and Yor, M. (1991) *Continuous Martingales and Brownian Motion.* Springer, Berlin.

[74] Rice, S.O. (1939) The distribution of the maxima of a random curve. *Amer. J. Math.* **61**, 409–416.

[75] Sigma (1995) Non-proportional reinsurance of losses due to natural disasters in 1995: Prices down despite insufficient cover. *Sigma publication No 6*, Swiss Re, Zürich.

[76] Sigma (1996) Natural catastrophes and major losses in 1995: decrease compound to previous year but continually high level of losses since 1989. *Sigma publication No 2*, Swiss Re, Zürich.

[77] Sigma (1997) World insurance in 1995: premium volume exceeds USD 2000 billion for the first time. *Sigma publication No 4*, Swiss Re, Zürich.

[78] Sigma (1998) Exceptionally few high losses from natural and man-made catastrophes in 1997. *Sigma publication No 3*, Swiss Re, Zürich.

[79] Smith, R.L. (1988) A counterexample concerning the extremal index. *Adv. Appl. Probab.* **20**, 681–683.

[80] Rogers, L.C.G. and Williams, D. (1987)
Diffusions, Markov Processes, and Martingales, Vol. 2.
Wiley, New York.

[81] Rootzén, H. (1988) Maxima and exceedances of stationary Markov chains. *Adv. Appl. Prob.* **20**, 371–390.

[82] Rydberg, T. (1996) Generalized hyperbolic diffusions with applications towards finance. Research Report No. 342. Department of Theoretical Statistics, University of Aarhus.

[83] Shephard, N. (1996) Statistical aspects of ARCH and stochastic volatility. In: Cox, D.R., Hinkley, D.V. and Barndorff-Nielsen, O.E. (Eds.), Likelihood, Time Series with Econometric and other Applications. Chapman and Hall, London.

[84] Sørensen, M. (1996)
Estimating functions for discretely observed diffusions: a review.
Research Report No 348, University of Aarhus.

[85] Taylor, S. (1995) *Modelling financial time series.* Wiley, Chichester.

[86] Tong, H. (1990) *Non-linear Time Series – A Dynamical System Approach* Oxford University Press, Oxford.

[87] Tweedie, R.L. (1976) Criteria for classifying general Markov chains. *Adv. Appl. Prob.* **8**, 737–771.

[88] Tweedie, R.L. (1983) Criteria for rates of convergence of Markov chains, with application to queuing and storage theory. In J.F.C. Kingman and G.E.H. Reuter, editors, *Probability, Statistics and Analysis.* London Mathematics Society Lecture Note Series, Cambridge University Press, Cambridge.

[89] Tweedie, R.L. (1983) The existence of moments for stationary Markov chains. *J. Appl. Prob.* **20**, 191–196.

[90] Vasicek, O.A. (1977)
An equilibrium characterization of the term structure. J. Fin. Econ. **5**, 177-188.

[91] Vervaat, W. (1979) On a stochastic difference equation and a representation of nonnegative infinitely divisible random variables. *Adv. Appl. Prob.* **11**, 750–783.

[92] Weiss, A.A. (1984) ARMA models with ARCH errors. *Journ. of Time Series Anal.* **3**, 129–143.

List of Figures

List of Tables

List of Symbols

AR(p)	autoregressive process of order p	7
ARCH(p)	autoregressive conditionally heteroskedastic process of order p	7
ARMA(p,q)	autoregressive moving-average process of order (p, q)	62
a.a.	amost all	113
a.s.	almost sure, almost surely, with probability 1	35
const	some constant	36
$D(u_n)$	mixing condition	3
$\Delta(u_n)$	mixing condition	??
ε_x	Dirac measure	58
$E(X)$	expectation of the random variable X	16
F	distribution function and distribution of a random variable	2
$F^{\leftarrow}(x)$	generalised inverse of F in x: $\inf\{s \in \mathbb{R} \mid F(s) \geq x\}$	79
$\overline{F}$	tail of the distribution function F: $\overline{F} = 1 - F$	24
f^{-1}	inverse of f	68
f', f''	first and second derivatives of f	69
Γ	gamma function: $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$	80
$\Gamma(a, b)$	gamma distribution with parameters a and b: $p(x) = b^a\, \Gamma(a)^{-1}\, x^{a-1}\, e^{-bx},\ x \geq 0$	80
GARCH(p,q)	generalised autoregressive conditional heteroske-	7

	dastic process of order (p, q)	
H	stationary distribution function of a diffusion process	65
h	stationary density of a diffusion process	63
i.i.d.	independent, identically distributed	2
inf	infimum	16
K_λ	modified Bessel function of the third kind: $K_\lambda(x) = \frac{1}{2}\int_0^\infty t^{\lambda-1}\exp(-\frac{1}{2}x(t+t^{-1}))dt$	89
Λ	Gumbel distribution function: $\Lambda(x) = \exp(-e^{-x}),\ x \in \mathbb{R}$	2
lim	limes	2
lim inf	limes inferior	29
lim sup	limes superior	30
lhs	left-hand side	65
$\ln x$	logarithm with basis e	2
$l(x)$	slowly varying function	34
$L_t^X(z)$	local time of the process (X_t) in state z up to time t	69
$l_t(z)$	$L_t^W(z)$	69
MDA	maximum domain of attraction	65
$MDA(Q)$	MDA of the extreme value distribution Q	65
$M_{k,n}$	maximum of $X_k, ..., X_n$	44
M_n	$M_{1,n}$	3
$M_{(s,t]}$	maximum of the continuous time process (X_u) in the interval $(s,t]$	71
M_t	$M_{(0,t]}$	10
m_X	speed measure of the diffusion process (X_t)	63
$\lvert m_X \rvert$	total speed measure of the diffusion process (X_t) with state space (l,r): $\int_{(l,r)} m(dz)$	63
$\mathbb{N}$	set of positive integers	3
$N, N(t), N_t$	point process	3
$N_{\varepsilon,u}^X$	point process of ε-upcrossings	5

	above the threshold u of the process (X_t)	
$N(0,1)$	standard normal distribution	18
$o(1)$	$a(x) = o(b(x))$ as $x \to x_0$ means that $\lim_{x\to x_0} a(x)/b(x) = 0$	16
$O(1)$	$a(x) = O(b(x))$ as $x \to x_0$ means that $\limsup_{x\to x_0} \lvert a(x)/b(x)\rvert < \infty$	21
ω	$\omega \in \Omega$ random outcome	40
OU	Ornstein-Uhlenbeck	70
$\varphi(\cdot\lvert\mu,\sigma^2)$	density of the normal distribution with mean μ and variance σ^2	18
Φ	standard normal distribution or distribution function	87
Φ_α	Frechet distribution function: $\Phi_\alpha(x) = \exp(-x^{-\alpha}),\ \alpha > 0$	2
Ψ_α	Weibull distribution function: $\Psi_\alpha(x) = \exp(-(-x)^\alpha),\ \alpha > 0$	2.
$\mathbb{R}$	real line	2
rhs	right-hand side	24
s_X	scale function of the diffusion process (X_t)	63
τ_B^X	$\inf\{t \geq 0 \mid X_t \in B\}$	21
τ_t	random time change at time t	69
u, u_n	threshold in extreme value theory	3
$\mathrm{var}(X)$	variance of the random variable X	78
W	standard Brownian motion	6
w.l.o.g.	without loss of generality	17
x_F	right endpoint of the distribution function F	109
$\underline{X}$	random vector $(X_1, ..., X_m)$ for some $m > 0$	58
1_A	indicator function of the set (event) A	2
$\sim$	$a(x) \sim b(x)$ as $x \to x_0$ means that $\lim_{x\to x_0} a(x)/b(x) = 1$	9
$\approx$	$a(x) \approx b(x)$ as $x \to x_0$ means that $a(x)$ is approximately (roughly) of the same	39

	order as $b(x)$ as $x \to x_0$. It is only used in a heuristic sense.	
$\lvert \cdot \rvert$	absolute value	5
B^c	complement of the set B	21
$\stackrel{a.s.}{\to}$	a.s. convergence	42
$\stackrel{d}{\to}$	convergence in distribution	3
$\stackrel{P}{\to}$	convergence in probability	42
$\stackrel{w}{\to}$	weak convergence	3
$\stackrel{d}{=}$	same distribution	44

Index

E

F

U

V

W

Y

Curriculum Vitae

of Milan Borkovec, Center of Mathematical Sciences, Munich University of Technology, D-80290 Munich.

Born the 02/12/70 in Basle, Switzerland.

Thesis advisor: Prof. C. Klüppelberg.

Fields of research: Extreme Value Theory for Stochstic Processes, Mathematical Finance.

Education

1997 – 1999	Teaching assistant and Ph.D. student at the University of Technology in Munich
1995 – 1997	Teaching assistant and Ph.D. student at the Johannes Gutenberg University in Mainz
1990 – 1995	Study of mathematics at ETH Zürich, specializing in probability theory. Subject of Diploma thesis: "Extremales Verhalten von stark oszillierenden Zeitreihen". Walter Saxer-Versicherungs-Hochschulpreis for diploma thesis
1986 – 1989	Grammar school in Muttenz BL, Matura Typus C
1982 – 1986	Secondary school in Muttenz BL
1977 – 1982	Primary school in Muttenz BL

Teaching Experience

04/01/97 – 03/31/2000 :	Teaching assistant, Munich University of Technology.
04/01/95 – 03/31/97 :	Teaching assistant, Johannes Gutenberg-University Mainz.

Lectures

- Oberseminar, University of Mainz, Germany, February 1999
- Data Analysis and Modelling Seminar, University of Freiburg, Germany, December 1998
- Financial and Insurance Mathematics Seminar, ETH Zurich, Switzerland, November 1998
- 22nd European Meeting of Statisticians 7th Vilnius Conference on Probability Theory and Mathematical Statistics, Vilnius, Lithuania, August 1998
- Münchener Stochastik-Tage 1998, Munich, Germany, March 1998
- Satellite Meeting to the 4th World Congress of the Bernoulli Society on "Stable Processes and other Heavy-Tailed Models for Highly Volatile Phenomena", Wroclaw, Poland, August 1996

Conferences attended

- 1st International Conference on Risk Management in Finance, Munich, Germany, October 1998
- Workshop "Extremes - Risk and Safety", Gothenburg, Sweden, August 1998
- Workshop on "The Interplay Between Finance, Insurance and Statistics", Cortona, Italy, June 1998
- ASTIN Congress, Winterthur, Switzerland, October 1997
- World Congress of the Bernoulli Society, Vienna, Austria, August 1996
- CIME Session on "Financial Mathematics", Bressanone, Italy, July 1996

Research Visits

- Rijksuniversiteit Groningen, Groningen, The Netherlands, November 1997 (1 week)
- University of Sussex, Brighton, England, September 1997 (1 week)

Publications

- Borkovec, M. (1998) Extremal behavior of the autoregressive process with ARCH(1) errors, Preprint.

- Borkovec, M., Klüppelberg, C. (1998) The tail of the stationary distribution of an autoregressive process with ARCH(1) errors, Preprint.

- Borkovec, M., Klüppelberg, C. (1998) Extremal behaviour of diffusion models in finance, Extremes 1, 47-80.

Referees

- Prof. C. Klüppelberg, Munich University of Technology, D-80290 Munich, Germany, cklu@mathematik.tu-muenchen.de

- Prof. H. Rootzén, Chalmers, S-412 96 Gothenburg, Sweden, rootzen@math.chalmers.se