Extremes of Lévy Driven Moving Average Processes with Applications in Finance

Vicky Maria Fasen

Center for Mathematical Sciences

Munich University of Technology

D-85747 Garching

2004

Zentrum Mathematik
Lehrstuhl für Mathematische Statistik
der Technischen Universität München

Extremes of Lévy Driven Moving Average Processes with Applications in Finance

Vicky Maria Fasen

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor der Naturwissenschaften (Dr. rer. nat)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Herbert Spohn
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Claudia Klüppelberg
	2.	Prof. Dr. Gennady Samorodnitsky Cornell University, Ithaca, USA

Die Dissertation wurde am 07. 10. 2004 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 03. 12. 2004 angenommen.

Zusammenfassung

Empirische Volatilität ist nicht konstant in der Zeit und weist Tails auf, die schwerer sind als normalverteilt. Des Weiteren sieht man oft Sprünge und Clusterverhalten. In dieser Arbeit wird das Extremwertverhalten verschiedener Volatilitätsmodelle untersucht: Subexponentielle Lévy getriebene MA Prozesse im Anziehungsbereich der Gumbel-Verteilung, regulär variierende gemischte MA Prozesse, Ornstein-Uhlenbeck Prozesse mit exponentiellem Tail und COGARCH Prozesse.

Der Schwerpunkt dieser Arbeit liegt in der Untersuchung von subexponentiellen Lévy getriebenen MA Prozessen $Y(t) = \int_{-\infty}^{\infty} f(t-s)\, dL(s)$ für $t \in \mathbb{R}$, wobei f eine deterministische Funktion und L ein Lévy Prozess ist. In Kapitel 1 beschäftigen wir uns mit dem extremalen Verhalten im Maximum-Anziehungsbereich der Gumbel-Verteilung und in Kapitel 2 im Anziehungsbereich der Fréchet-Verteilung. Das Verhalten in den beiden Anziehungsbereichen ist sehr unterschiedlich. Für beide Klassen werden hinreichende Bedingungen an die Kernfunktion f gegeben, so dass eine stationäre Version des MA Prozesses Y existiert. Wir berechnen das Tailverhalten der stationären Verteilung. Es stellt sich heraus, dass die stationäre Verteilung auch wieder subexponentiell ist und sogar im gleichen Anziehungsbereich wie der treibende Lévy Prozess L liegt. Somit modellieren sie schwere Tails und Volatilitätssprünge. Die Analyse des extremalen Verhaltens basiert auf einem zeit-diskreten Gitter, das bei den Sprungzeitpunkten des Lévy Prozesses L und den Extrema der Kernfunktion f geeignet gewählt wird. Nachdem die diskrete Folge mit Marken versehen worden ist, wird das Grenzwertverhalten des daraus entstehenden Punktprozesses berechnet. Dieser liefert vollständige Information über das extremale Verhalten. Unter Anderem ergibt sich daraus die Konvergenz der normalisierten, wachsenden Maxima. Beide Modelle weisen Volatilitätscluster auf. Regulär variierende MA Prozesse verweilen lange über einer hohen Schwelle, im Gegensatz dazu haben MA Prozesse im Anziehungsbereich der Gumbel-Verteilung nur in einzelnen Punkten Exzesse.

Desweiteren betrachten wir in Kapitel 3 das Extremwertverhalten von Ornstein-Uhlenbeck Prozessen mit exponentiell fallendem Tail. Es ist ähnlich zum subexponentiellem MA Prozess im Anziehungsbereich der Gumbel-Verteilung. Sie haben schwere Tails aber keine Volatilitätscluster. Als letzte Klasse an Volatilitätsmodellen wird der COGARCH Prozess untersucht. Getrieben von einen compound Poisson Prozess weist er auch regulär variierende Tails, Volatilitätssprünge und Cluster in den Extrema auf.

Abstract

Empirical volatility changes in time and exhibits tails, which are heavier than those of normal distributions. Moreover, empirical volatility has - sometimes quite substantial - upwards jumps and clusters on high levels. We investigate classical and non-classical stochastic volatility models with respect to their extreme behavior: subexponential Lévy driven MA processes in the maximum domain of attraction of the Gumbel distribution, regularly varying mixed MA processes, Ornstein-Uhlenbeck processes with exponentially decreasing tails and COGARCH processes.

The basic volatility models of this thesis are subexponential Lévy driven MA processes $Y(t) = \int_{-\infty}^{\infty} f(t-s)\, dL(s)$ for $t \in \mathbb{R}$ where f is a deterministic function and L is a Lévy process. In Chapter 1 we study the extremal behavior of subexponential MA processes in the maximum domain of attraction of the Gumbel distribution and in Chapter 2 of the Fréchet distribution. The behavior is quite different in these different regimes. For both classes we give sufficient conditions for the kernel function f, such that a stationary version of the MA process Y exists, which preserves the infinitely divisibility of L. We calculate the tail behavior of the stationary distribution, which is again subexponential and in the same maximum domain of attraction as the driving Lévy process L. Hence they capture heavy tails and volatility jumps. Our investigation on the extremal behavior of Y is based on a discrete-time skeleton of Y chosen to incorporate those times, where large jumps of the Lévy process L and extremes of the kernel function f occur. Adding marks to this discrete-time skeleton, we obtain, by the weak limit of marked point processes, complete information about the extremal behavior. A complementary result guarantees the convergence of running maxima. Both models have volatility clusters. Regularly varying MA processes have long high level excursion in contrast to subexponential MA processes in the maximum domain of attraction of the Gumbel distribution, where they collapse into single points.

Furthermore, in Chapter 3 we investigate the extremal behavior of Ornstein-Uhlenbeck processes with exponential tails. This is similar to subexponential Ornstein-Uhlenbeck processes in the maximum domain of attraction of the Gumbel distribution. They are heavy tailed, but do not exhibit volatility clusters. As the last class of continuous-time volatility models, we study a continuous-time GARCH(1,1) model. Driven by a compound Poisson process it exhibits regularly varying tails, volatility upwards jumps and clusters on high levels.

Acknowledgement

It is a particular pleasure for me to express my sincere thanks to my advisor Prof. Dr. Claudia Klüppelberg for having confidence in me and for her infinite help and support. I feel also very grateful that by her assistance, I came into contact with very distinguished scientists and that she gave me the opportunity to travel extensively.

It is a pleasure for me to thank Prof. Dr. Gennady Samorodnitsky for his encouragement especially at the Autumn School on Risk Management in September 2003. I have extremely benefited from discussions with him.

I acknowledge with thanks the hospitality of Prof. Dr. Holger Rootzén and his colleagues at the Department of Mathematical Statistics, Chalmers University of Technology in fall 2003. Discussions with him have been very fruitful and stimulating.

I would like to thank my colleagues at the Munich University of Technology for their support and for the enjoyable company during the last years.

Last but not least I thank my family for their support in all situations and all my friends for their encouragement especially during the last months.

Financial support by the Deutsche Forschungsgemeinschaft through the graduate program "Angewandte Algorithmische Mathematik" at the Munich University of Technology is gratefully acknowledged.

Contents

Introduction

Extreme value theory

Extreme value theory is a statistical discipline that develops techniques and models for describing rare events of extremes. The subject has a rich mathematical theory and also a long tradition of applications in a variety of areas. The interest in this theory has increased continuously during the last decades, partly due to the fact that catastrophic events are often followed by large pecuniary claims. Hence, extreme value models and techniques are widely applied for designing protection systems against the effects of extreme events in such diverse fields as financial risk management, insurance mathematics, engineering sciences, environmental engineering and environmental statistics. Many applications of extreme value theory are described in the monograph of Gumbel [73].

Classical extreme value theory is the asymptotic theory for maxima

$$M(n) = \max\{Y_1, \ldots, Y_n\}$$

of independent identically distributed (i. i. d.) random variables $\{Y_k\}_{k\in\mathbb{N}}$ with distribution function F. The central limit theory obtains an asymptotic normal distribution for the sum of many i. i. d. random variables with finite variance, whatever their common original distribution function is. The distribution function has not to be known precisely to apply the asymptotic theory. A similar situation exists in extreme value theory. The extremal types theorem, discovered by Fisher and Tippett [66], and discussed afterwards by Gnedenko [67], exhibits possible limit forms of the distribution of $M(n)$ under linear normalization. This means that, provided $a_n > 0$ and $b_n \in \mathbb{R}$ are sequences such that

$$\lim_{n\to\infty} \mathbb{P}(a_n^{-1}(M(n) - b_n) \le x) = \lim_{n\to\infty} F^n(a_n x + b_n) = G(x)\,, \quad x \in \mathbb{R}\,, \qquad (0.0.1)$$

for some non-degenerate distribution G (we say F is in the *maximum domain of attraction* of G and write $F \in \mathrm{MDA}(G)$), then there are constants $a > 0$, $b \in \mathbb{R}$ such that $x \mapsto G(ax+b)$ is one of the following three *extreme value distributions*:

- Fréchet: $\Phi_\alpha(x) = \begin{cases} 0, & x \le 0, \\ \exp\left(-x^{-\alpha}\right), & x > 0, \end{cases}$ for $\alpha > 0$.
- Gumbel: $\Lambda(x) = \exp\left(-\mathrm{e}^{-x}\right), \quad x \in \mathbb{R}\,.$
- Weibull: $\Psi_\alpha(x) = \begin{cases} \exp\left(-\left(-x\right)^{\alpha}\right), & x \le 0, \\ 1, & x > 0, \end{cases}$ for $\alpha > 0\,.$

Of great importance for the development of extreme value theory was the work of de Haan [56], who applied rigorously methods from regular variation as the appropriate analytic tool. In this framework the class of *regularly varying* distributions $\mathcal{R}_{-\alpha}$ is famous, where $F \in \mathcal{R}_{-\alpha}$ satisfies $\lim_{t\to\infty} \overline{F}(tx)/\overline{F}(t) = x^{-\alpha}$ for every $x > 0$ and some $\alpha \in \mathbb{R}$, where $\overline{F}(x) = 1 - F(x)$. The question arises now for criteria on a distribution tail to belong to a maximum domain of attraction and to characterize the norming constants a_n, b_n. This is based on the intuition that the far out tail of F determines completely the maximum domain of attraction of F. In the maximum domain of attraction of the Fréchet distribution are only distributions with regularly varying tails; examples are the stable, Cauchy, Pareto and Burr distribution. All essential results of regularly varying functions can be found in Bingham et al. [29]. Distributions in $\mathrm{MDA}(\Lambda)$ have lighter tails than regularly varying distributions and distributions in $\mathrm{MDA}(\Psi_\alpha)$ have bounded support to the right. For details we refer to the excellent monographs about extreme value theory of Embrechts et al. [60], Leadbetter et al. [95] and Resnick [125].

Extreme value theory is not only concerned with maximum domain of attractions and the calculation of normalizing constants a_n, b_n, but also with the more sophisticated behavior of exceedances. An *exceedance* of Y_k at level u_n means that $Y_k > u_n$. Instead of considering the times at which high-threshold exceedances occur and the excess values over the threshold as two separate processes, they are combined into one point process. Since exceedances occur randomly in time, a point process counts the number of events, where $k/n \in [s,t)$ and $Y_k > u_n = a_n x + b_n$ for any $0 \le s < t$, $x \in \mathbb{R}$. The time is normalized, since u_n is increasing and so the number of exceedances in a fixed time interval is decreasing. A sequence of such point processes converges weakly to a Poisson random measure. Hence we obtain additional information, for example, about the location of maxima, the behavior of upper order statistics, joint convergence of maxima and minima and records. Point processes of exceedances are essential to understand the structure of the extremal behavior of any sequence of random variables. The theory of point processes has been treated in detail in Kallenberg [80] and Resnick [125].

The results of extreme value theory for i. i. d. sequences as extremal types theorem and convergence of point processes were generalized for stationary sequences under weak dependence restrictions, well-known as $D(u_n)$ and $D'(u_n)$ conditions by Watson [143], Berman [23], Loynes [101], Leadbetter [93] and others. The $D(u_n)$ condition describes the degree of independence of maxima on separated intervals, the $D'(u_n)$ condition avoids clusters in extremes. The qualitative behavior of extremes of the stationary sequence and the associated i. i. d. sequence are the same.

In this thesis we concentrate on the extremes of continuous-time processes. The pioneering work started for Gaussian processes with Rice [127,128,129]. The present view of extremal behavior for continuous-time processes began with Cramér [48,49] and Pickands [119,118]. A wealth of other papers followed, which complemented the results of Cramér and Pickands. A general theory on the extremal behavior of stationary continuous-time processes $\{Y(t)\}_{t\in\mathbb{R}}$, which have almost surely continuous sample paths and continuous one-dimensional distribution is given in the seminal monograph of Leadbetter et al. [95], see also Leadbetter and Rootzén [96], and further references therein. Their approach is via a discrete-time skeleton, so that results of stationary sequences can be applied. To this end for a fixed $h > 0$ the sequence of submaxima

$$M_k = \sup_{(k-1)h\leq t<kh} Y(t) \quad \text{for } k \in \mathbb{N} \tag{0.0.2}$$

is defined. If the $D(u_n)$ and $D'(u_n)$ conditions are valid for the discrete-time sequence $\{M_k\}_{k\in\mathbb{N}}$, then the extremal behavior of the continuous-time process can be described by this. Among other results an extremal types theorem holds. However, it is clearly the tail of the distribution of M_1 rather than $Y(1)$, which determines the limiting type. The conditions $D(u_n)$ and $D'(u_n)$ are reformulated onto conditions on the continuous-time process Y, known as $C(u_n)$ and $C'(u_n)$ condition.

The best understood continuous-time processes are Gaussian processes. For stationary normal sequences and processes the D, D' and C, C' conditions reduce to conditions on the covariance function, known as Berman's condition, since the distribution of Gaussian processes are completely described by their mean value and covariance function. The early monographs on extreme value theory of Leadbetter et al. [95] and Berman [26] contain all basic results on this topic, and it is this source from which all specific results are derived. Berman has an alternative approach as the C, C' condition by considering excursions of upcrossings above high levels; such events are called sojourns. The extreme behavior of stationary diffusions driven by Brownian motion has been investigated by Newell [114], Berman [24, 26], Mandl [105], Davis [51] and Borkovec and Klüppelberg [32]. Such models, though driven by Brownian motion can fall outside the family of Gaussian processes. They have played an important role in physics, but also as financial models. Typical applications are term structure models such as the Vašiček model, the Cox-Ingersoll-Ross model and the generalized hyperbolic diffusion.

A substantial contribution to the extremal behavior of continuous-time processes came from Albin [2,3]. He developed a theory, which is independent of the maximum domain of attraction. These results he used for describing the extremal behavior of totally skewed stable processes [4, 5, 6], differentiable processes [7] and Ornstein-Uhlenbeck processes [1].

Stochastic volatility models

Continuous-time models play a crucial role in modern finance. They provide the basis of option pricing, asset allocation and term structure theory. The underlying process of asset prices, exchange rates, indices, or interest rates is often irregularly spaced, in particular, in the context of high frequently data. Consequently, one often works with continuous-time models. Moreover, such data have often leptokurtic marginal distributions; i. e. their histograms show a very pronounced peak around zero and are heavy tailed. Further stylized features exhibited by financial data are volatility clusters on high levels, large fluctuations and long-range dependence. The meaning of *heavy tailed* is that the marginal distribution is heavier tailed than a normal distribution. Evidence of heavy tails in financial asset returns distributions are plentiful since the seminal work of Mandelbrot [104] on cotton prices. Clustering in the data causes periods of high and low activity which imply periods of high and low volatility. Hence, volatility processes are typical examples with these stylized features. Such observations resulted in an enormous effort to develop empirically reasonable models, which can be integrated in financial theory. For an introduction and overview of stochastic volatility models we refer to Barndorff-Nielsen and Shephard [17] and Shephard [139].

A natural class of heavy tailed distributions is given by the class of *subexponential distributions* denoted by $\mathcal{S}$. A random variable Y is subexponential, if

$$\mathbb{P}(Y_1 + \ldots + Y_n > x) \sim \mathbb{P}(\max\{Y_1, \ldots, Y_n\} > x) \quad \text{for } x \to \infty,$$

where $\{Y_n\}_{n\in\mathbb{N}}$ is an i. i. d. sequence with the same distribution as Y. This clearly indicates the strong influence of the largest value on the total sum. A few large values are likely to determine the long term behavior of the system. The name arises from their property, that the tail of Y decreases slower than any exponential function and hence no exponential moment exists. The class of subexponential distributions includes regularly varying distributions, the semi-heavy tailed Weibull distribution with shape parameter less than one and the log-normal distribution. Subexponential distributions have gained popularity in various contexts of applied probability such as telecommunication models, branching theory, queueing models, insurance mathematics and financial risk management. Subexponential distributions can belong to two different maximum domain of attractions. All regularly varying distributions are subexponential and belong to $\text{MDA}(\Phi_\alpha)$. Other subexponential distributions like the lognormal and semi-heavy tailed Weibull distribution belong to $\text{MDA}(\Lambda)$ and have *rapidly varying* tails, i. e. $\lim_{t\to\infty} \overline{F}(tx)/\overline{F}(t) = 0$ for every $x > 1$ and the limit is ∞ for $0 < x < 1$. Distributions in $\text{MDA}(\Psi_\alpha)$ have bounded support to the right and hence can not be subexponential. A survey of the class of subexponential distributions is provided by Goldie and Klüppelberg [70], see also Embrechts et al. [60].

Certain time series models are very popular in financial econometrics, where they are designed to capture some of the distinctive features mentioned above. An approach is to derive from discrete-time models continuous-time models that arise naturally and intuitively to reflect the stylized facts of financial processes. Classical time series theory is mostly concerned with MA *(moving average)* processes

$$Y_n = \sum_{k=-\infty}^{\infty} c_k Z_{n-k} \quad \text{for } n \in \mathbb{N}, \tag{0.0.3}$$

where $\{Z_k\}_{k\in\mathbb{Z}}$ is an i. i. d. sequence, $\{c_k\}_{k\in\mathbb{Z}}$ is a sequence of constants, and the infinite sum is assumed to converge with probability one. This class includes ARMA processes used most frequently in applications in engineering, physics, chemistry and metrology; see Brockwell and Davis [38]. The right tail and the extremal behavior of a stationary MA process depend on the weights, the right and the left tail of the marginal distribution function of the noise variables. MA processes with subexponential noise variables Z exhibit heavy tails as well as clusters in extremes depending on the properties of the filter. The maximum domain of attraction of the subexponential noise variables Z plays a crucial role for the extremal behavior of Y. The quantitative behavior of extremes of MA processes in different maximum domain of attractions is similar, but the qualitative behavior is completely different. They are very well studied in Davis and Resnick [54, 55] and Rootzén [130]. As a result of clusters classical extreme value theory, as mentioned in the introduction, is not applicable.

The aim of this thesis is to provide and investigate a wide class of continuous-time models which reflects the extremal properties of empirical volatilities, namely heavy tails and high level volatility clustering. We are concerned with a continuous-time version of the MA process as given in (0.0.3) with respect to its extremal behavior. The continuous-time analogue to an i. i. d. sequence in discrete-time is constituted by the increments of a Lévy process $L = \{L(t)\}_{t\in\mathbb{R}}$, because of its independent and stationary increments. Moreover, we assume that $L(0) = 0$ and that L is càdlàg. Then similarly to the definition of the discrete-time MA process (0.0.3) we define the continuous-time MA process $\{Y(t)\}_{t\in\mathbb{R}}$ as

$$Y(t) = \int_{-\infty}^{\infty} f(t-s)\, dL(s) \quad \text{for } t \in \mathbb{R}\,, \tag{0.0.4}$$

where the *kernel function* $f : \mathbb{R} \to \mathbb{R}$ is measurable. A typical example is the *Ornstein-Uhlenbeck process* with $f(t) = \mathrm{e}^{-\lambda t}\,\mathbf{1}_{[0,\infty)}(t)$ for some $\lambda > 0$, propagated as stochastic volatility model by Barndoff-Nielsen and Shephard [15], and CARMA processes studied by Brockwell [35, 36]. The class of continuous-time MA processes is very flexible to design models to fit marginal features of the distribution of data as well as to deal separately with the observed dependence structure in data. Under certain conditions on the kernel function f the resulting process (0.0.4) can even exhibit long range dependence, another stylized fact often observed in empirical volatilities. Furthermore, MA processes can capture volatility jumps. The sample path behavior of the continuous-time MA process depends on both, the driving Lévy process L and the kernel function f. As a result of the simple structure of stable distributions, properties of stable MA processes are very well known; we refer to the excellent monograph of Samorodnitsky and Taqqu [137]. The extremal behavior of stable MA processes had already been investigated in 1978 by Rootzén [130].

In this thesis, we are concerned with the whole class of *subexponential Lévy driven MA processes*, this means MA processes driven by a Lévy process, whose increment distribution is subexponential. There have been many open questions concerning this class: starting with the existence of a stationary version of Y we proceed to extreme value problems as the tail behavior of the marginal distribution, tail behavior of functionals of Y, extremal behavior of local maxima, sample path behavior in the neighborhood of local maxima, running maxima and clustering of extremes. The sample path behavior of subexponential Lévy driven MA processes is similar to subexponential discrete-time MA processes: they are heavy tailed and have clusters in extremes, such that they constitute an ideal class for modelling high frequency financial data.

Necessary and sufficient conditions for the existence of Lévy driven MA processes are already given in Rajput and Rosinski [121], Theorem 2.7. They depend on both the generating triplet of the Lévy process L and the properties of the kernel function f. In our framework we calculate necessary conditions for a subexponential Lévy driven MA process depending only on the kernel function f. In the case of regularly MA processes they are nearly necessary.

Increments of Lévy processes are infinitely divisible, hence we are concerned with infinitely divisible subexponential distributions. An important characterization of infinitely divisible subexponential distributions is the tail equivalence of probability measure and Lévy measure. For calculating the tail behavior of the stationary distribution of the MA process we use this property. As intended they are also heavy tailed even subexponential and in the same maximum domain of attraction as the driving Lévy process. Hence, they are possible candidates for modelling volatility processes. A basic difference in the two regimes MDA(Φ_α) and MDA(Λ) is that regularly varying MA processes are tail equivalent to the increments of the Lévy process; but the tails of subexponential Lévy driven MA processes in MDA(Λ) is considerably lighter than the tails of the increments of the driving Lévy process.

Our investigation on the extremal behavior is based on a discrete-time skeleton of Y chosen to incorporate those times, where large jumps of the Lévy process and extremes of the kernel function simultaneously occur. In the following we give an intuitive explanation of this approach. For the sake of argument we restrict our attention to a subexponential MA process, where the Lévy process is positive and has countably many jumps. We denote the jump times by $\{\Gamma_k\}_{k\in\mathbb{N}}$ and assume that f has a unique maximum in η with value $f(\eta) = f^+$. As a result of the heavy tails, the Lévy process has jumps, which are larger than the typical ones. Hence Y achieves an extreme value nearly the largest jump of L multiplied with the largest value of f, which is f^+. This value is by (0.0.4) achieved in $Y(\Gamma_k+\eta)$ for some $k \in \mathbb{N}$. Motivated by this idea, which we will make mathematically precise also for general subexponential Lévy driven MA processes, the study of the extremal behavior of the continuous-time process Y is reduced to study the extremal behavior of the discrete-time skeleton $\{Y(\Gamma_k+\eta)\}_{k\in\mathbb{N}}$. This sequence has similar extremal behavior as a discrete-time MA process, such that we can use analog techniques as there. The extremal behavior of these sequence $\{Y(\Gamma_k+\eta)\}_{k\in\mathbb{N}}$ is described by point processes of high level exceedances of $Y(\Gamma_k+\eta)$. An interpretation of the limit process allows to provide an interpretation of the extreme behavior of the process. Our results are derived under a weak condition ensuring a Poisson type limit.

All these considerations concern in a first step the discrete-time skeleton only and ignore the fact that we deal with continuous-time processes. We furthermore introduce marks as the proper concept to derive the limiting behavior of the excursions of Y above high levels. In order to retain information about Y near $\Gamma_k+\eta$ we attach to each $Y(\Gamma_k+\eta)$ a mark, namely $\{Y(\Gamma_k+t_i)\}_{i=1,\ldots,d}$ for $t_1,\ldots,t_d \in \mathbb{R}$ and $\sup_{s\in I_k} Y(s)$, where I_k is a surrounding interval of $\Gamma_k+\eta$, such that the marked point process describes the sample path behavior near its extremes. Interesting questions, we investigate, concern the length of the excursion and the rate of "decrease" after $\Gamma_k+\eta$. Hence, we obtain perfect information about the extremal behavior of the continuous-time process.

The limit process of the marked point process turns out to be different in different regimes. As a measure for clusters in extremes we use the extremal index of the discrete-time sequence $\{M_k\}_{k\in\mathbb{N}}$ as given in (0.0.2). For regularly varying Lévy driven MA processes the extremal index is always less than one, which indicates clusters in extremes. In contrast to this clusters of subexponential Lévy driven MA processes in MDA(Λ) only occur, if the kernel function f has more than one local supremum or infimum, respectively. Hence from this point of view, subexponential Ornstein-Uhlenbeck processes in MDA(Λ) are not appropriate for modelling volatility clusters. Furthermore, a very important qualitative difference of the extremal behavior in these both regimes is that regularly varying MA processes have long high level excursions in contrast to MA processes in $\mathcal{S} \cap \mathrm{MDA}(\Lambda)$, where they collapse into single points. These points are also determined by the extremes of the kernel function f. Both cases have in common that extremes are caused by large jumps of the Lévy process and extremes of the kernel function. Thus, the normalizing constants of the running maxima as given in (0.0.1) are calculated by the right and left tails of $L(1)$ in combinations with the values of the supremum and the infimum of the kernel function f.

In the case of Lévy processes with $L(1) \in \mathcal{R}_{-\alpha}$ the class of processes we are studying is much larger than (0.0.4). There we replace L by an infinitely divisible, independently scattered random measure, which is a generalization of a Lévy process. Thus we obtain a larger class of continuous-time models, which is more flexible for modelling long memory in the data. In this case (0.0.4) are mixed MA processes and include in particularly Lévy driven MA processes. An interesting example is the superposition of Ornstein-Uhlenbeck processes, also applied as stochastic volatility model by Barndorff-Nielsen and Shephard [14].

In addition, we study the extremal behavior of Ornstein-Uhlenbeck processes with exponential tails, namely the Γ-Ornstein-Uhlenbeck process and Ornstein-Uhlenbeck processes driven by a convolution-equivalent Lévy process (Definition A.1.3). Convolution-equivalent distribution functions are generalizations of subexponential distribution functions, hence they share the important property of tail-equivalence of the probability measure and the Lévy measure. Moreover, they all belong to $\mathrm{MDA}(\Lambda)$. An important example in finance are the class of generalized inverse Gaussian distributions; the normal inverse Gaussian model has been prominent, see Barndorff-Nielsen [12]. In this thesis we include the Γ-Ornstein-Uhlenbeck process for its importance in the context of variance gamma models; see Madan and Seneta [103]. We will also compute the tail behavior of the stationary version of the Ornstein-Uhlenbeck process, the point process behavior of local maxima, the normalizing constants of running maxima and the extremal index function. The results are analog to subexponential Ornstein-Uhlenbeck-processes in $\mathrm{MDA}(\Lambda)$: they exhibit heavy tails, but can not model volatility clusters.

A completely different approach to obtain continuous-time volatility models starts with a GARCH model and derives from this discrete-time model a continuous-time model. GARCH models and their extensions have been in the limelight as appropriate models to capture certain empirical facts of volatility processes; see Engle [61] for an overview on GARCH modelling and Mikosch and Stărică [113] for their extremal behavior. We use the approach of Klüppelberg, Lindner and Maller [87], who started with a discrete-time GARCH(1,1) model and replaced the noise variables by a Lévy process with jumps $\Delta L_t = L_t - L_{t-}$, $t \geq 0$. The left continuous volatility process is defined by

$$dY(t+) \quad = \quad \beta dt + Y(t)\, \mathrm{e}^{X(t-)}\, d(\mathrm{e}^{-X(t)})\,,$$

where $\beta > 0$. The auxiliary càdlàg process X is defined by

$$X(t) = t \log \eta - \sum_{0<s\leq t} \log(1 + \lambda\eta(\Delta L(s))^2)\,,$$

for $\eta > 1$ and $\lambda \geq 0$. It is an essential feature of this model that the price process $S(t)$ is given by $dS(t) = \sqrt{Y(t)}\, dL(t)$. This continuous-time GARCH(1,1) model is called a COGARCH(1,1) model. COGARCH processes have the important feature to model upward jumps as well as to have a heavy tailed stationary distribution. Similar to the Lévy driven MA process, we choose a discrete-time skeleton of the process by the jump times of a compound Poisson process L for modelling the extremal behavior. Applications of classical results on extreme value theory of random recurrence equations by de Haan et al. [57] yield that they have clusters of extremes.

An excellent review article on volatility modelling of the COGARCH process against the Ornstein-Uhlenbeck process is Klüppelberg et al. [88].

A general guideline

This thesis is divided into three chapters, which are based on the papers [63, 62, 64]. Chapter 1 deals with the extremal behavior of subexponential Lévy driven MA processes in MDA(Λ) and Chapter 2 is devoted to subexponential Lévy driven mixed MA processes in MDA(Φ_α), which contains as special class the Lévy driven MA processes. Finally, Chapter 3 is concerned with financial applications: we investigate classical and non-classical volatility models: generalized Cox-Ingersoll-Ross models, subexponential Ornstein-Uhlenbeck processes, Ornstein-Uhlenbeck processes with exponential tails and COGARCH processes. Here we draw heavily from the more technical results in Chapter 1 and Chapter 2, complementing known results on the more traditional generalized Cox-Ingersoll-Ross models with new results.

For the technical results, in particular, of Chapter 1 and Chapter 2, an extensive notation was necessary. Consequently, it was not possible to use a consistent notation in the thesis. In order to avoid misunderstandings, the notation is explained in detail in each chapter. Moreover, each chapter starts with a detailed introduction including a guideline.

In the following we present a guideline to the thesis, summarized from the introduction to each chapter.

Chapter 1. We present in Section 1.1 the necessary definitions and basic results of Lévy driven MA processes Y in $\mathcal{S} \cap \mathrm{MDA}(\Lambda)$. Hence, Section 1.1.1 is devoted to subexponential distributions on the real line and the properties we need. A section on stationarity of MA processes in $\mathcal{S} \cap \mathrm{MDA}(\Lambda)$ follows. Section 1.1.3 is devoted to the most important examples of MA processes. A prominent example is a process driven by a compound Poisson process, which is also called Poisson shot noise process in the literature. Throughout this chapter and Chapter 2 Poisson shot noise processes form the basic structure of our results and proofs. A short summary on simple notions of extreme value theory concludes Section 1.1.

In Section 1.2 we derive general results for weak convergence of point processes (in Section 1.2.1), marks (in Section 1.2.2) and for marked point processes (in Section 1.2.3) for subexponential processes in the maximum domain of attraction of the Gumbel distribution. Weak convergence of point processes are fundamental for our continuous-time process as its extreme behavior is governed by a discrete-time skeleton. Our results are derived under a weak condition ensuring a Poisson type limit. Such results apply immediately to discrete-time MA processes. We furthermore introduce marks as the proper concept to derive the limiting behavior of excursions of continuous-time processes.

Section 1.3 deals with the tail behavior of $Y(0)$ and the tail behavior of certain functionals of Y. Proofs are based on the fact that Lévy measures and their corresponding distributions are tail-equivalent in the subexponential case and Rosinski and Samorodnitsky [132].

The results of Sections 1.2 and 1.3 are applied in Section 1.4 to the continuous-time process Y by choosing a proper discrete-time skeleton. In the case that f has a unique supremum we show in Section 1.4.2 that the running maxima of Y happen exactly at jump times of L shifted by the argument, where the supremum of f happens.

Furthermore, in Section 1.4.3 we calculate the limit distribution of running maxima of Y. Finally, in Section 1.4.4 we introduce the extremal index function as a proper concept for modelling clusters of exceedances for continuous-time processes.

Chapter 2. We start with a short introduction into multivariate regular variation in Section 2.1 followed by an investigation of regularly varying mixed MA processes Y in Section 2.2. This includes on the one hand sufficient conditions for the stationarity of the mixed MA process Y in Section 2.2.1 and on the other hand the tail behavior of the mixed MA process Y and the tail behavior of $M(h) = \sup_{t\in[0,h]} Y(t)$ for $h > 0$ in Section 2.2.2. Finally, Section 2.2.3 gives with FICARMA and supOU processes the most prominent examples for regularly varying mixed MA processes, which exhibit long range dependence. In Section 2.3 we obtain analogous results as for regularly varying mixed MA processes for regularly varying renewal shot noise processes.

Section 2.4 is concerned with the point process behavior of multivariate regularly varying stationary sequences. We present analog results as given in Section 1.2 for stationary sequences in $\mathcal{S}\cap\mathrm{MDA}(\Lambda)$ for multivariate regularly varying stationary sequences. Moreover, Davis and Mikosch [53] generalize results of Davis and Hsing [52] on the point process behavior of stationary processes with regularly varying tails to a multivariate setting. We give an overview of their results, which are then applied in Section 2.5. First, we study the asymptotic behavior of the embedded marked point process of local maxima of Y in Section 2.5.1. Afterwards, in Section 2.5.2, we present under less restrictive assumptions than in Section 2.5.1 a marked point process result, which also includes the behavior of large infima of Y. Moreover, we obtain the limit distribution of running maxima of the mixed MA process in Section 2.5.3, and compute the extremal index function in Section 2.5.4. The results are in particular valid for stationary renewal shot noise processes.

Chapter 3. This chapter is based on joint work with Claudia Klüppelberg and Alexander Lindner. We investigate classical and non classical stochastic volatility models with respect to their extreme behavior. Classical volatility models as the generalized Cox-Ingersoll-Ross model can model heavy tails, but are not able to model volatility jumps. We review, in Section 3.1, their extremal behavior, which can be in different maximum domain of attractions. Volatility jumps can be modelled by Lévy driven Ornstein-Uhlenbeck processes, which are content of Section 3.2. Their extremal behavior is characterized by the extremal behavior of the driving

Lévy process, so that we have to distinguish between different classes of driving Lévy processes. So Section 3.2.1 concerns the two subexponential models. Both can capture heavy tails but only the regularly varying Ornstein-Uhlenbeck process has clusters.

Then, in Section 3.2.2, we study Ornstein-Uhlenbeck processes with exponential tails. As a prominent example we investigate the Γ-Ornstein-Uhlenbeck process, i. e. the stationary volatility is gamma distributed. As an important larger class we study Ornstein-Uhlenbeck processes, whose driving Lévy process belongs to the class of convolution-equivalent tails. These classes extend subexponential Lévy processes in a natural way; see Definition A.1.3. It turns out that for all Ornstein-Uhlenbeck processes in Section 3.2, high level volatility clusters are exhibited only in the case of regularly varying Lévy processes.

The last class of models reviewed in this chapter concerns the COGARCH process in Section 3.3. In contrast to the Lévy driven Ornstein-Uhlenbeck process considered before, the COGARCH volatility has heavy tails under quite general conditions on the driving Lévy process L. Furthermore, the COGARCH exhibits high level volatility clusters.

Finally, in Section 3.4 a short conclusion is given. Here we compare the models introduced in Chapter 3. It turns out that there is a striking similarity concerning the extremal behavior of models with the same stationary distribution. Here we also discuss briefly some further empirical facts of volatility data.

Chapter 1

Extremes of subexponential Lévy driven MA processes

In this chapter we investigate the extremal behavior of a stationary *continuous-time moving average* (MA) process of the form

$$Y(t) = \int_{-\infty}^{\infty} f(t-s)\, dL(s) \qquad \text{for } t \in \mathbb{R}\,, \tag{1.0.1}$$

where $f : \mathbb{R} \to \mathbb{R}$, called *kernel function*, is measurable, and the driving process $L = \{L(t)\}_{t\in\mathbb{R}}$ is a *Lévy process*. We recall that a general Lévy process L has independent and stationary increments, $L(0) = 0$ and is càdlàg, i. e. L has a. s. (almost surely) right continuous sample paths with left hand limits. To make the definition of $Y = \{Y(t)\}_{t\in\mathbb{R}}$ meaningful we define the driving Lévy process on $\mathbb{R}$ by gluing together two independent Lévy processes $\{L_+(t)\}_{t\geq 0}$ and $\{L_-(t)\}_{t\geq 0}$ with identical distribution, so that $L(t) = L_+(t)$ for $t \geq 0$ and $L(t) = -L_-(-t-)$ for $t < 0$. A Lévy process $L = \{L(t)\}_{t\in\mathbb{R}}$ on the real line is characterized by the *Lévy-Khinchine representation* of its characteristic function $\mathbb{E}(\mathrm{e}^{iuL(t)}) = \exp(t\psi(u))$ for $t\,, u \in \mathbb{R}$ with

$$\psi(u) = ium - \frac{1}{2}u^2\sigma^2 + \int_{\mathbb{R}} \left(\mathrm{e}^{iux} - 1 - iu\kappa(x)\right)\, d\nu(x) \qquad \text{for } u \in \mathbb{R}\,, \tag{1.0.2}$$

and $\kappa(x) = x\,\mathbf{1}_{[-1,1]}(x)$, where $\mathbf{1}_A$ denotes the indicator function of the set A. The quantities (m, σ^2, ν) are called the *generating triplet* of the Lévy process L. Here $m \in \mathbb{R}$, $\sigma^2 \geq 0$ and ν is a measure on $\mathbb{R}$, called *Lévy measure*, satisfying $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}}(1 \wedge |x|^2)\,\nu(dx) < \infty$.

We shall decompose L in three independent Lévy processes according to its jump sizes, which are represented by ν: $L = L_1 - L_2 + L_3$, where $L_1 = \{L_1(t)\}_{t\in\mathbb{R}}$ and $L_2 = \{L_2(t)\}_{t\in\mathbb{R}}$ are independent positive Lévy processes, with Lévy measures

$$\nu_1(A) = \nu(A \cap (1,\infty)) \quad \text{and} \quad \nu_2(A) = \nu(-A \cap (-\infty,-1)) \quad \text{for } A \in \mathcal{B}(\mathbb{R})$$

and generating triplet $(0, 0, \nu_i)$ for $i = 1, 2$. The Lévy process $L_3 = \{L_3(t)\}_{t\in\mathbb{R}}$ has Lévy measure

$$\nu_3(A) = \nu(A \cap [-1,1]) \quad \text{for } A \in \mathcal{B}(\mathbb{R}),$$

i. e. it has finite support. Then L_1 and L_2 are increasing compound Poisson processes whose jumps are larger than 1, and L_3 has jumps with modulus only smaller than 1 and generating triplet (m, σ^2, ν_3). Throughout this chapter we shall need properties and results for Lévy processes; we refer to the three monographs Applebaum [9], Bertoin [27] and Sato [138].

This decomposition of L induces a decomposition of Y giving $Y = Y_1 - Y_2 + Y_3$, where for $i = 1, 2, 3$,

$$Y_i(t) = \int_{-\infty}^{\infty} f(t-s)\, dL_i(s) \quad \text{for } t \in \mathbb{R} \tag{1.0.3}$$

are independent MA processes. We have to be careful that Y_i is a stationary i. d. process. Note already here, that Y_1 is positive provided that f is positive.

The sample path behavior of Y depends on the driving Lévy process and on the kernel function. We also consider in this chapter only the short memory case, where the kernel function (which is also sometimes called *memory function*) decreases sufficiently fast; see Remark 1.1.8 *(b)*. Extreme value theory for continuous-time (and discrete-time) stochastic processes has meanwhile a long history. A theory for Gaussian processes was developed during the eighties; see e. g. Albin [2,3], Berman [25,26], Leadbetter and Rootzén [96, 97] and Leadbetter et al. [95]. Under appropriate mixing conditions, in the short memory case, when for instance Berman's condition holds, Gaussian processes behave like i. i. d. random variables with distribution of the maxima of the process over a fixed interval.

Extreme value theory for stable MA processes was derived by Rootzén [130]. Such models have a completely different extreme behavior than Gaussian models. Increments of L have infinite second moment, hence large values are likely to occur leading immediately to regularly varying stationary distributions. Moreover, large values are carried on in time by the kernel function causing long high level excursions of the process Y.

We are concerned with the extremal behavior of the process Y driven by a subexponential Lévy process. Subexponential models are typical models for situations, where extremely large values are likely to occur in comparison to the mean size of the data. In our context this concerns the increments of the process L, which are subexponential; see Definition 1.1.1. The thesis is on *subexponential Lévy driven MA processes.* They also include the class of stable distributions, but are a much richer class. We concentrate in this chapter on subexponential distributions in the *maximum domain of attraction of the Gumbel distribution*, i. e. there exist constants $a_T > 0$, $b_T \in \mathbb{R}$ for $T > 0$ such that $\lim_{T\to\infty} T\mathbb{P}(L(1) > a_T x + b_T) = \exp(-x) = -\log\Lambda(x)$. Those distributions have moments of every order and their tails decrease faster than polynomial. Examples include the lognormal and the heavy-tailed Weibull distributions; see Example 1.1.4. Under appropriate conditions it will turn out that the stationary distribution of Y is subexponential and *tail-equivalent* to $f(tU)L(1)$ for some $t > 0$ and an uniform r. v. U on $(-1, 1)$, i. e. both have the same tail behavior. High level excursions of Y are, in contrast to the stable model, no longer persistent; in the limit they collapse into singular time points.

Throughout the chapter we shall need the following conditions on the Lévy process.

Condition (L1).
The marginal distribution $L(1)$ of the Lévy process L satisfies $L(1) \in \mathcal{S}\cap\mathcal{R}_{-\infty}$ ($L(1)$ is subexponential (Definition 1.1.1) with rapidly varying right tail (Definition 1.1.3)).

Condition (L2).
There exists a $p \in (0, 1]$ such that the following tail balance condition holds:

$$\lim_{x\to\infty} \frac{\nu\,(-\infty, -x]}{\nu\,[x, \infty)} = \frac{1-p}{p}.$$

If the support of ν is bounded below we assume w. l. o. g. $\nu\,(-\infty, -1] = 0$, else we choose another decomposition of L. If the support of ν is not bounded below, then we assume further $p \in (0, 1)$.

If $p \in (0, 1)$ then $-L(1) \in \mathcal{S} \cap \mathcal{R}_{-\infty}$. In the case $p = 1$ the accessorily condition that the support of ν is bounded below can be relaxed, but then the proofs are more complex.

For the kernel function we require the following condition.

Condition (K1).
The kernel function $f : \mathbb{R} \to \mathbb{R}$ is a regulated function, i. e. left and right hand limits exist, and for some $\delta \in (0,1)$,

$$\int_{-\infty}^{\infty} \sup_{0<s<1} |f(s+t)|^{\delta} \, dt < \infty.$$

Hence f is bounded and has at most countable many discontinuities. We write $f^+(t) = \max\{0, f(t)\}$, $f^+ = \sup_{t\in\mathbb{R}} f^+(t)$, $f^-(t) = \max\{0, -f(t)\}$, $f^- = \sup_{t\in\mathbb{R}} f^-(t)$ and assume $f^+ \geq f^-$.

We shall show in Remark 1.1.8 *(b)* that such kernel functions imply short memory of the process Y: the covariance function is integrable.

Condition (K2).
Define

$$O_1 := \left\{\alpha \in \mathbb{R} : \quad f(\alpha) = f^+\right\}, \qquad O_2 := \left\{\alpha \in \mathbb{R} : f(\alpha) = -f^+\right\}.$$

Then we define for $i = 1,2$, $P^{(i)} := \operatorname{card} O_i < \infty$ and denote $O_i = \{\alpha_1^{(i)}, \ldots, \alpha_{P^{(i)}}^{(i)}\}$, with $\alpha_1^{(i)} < \ldots < \alpha_{P^{(i)}}^{(i)}$. The support of f is a connected interval. f is continuous in the interior of its support $[a,b]$, $[a,\infty)$, $(\infty, a]$, respectively with $a < b$ and right and left continuous at the boundary.

Thus kernel functions, which are piecewise constant in their extremes, are excluded.

Our investigation is based on a discrete-time skeleton $\{Y(t_n)\}_{n\in\mathbb{N}}$ of Y chosen as to incorporate those times, where big jumps of the Lévy process and extremes of the kernel function occur. We embed the normalized process $\{Y(t_n)\}_{n\in\mathbb{N}}$ in a point process and derive the weak limit of this sequence of point processes. Not surprisingly, we find a strong analogy to discrete-time MA processes and corresponding results of Davis and Resnick [55] and Rootzén [131]. We model the path behavior of the continuous-time process near high level excursions by a mark on the point process. Obviously marks are influenced by the kernel function and its local suprema. We shall show that they behave asymptotically like the deterministic functions f/f^+ or $-f/f^+$. Our findings show that specifically chosen discrete-time points determine the extremal behavior of the continuous-time process.

The chapter is organized as follows. In Section 1.1 we summarize the necessary definitions and basic results. Section 1.1.1 is devoted to subexponential distributions on the real line and the properties we shall need. A section on stationarity follows. Section 1.1.3 is devoted to the most important examples. Prominent in this chapter will be processes driven by a compound Poisson process, which are also called *Poisson shot noise processes* in the literature. Throughout this chapter, Poisson shot noise processes form the basic structure for our results and proofs. This is due to the fact that they correspond to the processes Y_1 and Y_2 in the decomposition (1.0.3). A short summary on simple notions of extreme value theory concludes Section 1.1.

In Section 1.2 we derive general results for weak convergence of point processes (in Section 1.2.1), marks (in Section 1.2.2) and for marked point processes (in Section 1.2.3). Weak convergence of point processes are fundamental for our continuous-time process as its extreme behavior is governed by a discrete-time skeleton. Our results are derived under a weak condition ensuring a Poisson type limit. Such results apply also immediately to discrete-time MA processes. We furthermore introduce marks as the proper concept to derive the limiting behavior of excursions of Y.

Section 1.3 is devoted to the tail behavior of $Y(0)$ and the tail behavior of certain functionals of Y. Proofs are based on the fact that Lévy measures and their corresponding distribution functions are tail-equivalent in the subexponential case and Rosinski and Samorodnitsky [132].

The results of Sections 1.2 and 1.3 are applied in Section 1.4.1 to the continuous-time process Y by choosing the proper discrete-time skeleton. In the case that f has an unique supremum we show in Section 1.4.2 that the running maxima of Y happen exactly at jump times of L shifted by the argument, where the supremum of f happens. Furthermore, in Section 1.4.3 we calculate the limit distribution of running maxima of Y. Finally, in Section 1.4.4 we introduce the extremal index function as proper concept for modelling clusters of exceedances for continuous-time processes.

1.1 Definitions and auxiliary results

Throughout the chapter we shall use the following notation: we write $\overline{F} = 1 - F$ for the right tail of the distribution function (d. f.) F, F^{2*} for the convolution $F * F$ and $\overline{F^{2*}} = 1 - F^{2*}$. For any random variable (r. v.) Z on $\mathbb{R}$ we write $Z^+ = Z \vee 0$ and $Z^- = -Z \vee 0$; $X \stackrel{d}{=} Y$, if the distributions of the random variables (r. v. s) X and Y coincide. For real functions g and h we abbreviate $g(t) \sim h(t)$ for $t \to \infty$, if $g(t)/h(t) \stackrel{t\to\infty}{\longrightarrow} 1$ and we denote $g^+(t) = \max\{0, g(t)\}$, $g^-(t) = \max\{0, -g(t)\}$, $g^+ = \sup_{t\in\mathbb{R}} g^+(t)$, $g^- = \sup_{t\in\mathbb{R}} g^-(t)$.

1.1.1 Subexponential distributions on the real line

Subexponentiality is a property of the right tail of a distribution. Consequently, it has been defined originally for positive r. v. s..

In the context of this chapter $L(1)$ has a distribution on the whole of $\mathbb{R}$, which has a subexponential right tail. The definition of a subexponential r. v. has been extended from a positive r. v. to a r. v. on $\mathbb{R}$ by Willekens [144], and we shall start with this definition.

Definition 1.1.1
Let F be a d. f. on $\mathbb{R}$ with $F(x) < 1$ for every $x \in \mathbb{R}$.

(i) F belongs to the class of long tailed distributions, denoted by $\mathcal{L}$, if for all $y \in \mathbb{R}$ locally uniformly $\lim_{x\to\infty} \overline{F}(x+y)/\overline{F}(x) = 1$.

(ii) F belongs to the class of subexponential distributions, denoted by $\mathcal{S}$, if $F \in \mathcal{L}$ and $\lim_{x\to\infty} \overline{F^{2*}}(x)/\overline{F}(x)$ *exists and is finite.*

If $F \in \mathcal{L}$ or $F \in \mathcal{S}$ and Z is a r. v. with d. f. F, then we write $Z \in \mathcal{L}$ or $Z \in \mathcal{S}$.

The class $\mathcal{S}$ is closed under *tail-equivalence*, i. e. if $F \in \mathcal{S}$ and G is a d. f. with $\lim_{x\to\infty} \overline{F}(x)/\overline{G}(x) = q \in (0, \infty)$, then also $G \in \mathcal{S}$. It helps to think of a subexponential r. v. Z on $\mathbb{R}$ as a r. v. whose positive part $Z^+ \in \mathcal{S}$. A survey of the class of subexponential distributions with support on $\mathbb{R}_+$ is provided by Goldie and Klüppelberg [70], see also Embrechts et al. [60]. The following result summarizes mostly known properties of subexponentials on $\mathbb{R}$ needed for this chapter.

Proposition 1.1.2

(i) *If* $F \in \mathcal{L}$, *then* $\overline{F}(x/2)^2 = o(\overline{F}(x))$ *for* $x \to \infty$ *and* $\lim_{x\to\infty} \mathrm{e}^{\epsilon x}\overline{F}(x) = \infty$ *for any* $\epsilon > 0$.

(ii) *If* $F \in \mathcal{S}$, *then* $\lim_{x\to\infty} \overline{F^{2*}}(x)/\overline{F}(x) = 2$.

(iii) *Suppose* $F \in \mathcal{S}$, F_i *d. f. with* $\lim_{x\to\infty} \overline{F_i}(x)/\overline{F}(x) = q_i \geq 0$ *for* $i = 1, 2$ *and* $G = F_1 * F_2$. *Then*

$$\lim_{x\to\infty} \frac{\overline{G}(x)}{\overline{F}(x)} = q_1 + q_2. \tag{1.1.1}$$

If $q_i > 0$ *for some* $i \in \{1, 2\}$, *then also* F_i, $G \in \mathcal{S}$. *Moreover, for* $q_1 > 0$,

$$\lim_{x\to\infty} \int_{-\infty}^{x/2} \frac{\overline{F}_2(x-u)}{\overline{F}_1(x)} F_1(du) = \frac{q_2}{q_1}, \quad \lim_{x\to\infty} \int_{-\infty}^{x/2} \frac{\overline{F}_1(x-u)}{\overline{F}_1(x)} F_2(du) = 1. \tag{1.1.2}$$

(iv) *Let* N *be a Poisson r. v. with mean* μ *and* $\{X_k\}_{k\in\mathbb{N}}$ *be an i. i. d. sequence with d. f.* $F \in \mathcal{S}$. *Then the r. v.* $Y = \sum_{k=1}^{N} X_k$ *has d. f.* $G = \mathrm{e}^{-\mu} \sum_{n=0}^{\infty} \frac{\mu^n}{n!} F^{*n}$, $G \in \mathcal{S}$ *and*

$$\overline{G}(x) \sim \mu\overline{F}(x) \quad \text{for } x \to \infty.$$

(v) *Let* F *be an infinitely divisible (i. d.) d. f. with Lévy measure* ν. *Then*

$$F \in \mathcal{S} \quad \Longleftrightarrow \quad \frac{\nu\,[1,x)}{\nu\,[1,\infty)} \in \mathcal{S} \ \text{ for } x > 1 \quad \Longleftrightarrow \quad \overline{F}(x) \sim \nu(x,\infty) \text{ for } x \to \infty.$$

(vi) *If* $X \in \mathcal{S}$ *then* $XY \in \mathcal{S}$ *if and only if* $X^+Y^+ \vee X^-Y^- \in \mathcal{S}$. *If* $X \in \mathcal{S}$ *has only support on* $\mathbb{R}_+$ *and* Y *is a bounded r. v., then* $XY \in \mathcal{S}$.

Proof.

(i) Cline [44], Lemma 2.

(ii) Pakes [115], Corollary 2.1.

(iii) Pakes [115], Lemma 2.4 and Lemma 5.1 proves (1.1.1) of *(iii)* and F_i, $G \in \mathcal{S}$ if $q_i > 0$ for some $i \in \{1, 2\}$. In the case $q_1 + q_2 = 0$ we calculate for $x \in \mathbb{R}$

$$\overline{F_1 * F_2}(x) = \int_{-\infty}^{x/2} \overline{F}_2(x-u)F_1(du) + \int_{-\infty}^{x/2} \overline{F}_1(x-u)F_2(du) + \overline{F}_1\,(x/2)\,\overline{F}_2\,(x/2)\,.$$

Taking into account that $F \in \mathcal{S}$ and *(i)* holds, we obtain

$$\frac{\overline{F}_1\,(x/2)\,\overline{F}_2\,(x/2)}{\overline{F}(x)} = \frac{\overline{F}_1\,(x/2)\,\overline{F}_2\,(x/2)}{\overline{F}(x/2)\overline{F}(x/2)} \frac{\overline{F}(x/2)\overline{F}(x/2)}{\overline{F}(x)} \overset{x\to\infty}{\longrightarrow} 0.$$

Let $\epsilon > 0$ be arbitrary. Then there exists an $x_1 > 0$ such that $0 \le \overline{F}_1(x)/\overline{F}(x) \le \epsilon$ for every $x \ge x_1$. Consequently,

$$\int_{-\infty}^{x/2} \frac{\overline{F}_1(x-u)}{\overline{F}(x)} F_2(du) \le \epsilon \int_{-\infty}^{x/2} \frac{\overline{F}(x-u)}{\overline{F}(x)} F_2(du) \le \epsilon \frac{\overline{F * F_2}(x)}{\overline{F}(x)} \stackrel{x\to\infty}{\longrightarrow} \epsilon,$$

where we used *(iii)*. Similarly, the second summand tends to zero.
If $q_1 > 0$, by *(i)*, for $x \to \infty$,

$$\begin{aligned} \overline{F_1 * F_2}(x) &= \int_{-\infty}^{x/2} \overline{F}_1(x-u)\, F_2(du) + \int_{-\infty}^{x/2} \overline{F}_2(x-u)\, F_1(du) + \overline{F}_1\left(\frac{x}{2}\right) \overline{F}_2\left(\frac{x}{2}\right) \\ &\sim \int_{-\infty}^{x/2} \overline{F}_1(x-u)\, F_2(du) + \int_{-\infty}^{x/2} \overline{F}_2(x-u)\, F_1(du). \end{aligned} \qquad (1.1.3)$$

Applying Fatou's Lemma, we obtain the lower bound

$$\liminf_{x\to\infty} \int_{-\infty}^{x/2} \frac{\overline{F}_2(x-u)}{\overline{F}_1(x)}\, F_1(du) \ge \frac{q_2}{q_1}, \quad \liminf_{x\to\infty} \int_{-\infty}^{x/2} \frac{\overline{F}_1(x-u)}{\overline{F}_1(x)}\, F_2(du) \ge 1 \qquad (1.1.4)$$

and hence by (1.1.1) and (1.1.3) the inequalities in (1.1.4) are equalities.

(iv) Pakes [115], Theorem 5.1.

(v) Pakes [115], Theorem 3.1.

(vi) $XY \in \mathcal{S} \iff X^+Y^+ \vee X^-Y^- = X^+Y^+ + X^-Y^- = (XY)^+ \in \mathcal{S}$. The second statement has been proved by Cline and Samorodnitsky [46], Corollary 2.5. □

We shall also need certain properties of regularly varying functions. For further definitions and properties we refer to Bingham et al. [29].

Definition 1.1.3
A positive measurable function $u : \mathbb{R} \to \mathbb{R}_+$ is regularly varying with index α, denoted by $u \in \mathcal{R}_\alpha$ for $\alpha \in \mathbb{R}$, if

$$\lim_{t\to\infty} \frac{u(tx)}{u(t)} = x^\alpha \quad \textit{for } x > 0\,.$$

u is said to be slowly varying if $\alpha = 0$ and rapidly varying, denoted by $u \in \mathcal{R}_{-\infty}$, if the above limit is equal to 0 for $x > 1$ and ∞ for $0 < x < 1$. (The case $\alpha = \infty$ is also possible, but not relevant in this thesis.)

The class of subexponential distributions includes all distributions with regularly varying tails, the lognormal distribution and heavy-tailed Weibull distributions. A prominent example in the context of this chapter is the following:

Example 1.1.4 (Extended heavy-tailed Weibull model)
Let the right tail of the d. f. F behave like

$$\overline{F}(x) \sim \exp(-u(x)) \quad \text{for } x \to \infty,$$

where $u \in \mathcal{R}_\alpha$ with $\alpha \in (0,1)$ and there exist a $v > 1$ such that $u(tx) \leq x^\delta u(t)$ for all $t \geq v$ and $x > 1$, then $F \in \mathcal{S}$. The heavy-tailed Weibull distribution $\overline{F}(x) = K \exp(-x^\alpha)$, $x \geq 0$, for $\alpha \in (0,1)$ and $K > 0$, or Benktander Type II are special examples.

Proof.
Without loss of generality suppose $\overline{F}(x) = \exp(-u(x))$ for $x \in \mathbb{R}$. The statement of Proposition 3.7 *(a)*, Baltrunas et al. [11] is satisfied. Thus $u(t)$ is absolutely continuous on $t \geq v$ with Lebesgue density $q(t) \to 0$ as $t \to \infty$, and $tq(t)/u(t) \leq \rho$ for $t \geq v$. In particular, $\overline{F}(x) = \exp(-u(v))\exp(-\int_v^x q(t)\,dt)$ for $x \geq v$. Hence F is tail-equivalent to an absolutely continuous d. f., which is by Lemma 3.8. *(a)*, Baltrunas et al. [11], subexponential. Then also $F \in \mathcal{S}$. □

Note that one can construct examples, where $u \in \mathcal{R}_\alpha$, but $u(tx) \leq x^\delta u(t)$ for all $t \geq v$ and $x > 1$ is not satisfied, such that $F \notin \mathcal{S}$, see Cline [44], pp. 536.

We present some consequences of conditions (L1) and (L2).

Remark 1.1.5 (Condition (L1) and (L2))
(a) Let $L(1), -L(1) \in \mathcal{S}$. Then holds $\mathbb{P}(L(1) > x) \sim \mathbb{P}(L_1(1) > x) \sim \nu(x,\infty)$ and $\mathbb{P}(L(1) < -x) \sim \mathbb{P}(L_2(1) > x) \sim \nu(\infty, -x)$ for $x \to \infty$ by Proposition 1.1.2 *(v)*. Especially, (see also Lemma 1.3.4)

$$\begin{aligned}\mathbb{P}(|L(1)| > x) &= \mathbb{P}(L(1) > x) + \mathbb{P}(L(1) < -x)\\ &\sim \mathbb{P}(L_1(1) > x) + \mathbb{P}(L_2(1) > x) \quad \text{for } x \to \infty.\end{aligned}$$

If additionally (L2) is satisfied then Proposition 1.1.2 *(iii)* implies $|L(1)| \in \mathcal{S}$ and

$$\mathbb{P}(|L(1)| > x) = \mathbb{P}(L(1) > x) + \mathbb{P}(-L(1) > x) \sim \frac{1}{p}\mathbb{P}(L(1) > x) \text{ for } x \to \infty.$$

Suppose only $L(1) \in \mathcal{S}$ and $p = 0$. Then

$$\begin{aligned}\mathbb{P}(L_1(1) - L_2(1) - |L_3(1)| > x) &\leq \mathbb{P}(|L(1)| > x) \quad (1.1.5)\\ &\leq \mathbb{P}(L_1(1) + L_2(1) + |L_3(1)| > x).\end{aligned}$$

The Lévy measure of $L_1 + L_2$ is $\nu_1 + \nu_2$, which behaves asymptotically like ν. We obtain $L_1(1) + L_2(1) \in \mathcal{S}$ with $\mathbb{P}(L_1(1) + L_2(1) > x) \sim \mathbb{P}(L(1) > x)$ as $x \to \infty$. By Sato [138], Theorem 26.1, and Proposition 1.1.2 *(iii)* we have

$$\mathbb{P}(L_1(1) + L_2(1) + |L_3(1)| > x) \sim \mathbb{P}(L(1) > x) \quad \text{for } x \to \infty.$$

Similarly we obtain the same behavior for the right side of (1.1.5), such that $\mathbb{P}(|L(1)| > x) \sim \mathbb{P}(L(1) > x)$ for $x \to \infty$. Taking

$$\mathbb{P}(|L(1)| > x) = \mathbb{P}(L(1) > x) + \mathbb{P}(L(1) < -x)$$

into account yields further $\mathbb{P}(L(1) < -x)/\mathbb{P}(L(1) > x) \to 0$ for $x \to \infty$.

(b) For $i = 1, 2$, the r. v. s $L_i(1) \in \mathcal{S} \cap \mathcal{R}_{-\infty}$ are positive with $\mathbb{E}L_i(1)^q < \infty$ for $q > 0$ (cf. Embrechts et al. [60]). Applying Sato [138], Corollary 25.8 yields $\int_{|x|>1} |x|^q \nu_i(dx) < \infty$, for $q > 0$ and thus $\int_{|x|>1} |x|^q \nu(dx) < \infty$. Again by Sato [138], Corollary 25.8 we have $\mathbb{E}|L(1)|^q < \infty$ for $q > 0$. □

1.1.2 Stationarity

Before investigating extremal properties of the process Y given in (1.0.1) we summarize some basic results of continuous-time MA processes. Under certain conditions $Y(0)$ is well-defined as a limit in probability of integrals of step functions approximating f. This has been shown by Rajput and Rosinski [121] who also gives conditions for $Y(0)$ to be i. d.. The conditions are formulated in terms of the kernel function f and the generating triplet of the driving Lévy process.

Proposition 1.1.6 (Rajput and Rosinski [121], Theorem 2.7)
Let L be a Lévy process with generating triplet (m, σ^2, ν). Then $Y(0)$ given in (1.0.1) is well-defined and i. d. if and only if

$$\int_{-\infty}^{\infty} \left| mf(s) + \int_{-\infty}^{\infty} (\kappa(xf(s)) - f(s)\kappa(x))\, \nu(dx) \right| ds < \infty,$$
$$\int_{-\infty}^{\infty} |\sigma^2 f(s)|^2\, ds < \infty,$$
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \min\{1, |f(s)x|^2\}\, \nu(dx)\, ds < \infty,$$

where $\kappa(x) = x\,\mathbf{1}_{[-1,1]}(x)$*. The generating triplet is* (m_Y, σ_Y^2, ν_Y)*, where*

$$\begin{aligned} m_Y &= \int_{-\infty}^{\infty} mf(s) + \int_{-\infty}^{\infty} \left(\kappa(xf(s)) - f(s)\kappa(x)\right) d\nu(x)\,ds, \\ \sigma_Y^2 &= \sigma^2 \int_{-\infty}^{\infty} f^2(s)\,ds, \qquad (1.1.6) \\ \nu_Y\,[x,\infty) &= \int_{f(s)>0} \nu\left[\frac{x}{f(s)},\infty\right)ds + \int_{f(s)<0} \nu\left(-\infty,\frac{x}{f(s)}\right]ds \quad \text{for } x>0. \end{aligned}$$

The above integrals are originally defined as Lebesgue integrals, but in our framework Lebesgue and Riemann integrals are the same, so we shall not distinguish between them. As can also be seen, if the driving Lévy process L has no Gaussian component, then the second condition is meaningless, i. e. $Y(0)$ has a Gaussian component if and only if L has one. In the next Lemma we give some simple sufficient conditions for a MA process to be i. d..

Proposition 1.1.7
Let L *be a Lévy process with* $\mathbb{E}|L(1)|^\delta < \infty$ *for some* $\delta \in (0,1]$ *and suppose that* $f : \mathbb{R} \to \mathbb{R}$ *is measurable. Assume that*

$$\int_{-\infty}^{\infty} \sup_{0<s<1} |f(s+t)|^\delta\,dt < \infty. \qquad (1.1.7)$$

Then $Y(0)$ *given in (1.0.1) is i. d.. Moreover, the MA process* Y *given in (1.0.1) is stationary. The finite dimensional distributions of* $\widehat{Y}(t) = \int_{-\infty}^{\infty} f(t+s)\,dL(s)$ *for* $t \in \mathbb{R}$ *and* Y *coincide. In particular,*

$$Y(0) = \int_{-\infty}^{\infty} f(-s)\,dL(s) \stackrel{d}{=} \int_{-\infty}^{\infty} f(s)\,dL(s) = \widehat{Y}(0).$$

Proof.
Note that

$$\int_{-\infty}^{\infty} \sup_{0<s<1} |f(s+t)|^\delta\,dt < \infty \quad \Longleftrightarrow \quad \sum_{k=-\infty}^{\infty} \sup_{k<s<k+1} |f(s)|^\delta < \infty. \qquad (1.1.8)$$

As f is bounded we assume without loss of generality $|f(t)| \le 1$ for $t \in \mathbb{R}$. From (1.1.8) we conclude

$$\sum_{k=-\infty}^{\infty} \sup_{k<s<k+1} |f(s)|^\gamma < \infty$$

for $\gamma \geq \delta$ and by the comparison test $|f|^\gamma$ is integrable. Recalling that $\kappa(x) = x\,\mathbf{1}_{[-1,1]}(x)$ we estimate

$$\begin{aligned}
&\int_{-\infty}^{\infty} \left| mf(s) + \int_{-\infty}^{\infty} \kappa(xf(s)) - f(s)\kappa(x)\,\nu(dx) \right| ds \\
&\qquad \leq m \int_{-\infty}^{\infty} |f(s)|\,ds + \int_{-\infty}^{\infty} \int_{\substack{|f(s)x|\leq 1\\ |x|>1}} |xf(s)|\,\nu(dx)\,ds \\
&\qquad \leq m \int_{-\infty}^{\infty} |f(s)|\,ds + \int_{-\infty}^{\infty} |f(s)|^\delta\,ds \int_{|x|>1} |x|^\delta\,\nu(dx) < \infty,
\end{aligned}$$

where we used Fubini's theorem and $\int_{|x|>1} |x|^\delta \nu(dx) < \infty$ by Remark 1.1.5 *(b)*. On the other hand we have

$$\begin{aligned}
&\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \min\{1, |f(s)|^2 x^2\}\,\nu(dx)\,ds \qquad (1.1.9)\\
&= \int_{-\infty}^{\infty} f(s)^2\,ds \int_{-1}^{1} x^2\,\nu(dx) + \int_{|x|>1} \left[\int_{f(s)^2x^2>1} 1\,ds + \int_{\substack{f(s)^2x^2\leq 1\\ |x|>1}} f(s)^2 x^2\,ds \right] \nu(dx) \\
&\leq \int_{-\infty}^{\infty} f(s)^2\,ds \int_{-1}^{1} x^2\,\nu(dx) + \int_{|x|>1} \left[\int_{f(s)^2x^2>1} 1\,ds + \int_{-\infty}^{\infty} |f(s)|^\delta |x|^\delta\,ds \right] \nu(dx).
\end{aligned}$$

By the standard property of Lévy measures $\int_{\mathbb{R}} (1 \wedge |x|^2)\,\nu(dx) < \infty$ and the quadratic integrability of $|f|^2$ we yield that the first summand of (1.1.9) is finite. Again by the integrability of $|f|^\delta$ and as $\int_{|x|>1} |x|^\delta \nu(dx) < \infty$ we also conclude that the third term is finite. For the remaining part we have

$$\int_{|x|>1} \int_{f(s)^2x^2>1} 1\,ds\,\nu(dx) \;=\; \int_{f(s)\neq 0} \nu\left(-\infty, -\frac{1}{|f(s)|}\right) + \nu\left(\frac{1}{|f(s)|}, \infty\right) ds.$$

Define $c_k = \sup_{t\in[k-1,k)} |f(t)|$ for $k \in \mathbb{Z}$ and $\overline{f}(t) = \sum_{k=-\infty}^{\infty} c_k\,\mathbf{1}_{[k-1,k)}(t)$ for $t \in \mathbb{R}$. Let $\{Z_k\}_{k\in\mathbb{Z}}$ be an i. i. d. sequence with d. f. $(\nu_1 + \nu_2)((-\infty, x])/(\mu_1 + \mu_2)$. Since $\mathbb{E}|Z_1|^\delta < \infty$ and $\sum_{k=-\infty}^{\infty} |c_k|^\delta < \infty$ by (1.1.8) and monotone convergence we have

$$\mathbb{E}\left[\sum_{k=-\infty}^{\infty} |c_k Z_k|^\delta\right] = \mathbb{E}|Z_1|^\delta \sum_{k=-\infty}^{\infty} |c_k|^\delta < \infty.$$

Thus $\sum_{k=-\infty}^{\infty} |c_k Z_k|^\delta < \infty$ a. s.. By Jensen's-inequality for sequences we obtain $\sum_{k=-\infty}^{\infty} |c_k Z_k| \leq (\sum_{k=-\infty}^{\infty} |c_k Z_k|^\delta)^{1/\delta} < \infty$. Thus $\left|\sum_{k=-\infty}^{\infty} c_k Z_k\right| < \infty$ a. s.. Applying the three-series theorem it is necessary that $\sum_{k=-\infty}^{\infty} \mathbb{P}(|c_k Z_k| > 1) < \infty$. Thus we obtain

$$\begin{aligned}
&\int_{f(s)\neq 0} \nu\left(-\infty, -\frac{1}{|f(s)|}\right) + \nu\left(\frac{1}{|f(s)|}, \infty\right) ds \\
&\quad \leq \int_{\overline{f}(s)\neq 0} \nu\left(-\infty, -\frac{1}{\overline{f}(s)}\right) + \nu\left(\frac{1}{\overline{f}(s)}, \infty\right) ds \\
&\quad = (\mu_1 + \mu_2) \sum_{k=-\infty}^{\infty} \mathbb{P}(|c_k Z_k| > 1) < \infty.
\end{aligned}$$

Hence also $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \min\{1, |f(s)|^2 x^2\}\, \nu(dx)\, ds < \infty$. Thus by Proposition 1.1.6 the r. v. $Y(0)$ is well-defined and i. d..

For $u_1, \ldots, u_n \in \mathbb{R}$, $-\infty < t_1 < \ldots < t_n < \infty$, $n \in \mathbb{N}$, define the kernel function $\widetilde{f}(t) = \sum_{k=1}^{n} u_k f(t_k - t)$, $t \in \mathbb{R}$. Since $\widetilde{f}$ satisfies (1.1.7), by the first statement of the proof the r. v. s

$$\begin{aligned}
u_1 Y(t_1) + \ldots + u_n Y(t_n) &= \int_{-\infty}^{\infty} \widetilde{f}(s)\, dL(s), \\
u_1 Y(t_1 + h) + \ldots + u_n Y(t_n + h) &= \int_{-\infty}^{\infty} \widetilde{f}(s + h)\, dL(s)
\end{aligned}$$

are i. d. and their characteristic triplets coincide. Hence by the Cramér-Wold device Y is stationary. Similarly, $\{\widehat{Y}(t)\}_{t\in\mathbb{R}}$ is stationary and Y and $\widehat{Y}$ have the same finite dimensional distributions. □

Remark 1.1.8
(a) Equivalent conditions to (1.1.7) and (1.1.8) can be found in Balkema and de Haan [10].

(b) Let Y be the MA process given in (1.0.1) satisfying (K1) and (L1). Then Y is stationary and $Y(0)$ is i. d., with covariance function

$$\mathrm{Cov}(Y(0), Y(h)) = \int_{-\infty}^{\infty} f(s) f(s + h)\, ds.$$

Thus Y is a *short memory process* with $\int_{-\infty}^{\infty} \mathrm{Cov}(Y(0), Y(h))\, dh = (\int_{-\infty}^{\infty} f(t))^2 < \infty$. More about second order properties of MA processes can be found in the monograph of Daley and Vere-Jones [50], Chapter 8; especially Proposition 8.5. IV about representation of MA processes is to mention. The computation of the mean and variance of a Poisson shot noise process is known as Campbell's formulae, see Sato [138], Proposition 19.5.

(c) The assumption $\delta < 1$ is not a necessary condition for Y to be stationary; see Lemma 2.2.3 in the case $|L(1)| \in \mathcal{R}_\alpha$, $\alpha > 0$. □

Proof of Remark 1.1.8 *(b)*.
Taking Remark 1.1.5 *(b)* into account gives

$$\int_{|x|>1} |x|^2 \, \nu_Y(dx) = \int_{-\infty}^{\infty} \int_{|f(s)x|>1} |f(s)x|^2 \, \nu(dx) \, ds \le \int_{-\infty}^{\infty} f(s)^2 \, ds \int_{|x|>1/f^+} \nu(dx) < \infty$$

and thus $\mathbb{E}Y(0)^2 < \infty$. For $h \ge 0$ denote the characteristic function of $(Y(0), Y(h))$ by $\varphi_{(Y(0),Y(h))}$, with ψ given in (1.0.2). Then Proposition 2.6 of Rajput and Rosinski [121] yields

$$\varphi_{(Y(0),Y(h))}(u_1, u_2) = \exp\left(\int_{-\infty}^{\infty} \psi(u_1 f(-s) + u_2 f(h-s)) \, ds \right) \qquad \text{for } u_1, u_2 \in \mathbb{R}.$$

Thus we obtain on the one hand

$$\begin{aligned} \operatorname{Cov}(Y(0), Y(h)) &= \mathbb{E}Y(0)Y(h) = \frac{\partial^2}{\partial u_1 \partial u_2} \varphi_{(Y(0),Y(h))}(u_1, u_2) \Big|_{(u_1,u_2)=(0,0)} \\ &= \int_{-\infty}^{\infty} f(s) f(h+s) \, ds \end{aligned}$$

and on the other hand

$$\int_{-\infty}^{\infty} \operatorname{Cov}(Y(0), Y(h)) \, dh = \int_{-\infty}^{\infty} \left[f(s) \int_{-\infty}^{\infty} f(h+s) \, dh \right] ds = \left(\int_{-\infty}^{\infty} f(s) \right)^2 .$$

□

1.1.3 Examples

Example 1.1.9 (Poisson shot noise process)
An important special case of Y given in (1.0.1) is a MA process driven by a compound Poisson process. Consider a compound Poisson process $L = \{L(t)\}_{t\in\mathbb{R}}$ with

$$L(t) = \sum_{j=1}^{N(t)} Z_j \quad \text{and} \quad L(-t) = \sum_{j=1}^{-N(-t-)} Z_{-j} \quad \text{for } t \ge 0, \tag{1.1.10}$$

where $N = \{N(t)\}_{t\in\mathbb{R}}$ is a Poisson process on $\mathbb{R}$ with intensity $\mu > 0$ and jump times $\Gamma = \{\Gamma_k\}_{k\in\mathbb{Z}\setminus\{0\}}$, $\Gamma_{-1} < 0 < \Gamma_1$, $\Gamma_k < \Gamma_{k+1}$, for $k \in \mathbb{Z}\setminus\{-1, 0\}$. The process N is independent of the i. i. d. sequence $Z = \{Z_k\}_{k\in\mathbb{Z}}$. The Lévy measure ν of L is $\nu\,[x,\infty) = \mu\mathbb{P}(Z_1 > x)$ for $x \in \mathbb{R}$, so that the generating triplet is $(\int \kappa(x)\,\nu(dx), 0, \nu)$ (cf. Sato [138], Theorem 4.3). Let f be continuous in the interior of its support $[a, b], (-\infty, a], [a,\infty)$, respectively, $-\infty \le a < b \le \infty$, right and left continuous at the boundary points and satisfies (K1). Favorable for Y with kernel function f driven by the compound Poisson process L given in (1.1.10), which satisfies (L1), is the simple representation

$$Y(t) = \int_{t-b}^{t-a} f(t-s)\,dL(s) = \sum_{\substack{j=N(t-b)+1\\ j\neq 0}}^{N(t-a)} f(t-\Gamma_j)Z_j \quad \text{for } t \in \mathbb{R} \text{ a.s.}. \tag{1.1.11}$$

Notice that the left and the right hand side have a. s. càdlàg sample paths by dominated convergence and $\int_{-\infty}^{\infty} \sup_{0<s<1} |f(t+s)|^\delta\,dt < \infty$. We call Y given in (1.1.11) a *Poisson shot noise process*. If additionally f is positive and Z_1 has only support on $\mathbb{R}_+$, then we call (1.1.11) a *positive Poisson shot noise process*. On the other hand, if the kernel function f has only support on $\mathbb{R}_+$, then

$$Y(t) = \sum_{\substack{j=-\infty\\ j\neq 0}}^{N(t)} f(t-\Gamma_j)Z_j \quad \text{for } t \in \mathbb{R} \text{ a.s.}.$$

In this case with $f(0) > 0$, Y jumps if and only if N jumps. In particular, if f is non-increasing, $f(0) > 0$, and has only support on $\mathbb{R}_+$, then the positive Poisson shot noise process is non-increasing between successive jumps of L, and thus Y has a local supremum in t if and only if $t \in \Gamma$.

More general shot noise processes can be defined by replacing the Poisson process by a point process, e. g. a renewal process, and $\{Z_j f(t)\}_{t\ge 0}$ by $\{X_j(t)\}_{t\ge 0}$, $j \in \mathbb{N}$, an i. i. d. sequence of stochastic processes. In that case $Y(t) = \sum_{\Gamma_j \le t} X_j(t-\Gamma_j)$ for $t \ge 0$. An introduction in shot noise processes can be found in Bondesson [30,31] and Parzen [116]. Such methodological papers have been complemented by application based work.

We conclude this Example with a summary of the relevant literature. The simplest class of shot noise processes are those driven by a Poisson process; i. e. jumps have deterministic size. Hsing and Teugels [78] were among the first to analyze the point process behavior of such processes with bounded positive non-increasing kernel functions supported on a finite interval; this work was continued by Doney and

O'Brien [58] using kernel functions with support on $\mathbb{R}_+$.
Another simple class arises, when the kernel function has bounded memory and the driving Lévy process is a point process. Then there exists a $t_0 > 0$ such that $Y(0)$ is independent of $Y(t)$ for $t \geq t_0$. The point process behavior of this class with non-negative, non-increasing kernel functions has been investigated by Homble and McCormick: when the jump size distribution has regularly varying tails, has been studied in McCormick [108], gamma or logconcave densities in Homble and McCormick [75, 76].
The times Γ are interpreted as random time points, where shocks arrive with a certain memory modelled by the kernel function. Such models are used in various branches of stochastic modelling like bunching in traffic, computer failure times, earthquake aftershocks, applications to workload input models in teletraffic and risk theory. We refer to Klüppelberg et al. [90] and references therein.
Applications fall basically into two regimes, stationary and non-stationary situations occur. We indicate two special economic applications. Stationary models play an important role in finance; for instance as volatility and log-price models. Samorodnitsky [135, 136] introduces a quite sophisticated heavy-tailed stationary shot noise process for log-prices or log-exchange rates. Models with stationary increments play an important role in finance. An extension of the Black-Scholes model has been considered in Klüppelberg and Kühn [85], who suggested a Poisson shot noise process to extend the Black-Scholes model by allowing for information coming as shocks into the market. The kernel function can be chosen to model short or long range dependence effects in the market.
On the other hand shot noise processes can be useful in the context of insurance as models for delay in claim settlement. For $j \in \mathbb{N}$ the process $\{X_j(t)\}$ models the payoff function of an insurance claim, i. e. it is an increasing random process. Then Y is called an *explosive shot noise process*; see Klüppelberg and Mikosch [89]. □

Example 1.1.10 (Discrete-time MA process)
Let $\xi = \{\xi_k\}_{k\in\mathbb{Z}}$ be an i. i. d. sequence of r. v. s and $\{c_k\}_{k\in\mathbb{Z}}$ be a sequence of real constants. Then we call a stochastic process $Y = \{Y_n\}_{n\in\mathbb{Z}}$ with

$$Y_n = \sum_{k=-\infty}^{\infty} c_{n-k}\xi_k \quad \text{for } n \in \mathbb{Z} \tag{1.1.12}$$

a *discrete-time MA process* (cf. Brockwell and Davis [38]). If ξ is i. d., this model can be considered as a special case of Y in (1.0.1): choose $f(t) = \sum_{k=-\infty}^{\infty} c_k \,\mathbf{1}_{[k-1,k)}(t)$

for $t \in \mathbb{R}$. The continuous-time MA process Y at discrete-time points

$$Y(n) = \sum_{k=-\infty}^{\infty} c_{n-k}[L(k+1) - L(k)] \quad \text{for } n \in \mathbb{Z},$$

is a discrete-time MA process.

By Proposition 1.1.7 the process Y and hence also the discrete-time MA process $\{Y(n)\}_{n\in\mathbb{Z}}$ is stationary, if $\sum_{k=-\infty}^{\infty} |c_k|^\delta < \infty$ for some $\delta \in (0,1)$. For discrete-time MA processes this is a typical sufficient condition for stationarity (see for example Davis and Resnick [55]).

We conclude again with a summary of the relevant literature about extreme behavior of discrete-time MA processes: Davis and Resnick [55] investigate a subexponential noise in the maximum domain of attraction of the Gumbel distribution and in [54] models with regularly varying noise. The case of light tailed noise is studied by Lindner and Klüppelberg [86]. Finally we mention Rootzén [131] with noise, whose right tail decreases approximately as $Kx^{-\beta}\exp(-x^\alpha)$ for $x \to \infty$ for some $0 < \alpha < \infty$. In the case $0 < \alpha < 1$ the noise is subexponential. The structure of this chapter is very similar to his paper. $\square$

Example 1.1.11 (CARMA process)

For $q < p$, $q, p \in \mathbb{N}_0$ define polynomials

$$b(z) := b_0 + b_1 z + \ldots + b_q z^q \text{ and } a(z) := z^p + a_1 z^{p-1} + \ldots + a_p \text{ for } z \in \mathbb{C}, \quad (1.1.13)$$

such that the zeros of a, denoted by λ_j, $j = 1, \ldots, p$, have strictly negative real parts, i. e. $\Re(\lambda_j) < 0$. Assume that the Lévy measure ν of the Lévy process L satisfies $\int_{|x|>2} \log|x|\nu(dx) < \infty$. Then the MA process Y given by (1.0.1) with kernel function

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{it\omega}\frac{b(i\omega)}{a(i\omega)}\,d\omega \quad \text{for } t \in \mathbb{R},$$

is stationary and is called *continuous-time* ARMA(p,q) (CARMA(p,q)) process. It is possible to interpret Y as solution of the stochastic differential equation

$$a(D)Y(t) = b(D)DL(t) \quad \text{for } t \geq 0,$$

where D denotes the differential operator with respect to t. If condition (L1) holds, then Remark 1.1.5 *(b)* implies that $\int_{|x|>2} \log|x|\nu(dx) < \infty$. Furthermore, the kernel function has the property that $f(t) = 0$ for $t \leq 0$ and the representation

$f(t) = \sum_{j=1}^{p} c_j \lambda_j^{m_j} \exp(\lambda_j t)$ for $t > 0$, where $c_j \in \mathbb{R}$, $m_j \in \mathbb{N}$ is the multiplicity of the zero λ_j of a. If λ denotes the zero of a with the largest real part of all zeros of a and multiplicity m, then $|f(t)| \sim K t^m \exp(\Re(\lambda)\, t)$ for $t \to \infty$, $K > 0$. Hence f satisfies conditions (K1) and (K2).

A famous special case is the CARMA(1,0) process with representation

$$Y(t) = \int_{-\infty}^{t} \mathrm{e}^{\lambda(t-s)}\, dL(s) \quad \text{for } t \in \mathbb{R},\ \lambda < 0, \tag{1.1.14}$$

which is nothing else but an *Ornstein-Uhlenbeck* (OU) *process* driven by a Lévy process. If the zeros of a are distinct, then we can interpret any CARMA process as a linear combination of p OU processes, driven by the same Lévy process.

An extremal analysis of OU processes driven by a Lévy process with light-tails has been provided by Albin [1]. Such a process differs qualitatively not much from the classical OU process driven by a Brownian motion. So again, it behaves more or less like i. i. d. random variables with d. f. $\mathbb{P}(\sup_{0\le t\le 1} Y(t) \le x)$.

Barndorff-Nielsen and Shephard [15, 14] use positive OU processes as a model for stochastic volatility, which is considered in Chapter 3. See Brockwell [35, 36] for a more general account on CARMA processes and their applications to financial time series. As empirical volatility often exhibits long memory, Barndorff-Nielsen and Shephard [13] extend their model by exactly the above-mentioned idea of summing up p different OU models driven by independent Lévy processes, in contrast to the same Lévy process by a CARMA process (Example 2.2.10). It can be shown that for $p \to \infty$ the limit process has the long memory property.

Brockwell and Marquardt [39], on the other hand, extends the class of CARMA processes to fractionally integrated CARMA (FICARMA) processes (Example 2.2.9), which also exhibit the long memory property. In Anh et al. [8] further models for continuous-time processes with long memory can be found. Such processes, however, fall out of the framework of this chapter by Remark 1.1.8 *(b)*, but are included in Chapter 2 about subexponential mixed MA processes in the domain of attraction of the Fréchet distribution. □

Example 1.1.12 (Stochastic differential delay equation)
Let $f : [-r, \infty) \to \mathbb{R}$ for some $r \geq 0$ be absolutely continuous on $[0, \infty)$ and the solution of the differential delay equation

$$\begin{aligned} df(t) &= \int_{[-r,0]} f(t+s)\, \vartheta(ds)\, dt \quad \text{ for Lebesgue almost all } t \geq 0, \\ f(t) &= 0 \quad \text{ for } t \in [-r, 0) \quad \text{ and } \quad f(0) = 1. \end{aligned}$$

Moreover, let ϑ be a finite signed measure on $[-r, 0]$, and we assume that

$$\max \left\{ \Re(\lambda) \,:\, \lambda \in \mathbb{C},\, \lambda - \int_{[-r,0]} \mathrm{e}^{\lambda s}\, \vartheta(ds) = 0 \right\} < 0.$$

Suppose further that the Lévy measure ν of the Lévy process L satisfies the condition $\int_{|x|>2} \log |x| \nu(dx) < \infty$. Then the MA process (1.0.1) with kernel function f is stationary (Gushchin and Küchler [74]). Since it is the solution of

$$dY(t) = \left(\int_{[-r,0]} Y(t+s)\, \vartheta(ds) \right) dt + dL(t) \quad \text{ for } t \geq 0$$

it is called affine *stochastic differential delay equation* (SDDE). If condition (L1) holds then, by the same argument as for CARMA processes, the assumption

$$\int_{|x|>2} \log |x| \nu(dx) < \infty$$

is satisfied.
Since $|f(t)| \leq C\mathrm{e}^{-\gamma t}$ for $t \in \mathbb{R}$ and some γ, $C > 0$, it satisfies also conditions (K1) and (K2).

A typical example for a weight measure is $\vartheta = \lambda\, \varepsilon_0$ for some $\lambda < 0$, with ε_0 the Dirac measure in zero. Then Y is again an OU process. Weight measures ϑ with exponential densities $\vartheta' = -\beta \mathrm{e}^{\alpha s}$ for α, $\beta \in \mathbb{R}$ can be found in Reiß [122].

Delay equations play an important role in biology, for example, in population models with the natural time delay due to a pregnancy period, or age dependent birth and death models. The regulation of water temperature of a shower is also a famous example. We refer to McDonald [109]. Other examples of delay equations can be found in Mackey [102] and Benhabib and Rustichini [22]. □

1.1.4 Extreme value theory

In this section we summarize some results on extreme value theory, where we restrict ourselves to the situation of subexponential distributions. All these are standard results in classical extreme value theory and can be found in many textbooks; we refer here to Embrechts et al. [60] or Leadbetter et al. [95].

Definition 1.1.13
A d. f. F is in the maximum domain of attraction of a non-degenerate d. f. G, we write $F \in \mathrm{MDA}(G)$, if there exist normalizing constants $a_n > 0$, $b_n \in \mathbb{R}$ such that for an i. i. d. sequence $\{Z_k\}_{k\in\mathbb{Z}}$ with d. f. F

$$\lim_{n\to\infty} \mathbb{P}(a_n^{-1}(\max_{k=1,\dots,n} Z_k - b_n) \le x) = G(x) \qquad \text{for } x \in \mathbb{R}\,. \tag{1.1.15}$$

We also write $Z_1 \in \mathrm{MDA}(G)$.

If $a_n > 0$, $b_n \in \mathbb{R}$, then, by setting $u_n = a_n x + b_n$ for $x \in \mathbb{R}$, relation (1.1.15) is equivalent to

$$\lim_{n\to\infty} n\overline{F}(u_n) = -\log G(x) \qquad \text{for } x \in \mathbb{R}\,. \tag{1.1.16}$$

The d. f. G (up to location and scale change) is by the Fisher and Tippett Theorem either a Fréchet d. f. $\Phi_\alpha(x) = \exp(-x^{\alpha})$ for $x > 0$ ($\alpha > 0$), a Weibull d. f. $\Psi_\alpha(x) = \exp(-(-x)^{\alpha})$ for $x < 0$ ($\alpha > 0$), or a Gumbel d. f. $\Lambda(x) = \exp(-\mathrm{e}^{-x})$ for $x \in \mathbb{R}$. If $F \in \mathcal{S}$ its support is unbounded above, hence G is either a Fréchet or Gumbel d. f.. The fact that subexponential d. f. may belong to $\mathrm{MDA}(\Phi_\alpha)$ or $\mathrm{MDA}(\Lambda)$ has consequences when studying extremal events. It is well-known that $F \in \mathrm{MDA}(\Phi_\alpha)$ if and only if $\overline{F} \in \mathcal{R}_{-\alpha}$.
In this chapter we concentrate on $F \in \mathrm{MDA}(\Lambda) \cap \mathcal{S}$; in that case $\overline{F} \in \mathcal{R}_{-\infty}$. For conditions and examples of such d. f. we refer to Goldie and Resnick [71] and Example 1.1.4.
We shall also use the following representation for $F \in \mathrm{MDA}(\Lambda) \cap \mathcal{S}$, which is a generalization of Karamata's theorem:

$$\overline{F}(x) = c(x) \exp\left[-\int_0^x \frac{1}{\omega(u)} du\right] \text{ for } x > 0,$$

where $c : \mathbb{R}_+ \to (0, \infty)$ with $\lim_{x\to\infty} c(x) = c$ and $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ absolutely continuous with $\lim_{x\to\infty} \omega(x) = \infty$. Furthermore, it is possible to choose

$$b_n = \inf\left\{x \,:\, F(x) \ge 1 - \frac{1}{n}\right\} \qquad \text{and} \qquad a_n = \omega(b_n)\,. \tag{1.1.17}$$

Thus the normalizing constants satisfy

$$b_n \stackrel{n\to\infty}{\longrightarrow} \infty, \quad a_n \stackrel{n\to\infty}{\longrightarrow} \infty \quad \text{and} \quad b_n/a_n \stackrel{n\to\infty}{\longrightarrow} \infty.$$

For the continuous-time process Y we are concerned with quantities like $\sup_{0\le t\le T} Y(t)$ and we shall use a continuous version of the Poisson condition (1.1.16). Since also $\lim_{T\to\infty} T\overline{F}(u_{\lfloor T\rfloor}) = -\log G(x)$, the discrete sequence u_n in (1.1.16) can be replaced by the continuous sequence $u_T := u_{\lfloor T\rfloor}$ with a_T and b_T chosen analogously.

As an example, consider a compound Poisson process L given as in (1.1.10) with generic jump size Z_1. Since the Lévy measure of L is $\nu[x,\infty) = \mu\mathbb{P}(Z_1 > x)$ for $x \in \mathbb{R}$, Proposition 1.1.2 *(v)* gives $L(1) \in \mathcal{S}$ if and only if $Z_1 \in \mathcal{S}$, and in that case,

$$\mathbb{P}(L(1) > x) \sim \mu\mathbb{P}(Z_1 > x) \quad \text{for } x \to \infty.$$

Hence the normalizing constants $a_T > 0$, $b_T \in \mathbb{R}$ of $L(1)$ are connected to the normalizing constants $\widetilde{a}_T, \widetilde{b}_T \in \mathbb{R}$ of Z_1 by $\widetilde{a}_T = a_{T/\mu}$ and $\widetilde{b}_T = b_{T/\mu}$. By defining $u_T = a_T x + b_T$, x and $\widetilde{u}_T = \widetilde{a}_T x + \widetilde{b}_T$ for $x \in \mathbb{R}$,

$$\lim_{T\to\infty} T\mathbb{P}(L(1) > u_T) = -\log G(x) \iff \lim_{T\to\infty} T\mathbb{P}(Z_1 > \widetilde{u}_T) = -\log G(x) \quad (1.1.18)$$

holds.

1.2 Results on marked point processes

In this section we begin our investigation on the extremal behavior of processes, not necessarily stationary, in $\mathcal{S} \cap \text{MDA}(\Lambda)$. As is intuitively clear, extremes of a continuous-time MA process driven by any subexponential Lévy process are caused by large jumps of the Lévy process in cooperation with extremes of the kernel function. We shall model this by a discrete-time skeleton $\{Y(t_n)\}_{n\in\mathbb{N}}$ of Y given by a point process, where high threshold exceedances of the Lévy process in combination with extremes of the kernel function happen, which create excursions over the high threshold, which can be modelled as marks on the exceedance points.

Concerning the discrete-time skeleton our approach shows similarity to Rootzén [131], where extremes of a discrete-time MA process as given in (1.1.12), with a specific model for the noise were investigated. That paper concentrates on noises of a subexponential model class given by $\mathbb{P}(\xi_k > x) \sim Kx^{-\beta} \exp(-x^\alpha)$ for $x \to \infty$, $\beta \in \mathbb{R}$, $0 < \alpha < \infty$, where $0 < \alpha < 1$ is a subclass of the extended heavy-tailed Weibull distribution, Example 1.1.4.

Consider the discrete-time MA process (1.1.12). An extreme value of ξ among the ξ_k determines the behavior of Y such that Y_n behaves roughly like $c_{n-k}\xi_k$. For models with right tail in $\mathcal{R}_{-\alpha}$, $\alpha > 0$, extremes are more pronounced than in $\mathcal{S} \cap \mathrm{MDA}(\Lambda)$. In the different domains of attractions this affects excursions above a high threshold in the following sense. If $\xi_k \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$, then the right tail of ξ_k is rapidly varying, implies

$$\lim_{T\to\infty} \mathbb{P}(c_{n-k}\xi_k > u_T | c^+\xi_k > u_T) = \mathbf{1}_{c_{n-k}=c^+} .$$

This is in contrast to the right tail of ξ_k belonging $\mathcal{R}_{-\alpha}$ resulting in

$$\lim_{T\to\infty} \mathbb{P}(c_{n-k}\xi_k > u_T | c^+\xi_k > u_T) = \left(c_{n-k}/c^+\right)^\alpha .$$

This gives a precise description of excursions in both subexponential models. For models in $\mathcal{S} \cap \mathrm{MDA}(\Lambda)$ high threshold excursions depend only on the times, where extremes of f occur; for models in $\mathcal{R}_{-\alpha}$ high threshold excursions depend on the whole kernel function. Thus in both maximum domains of attractions extremes of Y are caused by extremes of c and ξ but the length of an excursion above a high threshold is different.

In Section 1.2.1 we shall show under a general point process setup that extremes of Y_n are caused by extremes of the subexponential ξ. In Section 1.2.2 the behavior of the normalized process Y is studied whenever an exceedance occurs. Finally, we summarize the results of Section 1.2.1 and 1.2.2 by incorporating both effects into marked point processes in Section 1.2.3. Throughout, we continue with the example of a discrete-time MA process providing a good intuition. We then apply our theory in Section 1.4.1 and 1.4.2 to describe the extremal behavior of the continuous-time process Y given in (1.0.1) at a properly chosen discrete-time skeleton. This yields a full description of the extreme behavior of Y.

1.2.1 Point processes

We follow Resnick [124, 125] and introduce point processes to describe the extremal behavior precisely. Let S denote the locally compact and separable Hausdorff space

$[0, \infty) \times \mathbb{R}$ with the Borel σ-field $\mathcal{B}(S)$ and $M_P(S)$ denotes the class of *point measures* (integer-valued Radon measures) on S provided with the metric ρ that generates the topology of vague convergence (for every $f \in C_c(S)$ the map from $M_P(S) \to \mathbb{R}$ with $\mu \mapsto \int f \, d\mu$ is continuous). A *Radon measure* on a locally compact and separable Hausdorff space is a measure, which is finite on compact sets. A measure of the form $\sum_{k \in I} \varepsilon_{x_k}$, where $x_k \in S$, I is at most countable and ε_{x_k} denotes the Dirac measure in x_k, is a point measure. The space $(M_P(S), \rho)$ is a complete and separable metric space (cf. Bauer [21], Theorem 31.5) provided with the Borel σ-field $\mathcal{M}_P(S)$. A *point process* in S is a random element in $(M_P(S), \mathcal{M}_P(S))$, i. e. a measurable map from a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ into $(M_P(S), \mathcal{M}_P(S))$. A typical example in extreme value theory for a point process is a Poisson random measure. Given a Radon measure ϑ on $\mathcal{B}(S)$, a point process κ is called *Poisson random measure* with mean measure (or intensity measure) ϑ, denoted by PRM(ϑ), if $\kappa(A)$ is Poisson distributed with mean $\vartheta(A)$ for every $A \in \mathcal{B}(S)$ and if, for mutually disjoint sets $A_1, \dots, A_n \in \mathcal{B}(S)$, $n \in \mathbb{N}$, the r. v. s $\kappa(A_1), \dots, \kappa(A_n)$ are independent (independent increment property). More about point processes can be found in Daley and Vere-Jones [50] and Kallenberg [81]. We remark that $\overset{T\to\infty}{\Longrightarrow}$ denotes weak convergence.

The idea is now to consider a subexponential sequence and add a small r. v. to each r. v., where the meaning of a small r. v. is that its right tail decreases faster than that of the subexponential r. v.. Then this small r. v. has no influence on the point process behavior. A similar result also holds for regularly varying tails (see Lemma 2.4.4). Finally, in Proposition 1.2.5 we investigate the point process behavior of a sum of two subexponential sequences, which are tail-equivalent.

Theorem 1.2.1
Let $Z = \{Z_k\}_{k \in \mathbb{N}}$ be identically distributed r. v. s in $\mathcal{S} \cap \mathrm{MDA}(\Lambda)$ and $\theta = \{\theta_k\}_{k \in \mathbb{N}}$ be a sequence of r. v. s, where θ_k is independent of Z_k for $k \in \mathbb{N}$. Suppose $a_T > 0$, $b_T \in \mathbb{R}$ are constants such that

$$\lim_{T\to\infty} T\mathbb{P}(Z_1 > u_T) = \exp(-x)$$

for $u_T = a_T x + b_T$ with $x \in \mathbb{R}$ holds. Further assume that there exists a r. v. Θ such that for every $k \in \mathbb{N}$, $x \in \mathbb{R}$,

$$\mathbb{P}(\theta_k > x) \leq \mathbb{P}(\Theta > x) \qquad \text{and} \qquad \mathbb{P}(\Theta > x) = o(\mathbb{P}(Z_1 > x)) \quad \text{for } x \to \infty.$$

Denote by

$$\zeta_T = \sum_{k=1}^{\infty} \varepsilon_{\left(k/T, a_T^{-1}(Z_k + \theta_k - b_T)\right)} \qquad \text{and} \qquad \widetilde{\kappa}_T = \sum_{k=1}^{\infty} \varepsilon_{\left(k/T, a_T^{-1}(Z_k - b_T)\right)}$$

point processes in $M_P(S)$. *Suppose there exists a point process* κ *in* $M_P(S)$ *with* $\kappa([s,t)\times\{x\})=0$ *a. s. such that* $\widetilde{\kappa}_T \overset{T\to\infty}{\Longrightarrow} \kappa$. *Let* $I=[s,t)\times(x,\infty)\subseteq S$. *Then*

$$\lim_{T\to\infty} \mathbb{P}(\zeta_T(I)\neq\widetilde{\kappa}_T(I))=0.$$

In the following Corollary we see that the result is also valid under certain scalings of time.

Corollary 1.2.2

Suppose the assumptions of Theorem 1.2.1 hold. Let $\{\Gamma_k\}_{k\in\mathbb{N}}$ *be the jump times of a Poisson process with intensity* $\mu>0$. *Let* $\alpha\in\mathbb{R}$ *be arbitrary and* $s_k\in[\Gamma_{k-1}+\alpha,\Gamma_{k+1}+\alpha)$ *for* $k\in\mathbb{N}$, *setting* $\Gamma_0:=0$. *For* $T>0$ *denote by*

$$\widetilde{\kappa}_T=\sum_{k=1}^{\infty}\varepsilon_{\left(k/T,a_T^{-1}(Z_k-b_T)\right)} \qquad \text{and} \qquad \kappa_T=\sum_{k=1}^{\infty}\varepsilon_{\left((s_k\mu)/T,a_T^{-1}(Z_k+\theta_k-b_T)\right)}$$

point processes in $M_P(S)$. *Let* $I=[s,t)\times(x,\infty)\subseteq S$. *Then*

$$\lim_{T\to\infty}\mathbb{P}(\kappa_T(I)\neq\widetilde{\kappa}_T(I))=0$$

and in particular $\kappa_T \overset{T\to\infty}{\Longrightarrow} \kappa$. *Especially for* Z *i. i. d.* κ *is* PRM(ϑ) *with mean measure* $\vartheta(dt\times dx)=dt\,\times \mathrm{e}^{-x}\,dx$.

The main step of proving Theorem 1.2.1 is the following Lemma.

Lemma 1.2.3

Let $Z\in\mathcal{S}\cap\mathrm{MDA}(\Lambda)$ *be independent of the r. v. s* θ *and* X. *Suppose there exist constants* $a_T>0$, $b_T\in\mathbb{R}$, *such that for* $u_T=a_Tx+b_T$ *with* $x\in\mathbb{R}$,

$$\lim_{T\to\infty} T\mathbb{P}(Z>u_T)=\exp(-x). \tag{1.2.1}$$

For $\epsilon>0$ *define* $v_T=a_T\epsilon$.

(a) *Suppose* $\mathbb{P}(\theta>x)=o(\mathbb{P}(Z>x))$ *for* $x\to\infty$. *Then*

$$\lim_{T\to\infty} T\mathbb{P}(\theta+Z>u_T, Z\leq u_T-v_T) = 0, \tag{1.2.2}$$

$$\lim_{A\uparrow\infty}\lim_{T\to\infty} T\mathbb{P}\left(\theta+Z>u_T,|Z-u_T|>a_TA\right) = 0. \tag{1.2.3}$$

(b) *Then*

$$\lim_{T\to\infty} T\mathbb{P}(\theta+Z\leq u_T, Z>u_T+v_T)=0.$$

(c) *Suppose* $\mathbb{P}(X>x)\sim q\,\mathbb{P}(Z>x)$ *for* $x\to\infty$ *and* $q>0$. *Then*

$$\lim_{T\to\infty} T\mathbb{P}(X+Z>u_T, X\leq u_T-v_T, Z\leq u_T-v_T)=0.$$

Proof.
Denote by F_Z, F_θ and F_X the d. f. s of Z, θ and X, respectively.

(a) Note that $u_T \to \infty$, $v_T \to \infty$ and also $u_T/2 - v_T = (x/2-\epsilon)a_T + b_T/2 \to \infty$ for $T \to \infty$. Hence, we can assume that $u_T/2 < u_T - v_T$. Now, suppose for the moment that for $T \to \infty$,

$$\int_{u_T/2}^{u_T-v_T} \overline{F}_\theta(u_T-y)F_Z(dy) = o(\overline{F}_Z(u_T)), \tag{1.2.4}$$

$$\int_{-\infty}^{u_T/2} \overline{F}_\theta(u_T-y)F_Z(dy) = o(\overline{F}_Z(u_T)), \tag{1.2.5}$$

$$\overline{F}_Z\,(u_T/2)\,\overline{F}_\theta\,(u_T/2) = o(\overline{F}_Z(u_T)). \tag{1.2.6}$$

Then we estimate and obtain for $T \to \infty$,

$$\mathbb{P}\,(\theta+Z>u_T, Z\le u_T-v_T, \theta\le u_T/2) \le \int_{u_T/2}^{u_T-v_T} \overline{F}_\theta(u_T-y)F_Z(dy) = o(\overline{F}_Z(u_T))$$

and

$$\mathbb{P}\,(\theta+Z>u_T, Z\le u_T/2) = \int_{-\infty}^{u_T/2} \overline{F}_\theta(u_T-y)F_Z(dy) = o(\overline{F}_Z(u_T)).$$

Then

$$\begin{aligned}
&\mathbb{P}(\theta+Z>u_T, Z\le u_T-v_T)\\
&\le \mathbb{P}\,(\theta+Z>u_T, Z\le u_T-v_T, \theta\le u_T/2)\\
&\quad +\mathbb{P}\,(\theta+Z>u_T, Z\le u_T/2) + \mathbb{P}\,(Z>u_T/2, \theta>u_T/2)\\
&= o(\overline{F}_Z(u_T)) \quad \text{for } T\to\infty.
\end{aligned}$$

Applying (1.2.1) yields (1.2.2). On the other hand, we estimate

$$\begin{aligned}
&\frac{\mathbb{P}(\theta+Z>u_T, |Z-u_T|>a_TA)}{\mathbb{P}(Z>u_T)}\\
&= \int_{-\infty}^{u_T-a_TA} \frac{\overline{F}_\theta(u_T-y)}{\overline{F}_Z(u_T)}F_Z(dy) + \int_{u_T+a_TA}^{\infty} \frac{\overline{F}_\theta(u_T-y)}{\overline{F}_Z(u_T)}F_Z(dy)\\
&\le \sup_{z>a_TA} \frac{\overline{F}_\theta(z)}{\overline{F}_Z(z)} \frac{\overline{F_Z^{2*}}(u_T)}{\overline{F}_Z(u_T)} + \frac{\overline{F}_Z(u_T+a_TA)}{\overline{F}_Z(u_T)}.
\end{aligned} \tag{1.2.7}$$

For the first summand in (1.2.7) the assumption in *(a)* and the fact that $u_T \to \infty$, $a_T \to \infty$ for $T \to \infty$ gives

$$\lim_{T\to\infty} \sup_{z>a_TA} \frac{\overline{F}_\theta(z)}{\overline{F}_Z(z)} \frac{\overline{F_Z^{2*}}(u_T)}{\overline{F}_Z(u_T)} = 0. \tag{1.2.8}$$

Applying (1.2.1) again gives for the second summand in (1.2.7)

$$\lim_{T\to\infty}\frac{\overline{F}_Z(u_T+a_TA)}{\overline{F}_Z(u_T)}=\lim_{T\to\infty}\frac{\overline{F}_Z\left(a_T(x+A)+b_T\right)}{\overline{F}_Z\left(a_Tx+b_T\right)}=\frac{\exp(-x-A)}{\exp(-x)}\overset{A\to\infty}{\longrightarrow}0. \quad (1.2.9)$$

The result (1.2.3) follows then by (1.2.7)-(1.2.9).

Next we prove (1.2.4)-(1.2.6). By the same argument as used for (1.2.8) and the fact $u_T,\, v_T\to\infty$ for $T\to\infty$ we obtain (1.2.4):

$$\int_{u_T/2}^{u_T-v_T}\frac{\overline{F}_\theta(u_T-y)}{\overline{F}_Z(u_T)}F_Z(dy)\le\sup_{z\ge v_T}\frac{\overline{F}_\theta(z)}{\overline{F}_Z(z)}\frac{\overline{F_Z^{2*}}(u_T)}{\overline{F}_Z(u_T)}\overset{T\to\infty}{\longrightarrow}0.$$

Similarly, we obtain (1.2.5). As $u_T\to\infty$ for $T\to\infty$, we have

$$\int_{-\infty}^{u_T/2}\frac{\overline{F}_\theta(u_T-y)}{\overline{F}_Z(u_T)}F_Z(dy)\quad=\quad\sup_{z\ge u_T/2}\frac{\overline{F}_\theta(z)}{\overline{F}_Z(z)}\frac{\overline{F_Z^{2*}}(u_T)}{\overline{F}_Z(u_T)}\overset{T\to\infty}{\longrightarrow}0.$$

Finally, (1.2.6) follows from Proposition 1.1.2 *(i)*, which gives

$$0\le\lim_{T\to\infty}\frac{\overline{F}_\theta\left(u_T/2\right)\overline{F}_Z\left(u_T/2\right)}{\overline{F}_Z(u_T)}=\lim_{T\to\infty}\frac{\overline{F}_\theta\left(u_T/2\right)}{\overline{F}_Z\left(u_T/2\right)}\lim_{T\to\infty}\frac{\overline{F}_Z\left(u_T/2\right)\overline{F}_Z\left(u_T/2\right)}{\overline{F}_Z(u_T)}=0.$$

Statement (1.2.6) also holds, if θ and Z are tail-equivalent.

(b) We have

$$\lim_{T\to\infty}T\mathbb{P}(Z>u_T+v_T,\theta+Z\le u_T)\le\lim_{T\to\infty}T\mathbb{P}(Z>u_T+v_T)\mathbb{P}(\theta\le -v_T)=0.$$

(c) Since $F_X\in\mathcal{L}$, we know that $\overline{F}_X(u_T-y)/\overline{F}_Z(u_T)\to q$ for $T\to\infty$ locally uniformly in y. Moreover, by Proposition 1.1.2 *(iii)*

$$\lim_{T\to\infty}\int_{-\infty}^{u_T/2}\frac{\overline{F}_X(u_T-y)}{\overline{F}_Z(u_T)}F_Z(dy)=q.$$

Thus by Pratt's Lemma (Gänssler and Stute [68], Lemma 1.11.16)

$$\begin{aligned}
&\lim_{T\to\infty}\frac{\mathbb{P}\left(X+Z>u_T,X\le u_T-v_T,Z\le u_T/2\right)}{\mathbb{P}(Z>u_T)}\\
&=\lim_{T\to\infty}\int_{v_T}^{u_T/2}\frac{\overline{F}_X(u_T-y)}{\overline{F}_Z(u_T)}F_Z(dy)\\
&=\int_{-\infty}^{\infty}\lim_{T\to\infty}\frac{\overline{F}_X(u_T-y)}{\overline{F}_Z(u_T)}1_{[v_T,\frac{u_T}{2}]}(y)F_Z(dy)=0.
\end{aligned}$$

By symmetry also $\mathbb{P}(X+Z>u_T, X\le u_T/2, Z\le u_T-v_T)=o(\overline{F}_Z(u_T))$ for $T\to\infty$. Thus by (1.2.6)

$$\begin{aligned}
&\mathbb{P}\left(X+Z>u_T, X\le u_T-v_T, Z\le u_T-v_T\right)\\
&\quad\le \mathbb{P}\left(X+Z>u_T, X\le u_T-v_T, Z\le u_T/2\right)\\
&\qquad +\mathbb{P}\left(X+Z>u_T, X\le u_T/2, Z\le u_T-v_T\right)+\mathbb{P}\left(X>u_T/2\right)\mathbb{P}\left(Z>u_T/2\right)\\
&\quad= o(\overline{F}_Z(u_T)) \quad \text{for } T\to\infty.
\end{aligned}$$

□

Proof of Theorem 1.2.1.
Let $\epsilon>0$ be arbitrary. Write $I_\epsilon=[s,t)\times(x-\epsilon,x+\epsilon]$. Then

$$\begin{aligned}
\{\zeta_T(I)\neq\widetilde{\kappa}_T(I)\} \;\subseteq\; & \bigcup_{k\in[Ts,Tt)}\{\theta_k+Z_k>u_T, Z_k\le u_T-v_T\}\\
& \cup\bigcup_{k\in[Ts,Tt)}\{\theta_k+Z_k\le u_T, Z_k>u_T+v_T\}\cup\{\widetilde{\kappa}_T(I_\epsilon)>0\}.
\end{aligned}$$

Hence by Lemma 1.2.3 *(a)* and *(b)*,

$$\begin{aligned}
&\mathbb{P}\left(\zeta_T(I)\neq\widetilde{\kappa}_T(I)\right)\\
&\quad\le T(t-s)\mathbb{P}(\Theta+Z_1>u_T, Z_1\le u_T-v_T)\\
&\qquad +T(t-s)\mathbb{P}(\Theta+Z_1\le u_T, Z_1>u_T+v_T)+\mathbb{P}(\widetilde{\kappa}_T(I_\epsilon)>0)\\
&\quad\overset{T\to\infty}{\longrightarrow} 0+\mathbb{P}(\kappa(I_\epsilon)>0)\overset{\epsilon\to 0}{\longrightarrow} 0.
\end{aligned}$$

□

To prove Corollary 1.2.2 we modify an argument of Hsing and Teugels [78]; see their proof of Theorem 4.2 and Lemma 2.1. We include the proof here, since it keeps the thesis self-contained.

Lemma 1.2.4
Let $\{Y_k\}_{k\in\mathbb{N}}$ be a sequence of r. v. s and assume there exists a r. v. Y such that for some $x_0\in\mathbb{R}$,

$$\mathbb{P}(Y_k>x)\le\mathbb{P}(Y>x)\quad \textit{for } x\ge x_0.$$

Suppose there exist constants $a_T > 0$, $b_T \in \mathbb{R}$ *such that* $u_T = a_T x + b_T \to \infty$ *for* $x \in \mathbb{R}$ *and* $\mathbb{P}(Y > u_T) = O(T^{-1})$ *for* $T \to \infty$. *Assume* $\{\Gamma_k\}_{k\in\mathbb{N}}$ *are the jump times of a Poisson process* $N = \{N(t)\}_{t\geq 0}$ *with intensity* $\mu > 0$. *Let* $\alpha \in \mathbb{R}$ *be arbitrary and* $s_k \in [\Gamma_{k-1} + \alpha, \Gamma_{k+1} + \alpha)$ *for* $k \in \mathbb{N}$, *setting* $\Gamma_0 := 0$. *For* $T > 0$ *denote by*

$$\zeta_T = \sum_{k=1}^{\infty} \varepsilon_{\left(k/T, a_T^{-1}(Y_k - b_T)\right)} \qquad \text{and} \qquad \kappa_T = \sum_{k=1}^{\infty} \varepsilon_{\left((s_k\mu)/T, a_T^{-1}(Y_k - b_T)\right)}$$

point processes in $M_P(S)$. *Let* $I = [s,t) \times (x, \infty) \subseteq S$. *Then*

$$\lim_{T\to\infty} \mathbb{P}(\zeta_T(I) \neq \kappa_T(I)) = 0.$$

Proof.
Without loss of generality we assume $\alpha \geq 0$. For $T > 0$ define the point processes

$$\eta_T = \sum_{k=0}^{\infty} \varepsilon_{\left((k+\alpha\mu)/T, a_T^{-1}(Y_k - b_T)\right)} \quad \text{and} \quad \widetilde{\eta}_T = \sum_{k=0}^{\infty} \varepsilon_{\left((\Gamma_k\mu+\alpha\mu)/T, a_T^{-1}(Y_k - b_T)\right)}.$$

Then

$$\begin{aligned} \mathbb{P}(\zeta_T(I) \neq \eta_T(I)) &\leq \sum_{Ts-\alpha\mu \leq k \leq Ts} \mathbb{P}(Y_k > u_T) + \sum_{Tt-\alpha\mu\leq k\leq Tt} \mathbb{P}(Y_k > u_T) \\ &\leq 2\alpha\mu\mathbb{P}(Y > u_T) \overset{T\to\infty}{\longrightarrow} 0. \end{aligned} \tag{1.2.10}$$

Next, define the events

$$\begin{aligned} A_T &= \left\{ \left| N\left(\frac{sT-\alpha\mu}{\mu}\right) - (sT - \alpha\mu) \right| > \epsilon T \right\}, \\ B_T &= \left\{ \left| N\left(\frac{tT-\alpha\mu}{\mu}\right) - (tT - \alpha\mu) \right| > \epsilon T \right\}. \end{aligned}$$

The LLN (Mikosch [111], Theorem 2.2.4) gives $\lim_{T\to\infty} \mathbb{P}(A_T) = \lim_{T\to\infty} \mathbb{P}(B_T) = 0$. From this

$$\mathbb{P}(\eta_T(I) \neq \widetilde{\eta}_T(I)) \leq \mathbb{P}(A_T) + \mathbb{P}(B_T) + 2\epsilon T\mathbb{P}(Y > u_T) \overset{T\to\infty}{\longrightarrow} 0 \tag{1.2.11}$$

follows. If $s_k \leq tT \leq s_{k+1}$, then $\Gamma_{k-1} + \alpha \leq s_k \leq tT \leq s_{k+1} \leq \Gamma_{k+2} + \alpha$. Hence

$$\mathbb{P}(\widetilde{\eta}_T(I) \neq \kappa_T(I)) \leq 2\mathbb{P}(Y > u_T) + 2\mathbb{P}(Y > u_T) \overset{T\to\infty}{\longrightarrow} 0. \tag{1.2.12}$$

From (1.2.10)-(1.2.12) we conclude $\lim_{T\to\infty} \mathbb{P}(\zeta_T(I) \neq \kappa_T(I)) = 0$. □

Proof of Corollary 1.2.2.
A consequence of Theorem 1.2.1 is $\lim_{T\to\infty}\mathbb{P}(\zeta_T(I) \neq \widetilde{\kappa}_T(I)) = 0$. Applying Lemma 1.2.4 we conclude $\lim_{T\to\infty}\mathbb{P}(\zeta_T(I) \neq \kappa_T(I)) = 0$. Thus

$$\lim_{T\to\infty}\mathbb{P}(\widetilde{\kappa}_T(I) \neq \kappa_T(I)) = 0.$$

As $\widetilde{\kappa}_T \Longrightarrow \kappa$ for $T \to \infty$ and using Lemma 3.3 of Rootzén [131] we conclude $\kappa_T \Longrightarrow \kappa$ for $T \to \infty$.

In the case Z i. i. d, like in the proof of Theorem 4.2 of Hsing and Teugels [78] on the one hand

$$\mathbb{P}(\widetilde{\kappa}_{\lfloor T\rfloor}(I) \neq \widetilde{\kappa}_T(I)) \leq (tT - \lfloor tT\rfloor + sT - \lfloor sT\rfloor)\mathbb{P}(Z_1 > u_T) \stackrel{T\to\infty}{\longrightarrow} 0$$

and on the other hand $\widetilde{\kappa}_n \stackrel{n\to\infty}{\Longrightarrow} \kappa$. Thus $\widetilde{\kappa}_T \stackrel{T\to\infty}{\Longrightarrow} \kappa$ and $\kappa_T \stackrel{T\to\infty}{\Longrightarrow} \kappa$. □

The last Proposition 1.2.5 of this section regards point processes of independent subexponential sequences. It is a generalization of Goldie and Resnick [72], Theorem 2.3, who study the case of i. i. d. sequences.

Proposition 1.2.5
For $i = 1, 2$, let $\{Z_k^{(i)}\}_{k\in\mathbb{N}}$ be identically distributed r. v. s with $Z_k^{(i)} \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$ satisfying $Z_k^{(1)}$ is independent of $Z_k^{(2)}$ for $k \in \mathbb{N}$. Suppose there exist constants $a_T > 0$, $b_T \in \mathbb{R}$ such that for $u_T = a_T x + b_T$ with $x \in \mathbb{R}$,

$$\lim_{T\to\infty} T\mathbb{P}(Z_1^{(i)} > u_T) = K_i \exp(-x) \quad \text{for some } K_i > 0. \tag{1.2.13}$$

For $T > 0$ denote by

$$\kappa_T^{(i)} = \sum_{k=1}^{\infty} \varepsilon_{\left(k/T, a_T^{-1}(Z_k^{(i)} - b_T)\right)} \quad \text{and} \quad \kappa_T = \sum_{k=1}^{\infty} \varepsilon_{\left(k/T, a_T^{-1}(Z_k^{(1)} + Z_k^{(2)} - b_T)\right)}$$

point processes in $M_P(S)$, where $\kappa_T^{(1)} + \kappa_T^{(2)} \stackrel{T\to\infty}{\Longrightarrow} \kappa^{(1)} + \kappa_T^{(2)}$ for some point processes $\kappa^{(i)}$ with $\kappa^{(i)}([s,t) \times \{x\}) = 0$ a. s.. Suppose $I = [s,t) \times (x, \infty) \subseteq S$. Then

$$\lim_{T\to\infty}\mathbb{P}(\kappa_T(I) \neq \kappa_T^{(1)}(I) + \kappa_T^{(2)}(I)) = 0$$

and $\kappa_T \stackrel{T\to\infty}{\Longrightarrow} \kappa^{(1)} + \kappa^{(2)}$.

Proof.
For $\epsilon > 0$ set $v_T = a_T\epsilon$ and $I_\epsilon = [s,t) \times [x-\epsilon, x+\epsilon)$. Then

$$\begin{aligned}
&\{\kappa_T(I) \neq \kappa_T^{(1)}(I) + \kappa_T^{(2)}(I)\} \\
&\subseteq \{\kappa_T^{(1)}(I_\epsilon) > 0\} \cup \{\kappa_T^{(2)}(I_\epsilon) > 0\} \\
&\quad \bigcup_{k\in[Ts,Tt)} \{Z_k^{(1)} + Z_k^{(2)} > u_T, Z_k^{(1)} \leq u_T - v_T, Z_k^{(2)} \leq u_T - v_T\} \\
&\quad \cup\{Z_k^{(1)} + Z_k^{(2)} < u_T, Z_k^{(1)} > u_T + v_T\} \cup \{Z_k^{(1)} + Z_k^{(2)} < u_T, Z_k^{(2)} > u_T + v_T\} \\
&\quad \cup\{Z_k^{(1)} > u_T, Z_k^{(2)} > u_T\}.
\end{aligned} \tag{1.2.14}$$

So by Lemma 1.2.3 *(c)* we have

$$\begin{aligned}
&\mathbb{P}\left(\bigcup_{k\in[Ts,Tt)} \{Z_k^{(1)} + Z_k^{(2)} > u_T, Z_k^{(1)} \leq u_T - v_T, Z_k^{(2)} \leq u_T - v_T\}\right) \\
&\quad \leq T(t-s)\mathbb{P}(Z_1^{(1)} + Z_1^{(2)} > u_T, Z_1^{(1)} \leq u_T - v_T, Z_1^{(2)} \leq u_T - v_T) \overset{T\to\infty}{\longrightarrow} 0.
\end{aligned} \tag{1.2.15}$$

On the other hand, by (1.2.13)

$$\mathbb{P}\Big(\bigcup_{k\in[Ts,Tt)}\{Z_k^{(1)} > u_T, Z_k^{(2)} > u_T\}\Big) \leq T(t-s)\mathbb{P}(Z_1^{(1)} > u_T)\mathbb{P}(Z_1^{(2)} > u_T) \overset{T\to\infty}{\longrightarrow} 0. \tag{1.2.16}$$

Finally,

$$\begin{aligned}
&\mathbb{P}\left(\bigcup_{k\in[Ts,Tt)} \{Z_k^{(1)} + Z_k^{(2)} < u_T, Z_k^{(1)} > u_T + v_T\}\right) \\
&\quad \leq T(t-s)\mathbb{P}(Z_1^{(1)} + Z_1^{(2)} < u_T, Z_1^{(1)} > u_T + v_T) \\
&\quad \leq T(t-s)\mathbb{P}(Z_1^{(1)} > u_T + v_T, Z_1^{(2)} < -v_T) \\
&\quad = T(t-s)\ \mathbb{P}(Z_1^{(1)} > u_T + v_T)\mathbb{P}(Z_1^{(2)} < -v_T) \overset{T\to\infty}{\longrightarrow} 0.
\end{aligned} \tag{1.2.17}$$

Furthermore $\lim_{\epsilon\to 0}\lim_{T\to\infty}\mathbb{P}(\kappa_T^{(i)}(I_\epsilon) > 0) = 0$ for $i = 1,2$. Using additionally (1.2.14)-(1.2.17) yields

$$\lim_{T\to\infty} \mathbb{P}(\kappa_T(I) \neq \kappa_T^{(1)}(I) + \kappa_T^{(2)}(I)) = 0. \tag{1.2.18}$$

Thus by (1.2.18) and Rootzén [131], Lemma 3.3, also $\kappa_T \overset{T\to\infty}{\Longrightarrow} \kappa^{(1)} + \kappa^{(2)}$. □

Note $Z^{(1)}$, $Z^{(2)}$ and $\kappa^{(1)}$, $\kappa^{(2)}$ independent imply $\kappa_T^{(1)}$, $\kappa_T^{(2)}$ independent and thus we get from $\kappa_T^{(i)} \overset{T\to\infty}{\Longrightarrow} \kappa^{(i)}$, $i = 1,2$, also $\kappa_T^{(1)} + \kappa_T^{(2)} \overset{T\to\infty}{\Longrightarrow} \kappa^{(1)} + \kappa^{(2)}$.

Example 1.2.6 (Continuation of Example 1.1.10)
Assume $\xi_1 \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$ with $\lim_{x\to\infty} \mathbb{P}(\xi_1 < -x)/\mathbb{P}(\xi_1 > x) = (1-p)/p$, $p \in (0,1]$ and $\sum_{k=-\infty}^{\infty} |c_k|^\delta < \infty$ for some $\delta \in (0,1)$. Without loss of generality we assume

$$c_1 = \ldots = c_{P^{(1)}} = c^+, \qquad c_{-P^{(2)}+1} = \ldots = c_0 = -c^+$$

and $|c_k| < c^+ < \infty$ for $k \leq -P^{(2)}, k \geq P^{(1)}+1$, $P^{(1)} \in \mathbb{N}$, $P^{(2)} \in \mathbb{N}_0$. In the case $p = 0$ set $P^{(2)} := 0$. Let $a_T > 0$, $b_T \in \mathbb{R}$ and $u_T = a_T x + b_T$ for $x \in \mathbb{R}$ such that

$$\lim_{T\to\infty} T\mathbb{P}(c^+\xi_1 > u_T) = \exp(-x).$$

For $k \in \mathbb{Z}$ define the stationary processes

$$\overline{\xi}_k := -\xi_{k-P^{(2)}+1} - \ldots - \xi_k + \xi_{k+1} + \ldots + \xi_{k+P^{(1)}} \quad \text{and} \quad \theta_k := Y_k - c^+\overline{\xi}_k.$$

Let $\kappa^{(1)}$, $\kappa^{(2)}$ be independent PRM(ϑ_i), $i = 1,2$, with $\vartheta_1(dt \times dx) = dt \times \mathrm{e}^{-x}\, dx$, $\vartheta_2(dt \times dx) = dt \times (1-p)/p\, \mathrm{e}^{-x}\, dx$, respectively. By by similar arguments as in Proposition 1.2.5

$$\sum_{k=1}^{\infty} \varepsilon_{\left(\frac{k}{T}, a_T^{-1}(c^+\overline{\xi}_k - b_T)\right)} \overset{T\to\infty}{\Longrightarrow} P^{(1)}\kappa^{(1)} + P^{(2)}\kappa^{(2)}. \tag{1.2.19}$$

Furthermore $\mathbb{P}(\theta_k > x) = o(\mathbb{P}(c^+\overline{\xi}_k > x))$ and $\mathbb{P}(\theta_k < -x) = o(\mathbb{P}(c^+\overline{\xi}_k > x))$ for $x \to \infty$. Hence by Theorem 1.2.1

$$\mathbb{P}\left(\sum_{k=1}^{\infty} \varepsilon_{\left(\frac{k}{T}, a_T^{-1}(Y_k - b_T)\right)}(I) \neq \sum_{k=1}^{\infty} \varepsilon_{\left(\frac{k}{T}, a_T^{-1}(c^+\overline{\xi}_k - b_T)\right)}(I)\right) \overset{T\to\infty}{\longrightarrow} 0, \tag{1.2.20}$$

and by (1.2.19)

$$\sum_{k=1}^{\infty} \varepsilon_{\left(\frac{k}{T}, a_T^{-1}(Y_k - b_T)\right)} \overset{T\to\infty}{\Longrightarrow} P^{(1)}\kappa^{(1)} + P^{(2)}\kappa^{(2)}.$$

This result has been proven by Davis and Resnick [55], Theorem 3.3 with completely different methods. □

1.2.2 Marks

In the last section we have proven point process convergence of a discrete-time subexponential sequence. In the continuous-time setting, we choose the discrete-time points of the point process properly to capture large jumps of the Lévy process and extremes of the kernel function. Now, also the behavior of the continuous-time process between the discrete-time skeleton matters. The question arises how long the sample path of the continuous-time process stays on a high level, and how it reverts to its mean, if a high level exceedance at a certain time point occurs. The sample path behavior near high level excursions will be modelled by *marks*. The mark is the stochastic process Y under the conditional probability that $Y(\alpha) > u_T$. We shall see in Section 1.4 that our discrete-time skeleton captures fully the extremal behavior of Y.

Lemma 1.2.7
Let $Y = \{Y(t)\}_{t\in\mathbb{R}}$ be a stochastic process in $\mathbb{R}$, which is a. s. bounded on every compact set and has the decomposition

$$Y(t) = f(t)Z + \widetilde{Y}(t) \quad \text{for } t \in \mathbb{R},$$

where $f : \mathbb{R} \to \mathbb{R}$ is a deterministic function with $f^+ < \infty$ and $Z \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$ is a r. v. independent of the stochastic process $\widetilde{Y} = \{\widetilde{Y}(t)\}_{t\in\mathbb{R}}$, which is also a. s. bounded on every compact set. Assume furthermore that there exist constants $a_T > 0$, $b_T \in \mathbb{R}$ and $u_T = a_T x + b_T$ for $x \in \mathbb{R}$ such that

$$\lim_{T\to\infty} T\mathbb{P}(f^+ Z > u_T) = \exp(-x).$$

Let $\tau = f^+ Z + \theta$ with

$$\mathbb{P}(\theta > x) = o(\mathbb{P}(f^+ Z > x)) \quad \text{for } x \to \infty. \tag{1.2.21}$$

For $\epsilon > 0$ arbitrary the following assertions hold:

(a) Let $m > 0$ be fixed, then $\lim_{T\to\infty} \mathbb{P}\left(\sup_{-m\le t\le m} \left| \frac{Y(t)}{b_T} - \frac{f(t)}{f^+} \right| > \epsilon \,\middle|\, \tau > u_T \right) = 0.$

(b) Let $O = \{\alpha_1, \dots, \alpha_P\}$ be a finite set in $\mathbb{R}$ such that $f(t) = f^+$ for $t \in O$. For $y_1, \dots, y_P \in \mathbb{R}$, and $y = \max\{0, y_1, \dots, y_P\}$ we have

$$\lim_{T\to\infty} \mathbb{P}\left(Y(\alpha_1) > u_T + a_T y_1, \dots, Y(\alpha_P) > u_T + a_T y_P \middle| \tau > u_T \right) = \exp(-y).$$

(c) Let $|f(t)| < f^+$ and $\mathbb{P}(\widetilde{Y}(t) > x) = o(\mathbb{P}(f^+ Z > x))$ for $x \to \infty$. Then

$$\lim_{T\to\infty} \mathbb{P}\left(Y(t) > u_T + a_T y \middle| \tau > u_T \right) = 0 \quad \text{for } y \in \mathbb{R}.$$

Remark 1.2.8
(i) Let $\alpha \in \mathbb{R}$ with $f(\alpha) = f^+$, $\mathbb{P}(\widetilde{Y}(\alpha) > x) = o(\mathbb{P}(f^+Z > x))$ for $x \to \infty$ and $\tau = Y(\alpha)$, where Y and $\widetilde{Y}$ are a. s. bounded on every compact set. Then, Lemma 1.2.7 describes the sample path behavior of Y, if it has an exceedance over the threshold u_T at time point α. Let X_T for $T > 0$ be processes in some measurable metric space $(\mathbb{D}, \mathcal{D})$, where uniform convergence on compacta is sufficient for convergence. The process X_T with distribution

$$\mathbb{P}(X_T \in D) = \mathbb{P}(Y \in D | Y(\alpha) > u_T) \quad \text{for } D \in \mathcal{D}$$

is defined as a mark on Y. Lemma 1.2.7 *(a)* then states that the normalized marks converge weakly to the deterministic function f/f^+. Moreover, Lemma 1.2.7 *(b)* and *(c)* give the distributional limit of the *scaled excess over threshold.* For $P = 1$, the exponential limit in *(b)* corresponds to the limiting generalized Pareto distribution for scaled excesses in MDA(Λ).

(ii) If $f(\alpha\pm) = f^+$ or $f(\alpha_l\pm) = f^+$ for $l \in \{1, \ldots, P\}$ and $\widetilde{Y}$ is a. s. continuous in α or α_l, respectively, with $\mathbb{P}(\widetilde{Y}(\alpha) > x) = o(\mathbb{P}(f^+Z > x))$ for $x \to \infty$, then Lemma 1.2.7 also holds by replacing τ or $Y(\alpha_l)$ by $Y(\alpha\pm)$ or $Y(\alpha_l\pm)$, respectively. □

Proof of Lemma 1.2.7.
Let $\epsilon > 0$ be arbitrary.

(a) We decompose the probability:

$$\begin{aligned} &\mathbb{P}\left(\sup_{-m \le t \le m} \left| \frac{Y(t)}{b_T} - \frac{f(t)}{f^+} \right| > \epsilon \,\middle|\, \tau > u_T \right) \\ &= \mathbb{P}\left(\sup_{-m \le t \le m} \left| \frac{Y(t)}{b_T} - \frac{f(t)}{f^+} \right| > \epsilon, |f^+Z - u_T| > a_T A \,\middle|\, \tau > u_T \right) \\ &\quad + \mathbb{P}\left(\sup_{-m \le t \le m} \left| \frac{Y(t)}{b_T} - \frac{f(t)}{f^+} \right| > \epsilon, |f^+Z - u_T| \le a_T A \,\middle|\, \tau > u_T \right). \end{aligned} \tag{1.2.22}$$

The first term in (1.2.22) satisfies the inequality

$$\begin{aligned} &\mathbb{P}\left(\sup_{-m \le t \le m} \left| \frac{Y(t)}{b_T} - \frac{f(t)}{f^+} \right| > \epsilon, |f^+Z - u_T| > a_T A \,\middle|\, \tau > u_T \right) \\ &\le \frac{\mathbb{P}\left(|f^+Z - u_T| > a_T A, \tau > u_T\right)}{\mathbb{P}(\tau > u_T)}. \end{aligned} \tag{1.2.23}$$

Furthermore, by (1.2.21) and Proposition 1.1.2 *(iii)*,

$$\lim_{T\to\infty} T\mathbb{P}(\tau > u_T) = \lim_{T\to\infty} T\mathbb{P}(f^+Z + \theta > u_T) = \exp(-x). \tag{1.2.24}$$

Then, by using Lemma 1.2.3 *(a)* we conclude

$$\lim_{A\uparrow\infty}\lim_{T\to\infty} \frac{\mathbb{P}\left(|f^+Z - u_T| > a_T A, \tau > u_T\right)}{\mathbb{P}(\tau > u_T)} = 0. \tag{1.2.25}$$

For the second term in (1.2.22) we have

$$\begin{aligned}
&\mathbb{P}\left(\sup_{-m\le t\le m}\left|\frac{Y(t)}{b_T} - \frac{f(t)}{f^+}\right| > \epsilon, |f^+Z - u_T| \le a_T A\right)\\
&\quad\le \mathbb{P}\left(\sup_{-m\le t\le m}|Y(t) - f(t)Z| > b_T\epsilon - a_T(A+x), |f^+Z - u_T| \le a_T A\right)\\
&\quad= \mathbb{P}\left(\sup_{-m\le t\le m}|\widetilde{Y}(t)| > b_T\epsilon - a_T(A+x)\right)\mathbb{P}\left(\left|f^+Z - u_T\right| \le a_T A\right), \qquad (1.2.26)
\end{aligned}$$

where we used the independence of $\widetilde{Y}$ and Z in the last step. Furthermore, for $T \to \infty$, on the one hand $b_T\epsilon - a_T(A+x) \to \infty$ holds and on the other hand

$$\mathbb{P}\left(\left|f^+Z - u_T\right| \le a_T A\right) \le \mathbb{P}\left(f^+Z > u_T - a_T A\right) = O\left(T^{-1}\right). \tag{1.2.27}$$

Thus by (1.2.26) and (1.2.27) for $T \to \infty$,

$$\mathbb{P}\left(\sup_{-m\le t\le m}\left|\frac{Y(t)}{b_T} - \frac{f(t)}{f^+}\right| > \epsilon, |f^+Z - u_T| \le a_T A\right) = o\left(T^{-1}\right)$$

and by (1.2.24)

$$\mathbb{P}\left(\sup_{-m\le t\le m}\left|\frac{Y(t)}{b_T} - \frac{f(t)}{f^+}\right| > \epsilon, |f^+Z - u_T| \le a_T A \,\middle|\, \tau > u_T\right) \stackrel{T\to\infty}{\longrightarrow} 0. \tag{1.2.28}$$

Combining (1.2.22), (1.2.23), (1.2.25) and (1.2.28) yields the assertion.

(b) First we show

$$\lim_{T\to\infty} \mathbb{P}\left(\sup_{t\in O}|Y(t) - \tau| > a_T\epsilon \,\middle|\, \tau > u_T\right) = 0. \tag{1.2.29}$$

Define $v_T = a_T\epsilon$. We proceed as in *(a)* and decompose the probability

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\in O}|Y(t) - \tau| > a_T\epsilon \,\middle|\, \tau > u_T\right) \qquad (1.2.30)\\
&\quad= \mathbb{P}\left(\sup_{t\in O}|\widetilde{Y}(t) - \theta| > a_T\epsilon, f^+Z > u_T - v_T \,\middle|\, \tau > u_T\right)\\
&\qquad+\mathbb{P}\left(\sup_{t\in O}|\widetilde{Y}(t) - \theta| > a_T\epsilon, f^+Z \le u_T - v_T \,\middle|\, \tau > u_T\right).
\end{aligned}$$

For the first summand of (1.2.30) we get

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\in O}|\widetilde{Y}(t)-\theta|>a_T\epsilon, f^+Z>u_T-v_T\,\middle|\,\tau>u_T\right) \qquad (1.2.31)\\
&\leq \frac{\mathbb{P}\left(\sup_{t\in O}|\widetilde{Y}(t)-\theta|>a_T\epsilon, f^+Z>u_T-v_T\right)}{\mathbb{P}(\tau>u_T)}\\
&= \frac{\mathbb{P}\left(\sup_{t\in O}|\widetilde{Y}(t)-\theta|>a_T\epsilon\right)\mathbb{P}\left(f^+Z>u_T-v_T\right)}{\mathbb{P}(\tau>u_T)} \overset{T\to\infty}{\longrightarrow} 0
\end{aligned}$$

by the independence of $\widetilde{Y}-\theta$ and Z. The last term tends to zero, since $a_T\to\infty$, $T\mathbb{P}(f^+Z>u_T-v_T)\to\exp(-x+\epsilon)$ for $T\to\infty$ and (1.2.24) holds.

Using Lemma 1.2.3 *(a)* and (1.2.24) we get for the second summand of (1.2.30)

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\in O}|Y(t)-\tau|>a_T\epsilon, f^+Z\leq u_T-v_T\,\middle|\,\tau>u_T\right)\\
&\leq \frac{\mathbb{P}\left(\tau>u_T, f^+Z\leq u_T-v_T\right)}{\mathbb{P}\left(\tau>u_T\right)} \overset{T\to\infty}{\longrightarrow} 0. \qquad (1.2.32)
\end{aligned}$$

Therefore (1.2.29) is proven by (1.2.30)-(1.2.32). Invoking again (1.2.24) we see that

$$\begin{aligned}
\mathbb{P}\left(\tau>u_T+a_Ty_i|\,\tau>u_T\right) &= \frac{\mathbb{P}\left(\tau>a_T\max\{x,x+y_i\}+b_T\right)}{\mathbb{P}\left(\tau>a_Tx+b_T\right)}\\
&\overset{T\to\infty}{\longrightarrow} \exp(-\max\{y_i,0\}). \qquad (1.2.33)
\end{aligned}$$

Taking (1.2.29) into account we obtain the second statement of *(b)*.

(c) By considering (1.2.24) and the fact that for $|f(t)|<f^+$

$$\mathbb{P}\left(Y(t)>a_T(x+y)+b_T\right)=o\left(\mathbb{P}\left(f^+Z>a_T(x+y)+b_T\right)\right) \quad \text{for } T\to\infty,$$

by Proposition 1.1.2 *(iii)* we conclude

$$\begin{aligned}
&\mathbb{P}\left(Y(t)>u_T+a_Ty|\,\tau>u_T\right)\\
&\leq \frac{\mathbb{P}\left(Y(t)>a_T(x+y)+b_T\right)}{\mathbb{P}\left(\tau>a_Tx+b_T\right)}\\
&= \frac{\mathbb{P}(Y(t)>a_T(x+y)+b_T)}{\mathbb{P}(\tau>a_T(x+y)+b_T)}\frac{\mathbb{P}\left(\tau>a_T(x+y)+b_T\right)}{\mathbb{P}\left(\tau>a_Tx+b_T\right)} \overset{T\to\infty}{\longrightarrow} 0.
\end{aligned}$$

□

Example 1.2.9 (Continuation of Example 1.2.6)
Suppose $P^{(1)} = 1$ and $P^{(2)} = 0$. Let $k \in \mathbb{Z}$ be fixed. Define the discrete-time process $Y(n) := Y_n$, $f(n) := c_{n-k}$ for $n \in \mathbb{Z}$ and $Z := \xi_k$. Let X_T be a stochastic process with $\mathbb{P}(X_T \in D) = \mathbb{P}(Y/b_T \in D | Y_k > u_T)$ for $D \in \mathcal{B}(\mathbb{R}^{\mathbb{Z}})$. Then we get $X_T \overset{T\to\infty}{\Longrightarrow} \{c_{n-k}/c^+\}_{n\in\mathbb{Z}}$ by Lemma 1.2.7 *(a)*. For a subclass of the extended heavy tailed Weibull distribution, Example 1.1.4, this result can be found in Rootzén [131], Theorem 8.6. □

1.2.3 Marked point processes

In this section we introduce *marked point processes*, aiming at a description of the point process of high level exceedances and at the same time the excursion following an exceedance. First we define the proper sample path space. We use the notation of Section 1.2.1.

Let $\mathbb{D}(\mathbb{R})$ denote the space of functions on $\mathbb{R}$ which are right continuous and have left-hand limits. The space $\mathbb{D}(\mathbb{R})$ is provided with a metric d that generates the Skorohod topology, such that $\mathbb{D}(\mathbb{R})$ is separable and complete (see Billingsley [28], Section 16). Convergence with respect to d is equivalent to convergence of the restriction on every compact set $[-m, m]$ in $\mathbb{D}[-m, m]$ with metric d_m, which also generates the Skorohod topology in $\mathbb{D}[-m, m]$, since

$$d(x, y) = \sum_{m=1}^{\infty} 2^{-m} \min\{1, d_m(x|_{[-m,m]}, y|_{[-m,m]})\} \quad \text{for } x, y \in \mathbb{D}(\mathbb{R}).$$

Let $\mathcal{D}$ be the Borel σ-field in $\mathbb{D}(\mathbb{R})$. Sufficient for convergence in $\mathbb{D}(\mathbb{R})$ is uniform convergence on compacta. Denote $S_0 = M_P(S)$, $\widetilde{d}_0 = \rho$ and $S_k = \mathbb{D}(\mathbb{R})$, $\widetilde{d}_k = d$ for $k \in \mathbb{N}$. Then for every $k \in \mathbb{N}_0$ the metric space $(S_k, \widetilde{d}_k)$ is separable and complete. Consider the infinite Cartesian product $E = S_0 \times S_1 \times S_2 \times \cdots$. A random element $\eta = (\kappa, X_1, X_2, \ldots)$ in E is called a *marked point process*. It is clear that E provided with the product metric

$$\widetilde{d}(x, y) = \sum_{k=0}^{\infty} 2^{-k} \min\{1, \widetilde{d}_k(x_k, y_k)\} \quad \text{for } x, y \in E$$

of coordinatewise convergence is complete and separable. Convergence of random elements in E is described by weak convergence. For further details see Billingsley [28], Pollard [120] or Lindvall [100]. We need the following criteria for weak convergence in E, which is a modification of Billingsley [28], Theorem 3.1.

Lemma 1.2.10
Assume that $\widetilde{\zeta}_T = (\widetilde{\kappa}_T, \widetilde{X}_{1,T}, \widetilde{X}_{2,T}, \ldots)$, $\zeta_T = (\kappa_T, X_{1,T}, X_{2,T}, \ldots)$ *and* ζ *are random elements in* $(E, \widetilde{d})$ *satisfying* $\widetilde{\zeta}_T \stackrel{T\to\infty}{\Longrightarrow} \zeta$ *in* E. *Assume furthermore that, for any* $\epsilon > 0$, $k, m \in \mathbb{N}$,

$$\lim_{T\to\infty} \mathbb{P}\left(\rho(\kappa_T, \widetilde{\kappa}_T) > \epsilon\right) = 0 \quad \text{and} \quad \lim_{T\to\infty} \mathbb{P}\left(\sup_{-m\le t\le m} |X_{k,T}(t) - \widetilde{X}_{k,T}(t)| > \epsilon\right) = 0.$$

Then

$$\zeta_T \stackrel{T\to\infty}{\Longrightarrow} \zeta \quad \text{in } E.$$

Proof.
Sufficient for convergence in $\mathbb{D}(\mathbb{R})$ is uniform convergence with respect to the supremum norm on every set $[-m, m]$ for $m \in \mathbb{N}$. Hence the assumption implies

$$\lim_{T\to\infty} \mathbb{P}\left(d_m\left(X_{k,T}|_{[-m,m]}, \widetilde{X}_{k,T}|_{[-m,m]}\right) > \epsilon\right) = 0. \tag{1.2.34}$$

For a given $\epsilon > 0$ choose $i \in \mathbb{N}$ such that $2^{-i} < \epsilon/2$, then

$$\sum_{m=i+1}^{\infty} 2^{-m} < \frac{\epsilon}{2}.$$

Moreover,

$$\mathbb{P}(d(X_{k,T}, \widetilde{X}_{k,T}) > \epsilon) \le \sum_{m=0}^{i} \mathbb{P}\left(d_m(X_{k,T}|_{[-m,m]}, \widetilde{X}_{k,T}|_{[-m,m]}) > \frac{\epsilon}{2(i+1)}\right) \stackrel{T\to\infty}{\longrightarrow} 0,$$

where we applied (1.2.34). Similarly we get

$$\mathbb{P}(\widetilde{d}(\zeta_T, \widetilde{\zeta}_T) > \epsilon) \le \mathbb{P}\left(\rho\left(\kappa_T, \widetilde{\kappa}_T\right) > \frac{\epsilon}{2(i+1)}\right) + \sum_{k=1}^{i} \mathbb{P}\left(d(X_{k,T}, \widetilde{X}_{k,T}) > \frac{\epsilon}{2(i+1)}\right),$$

which tends to zero for $T \to \infty$. The assertion is now a consequence of Theorem 3.1 of Billingsley [28]. □

Now we formulate the main result of this section.

Theorem 1.2.11
For $k \in \mathbb{N}$ *let* $Y_k = \{Y_k(t)\}_{t\in\mathbb{R}}$ *be stochastic processes in* $(\mathbb{D}(\mathbb{R}), \mathcal{D})$ *with the decomposition*

$$Y_k(t) = f_k(t) Z_k + \widetilde{Y}_k(t) \quad \text{for } t \in \mathbb{R},$$

where $f_k : \mathbb{R} \to \mathbb{R}$ *is a deterministic function in* $\mathbb{D}(\mathbb{R})$ *with* $f_k^+ < \infty$, $f_k(\alpha_k) = f_k^+$ *for some* $\alpha_k \in \mathbb{R}$ *and* $Z_k \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$ *is a r. v. independent of* $\widetilde{Y}_k = \{\widetilde{Y}_k(t)\}_{t\in\mathbb{R}}$ *in* $\mathbb{D}(\mathbb{R})$*. Let* $a_{k,T} > 0$, $b_{k,T} \in \mathbb{R}$, $k \in \mathbb{N}$, $T > 0$, *be constants such that for* $u_{k,T} = a_{k,T}x + b_{k,T}$ *with* $x \in \mathbb{R}$,

$$\lim_{T\to\infty} T\mathbb{P}(f_k^+ Z_k > u_{k,T}) = \exp(-x)$$

holds. Furthermore for $k \in \mathbb{N}$*, let* $\tau_k = f_k^+ Z_k + \theta_k$*, where*

$$\mathbb{P}(\theta_k > x) = o(\mathbb{P}(f_k^+ Z_k > x)) \quad \text{for } x \to \infty,$$

and $X_{k,T}$ *be a random element in* $\mathbb{D}(\mathbb{R})$ *with distribution given by*

$$\mathbb{P}\left(X_{k,T} \in D\right) = \mathbb{P}\left(Y_k \in D \middle|\, \tau_k > u_{k,T}\right) \quad \text{for } D \in \mathcal{D}.$$

Finally, let κ_T *and* κ *be point processes in* $M_P(S)$ *with* $\kappa_T \overset{T\to\infty}{\Longrightarrow} \kappa$ *and*

$$\eta_T = \left(\kappa_T, \frac{X_{1,T}}{b_{1,T}}, \frac{X_{1,T}}{X_{1,T}(\alpha_1)}, \frac{X_{2,T}}{b_{2,T}}, \frac{X_{2,T}}{X_{2,T}(\alpha_2)}, \dots\right), \quad \eta = \left(\kappa, \frac{f_1}{f_1^+}, \frac{f_1}{f_1^+}, \frac{f_2}{f_2^+}, \frac{f_2}{f_2^+}, \dots\right)$$

be marked point processes in E*. Then*

$$\eta_T \overset{T\to\infty}{\Longrightarrow} \eta \quad \text{in } E.$$

The normalized marks $X_{k,T}/b_{k,T}$ characterize the normalized behavior of the process after an exceedance and $X_{k,T}/X_{k,T}(\alpha_k)$ the relative behavior.

Proof.
Define the continuous mapping $h_1 : M_P(S) \to E$ with

$$h_1(y) := \left(y, \frac{f_1}{f_1^+}, \frac{f_1}{f_1^+}, \frac{f_2}{f_2^+}, \frac{f_2}{f_2^+}, \dots\right)$$

and the marked point process

$$\zeta_T := \left(\kappa_T, \frac{X_{1,T}}{b_{1,T}}, \frac{X_{1,T}}{b_{1,T}}, \frac{X_{2,T}}{b_{2,T}}, \frac{X_{2,T}}{b_{2,T}}, \dots\right)$$

in E. Applying the continuous mapping theorem, Billingsley [28], Theorem 2.7, we obtain

$$\widetilde{\zeta}_T := h_1(\kappa_T) \overset{T\to\infty}{\Longrightarrow} h_1(\kappa) = \eta. \tag{1.2.35}$$

Define the set $D = \{g \in \mathbb{D}(\mathbb{R}) : \sup_{-m\le t\le m} |g(t) - f_k(t)/f_k^+| > \epsilon\} \in \mathcal{D}$, then by Lemma 1.2.7 *(a)* we know

$$\begin{aligned}
&\lim_{T\to\infty} \mathbb{P}\left(\sup_{-m\le t\le m} \left| \frac{X_{k,T}(t)}{b_{k,T}} - \frac{f_k(t)}{f_k^+} \right| > \epsilon \right) \\
&\quad = \lim_{T\to\infty} \mathbb{P}\left(\frac{X_{k,T}}{b_{k,T}} \in D \right) \\
&\quad = \lim_{T\to\infty} \mathbb{P}\left(\sup_{-m\le t\le m} \left| \frac{Y_k(t)}{b_{k,T}} - \frac{f_k(t)}{f_k^+} \right| > \epsilon \,\middle|\, \tau_k > u_{k,T} \right) = 0. \qquad (1.2.36)
\end{aligned}$$

The limit relations (1.2.35) and (1.2.36) ensure that the assumptions of Lemma 1.2.10 are satisfied and we get $\zeta_T \overset{T\to\infty}{\Longrightarrow} \eta$ in E. There also exists a measurable and a. s. continuous mapping $h_2 : E \to E$ with $h_2(\zeta_T) = \eta_T$ regarding to the probability measure of η. Applying the continuous mapping theorem, Billingsley [28], Theorem 2.7, again we obtain $\eta_T = h_2(\zeta_T) \overset{T\to\infty}{\Longrightarrow} h_2(\eta) = \eta$. □

1.3 Tail behavior

In this section we estimate the tail behavior of a subexponential Lévy driven MA process Y given in (1.0.1). As one would expect any marginal distribution of a MA process with a kernel function, which has a finite number of local extremes, has a lighter tail than of $f^+L(1)$. In addition to prove this, we will present sufficient conditions for $Y(0)$ to be subexponential.

Note, that the Lévy measure of $Y(0)$ is by (1.1.6) for $x > f^+$ as given by

$$\nu_Y[x,\infty) = \int_{f^+(s)>0} \nu_1 \left[\frac{x}{f^+(s)}, \infty\right) ds + \int_{f^-(s)>0} \nu_2 \left(-\infty, -\frac{x}{f^-(s)}\right] ds \qquad (1.3.1)$$

and hence, coincides with the Lévy measure of $\int_{-\infty}^{\infty} f^+(s)\, dL_1(s) + \int_{-\infty}^{\infty} f^-(s)\, dL_2(s)$. By the tail-equivalence of the Lévy measure and its corresponding probability distribution, in the case of subexponential distributions (Proposition 1.1.2 *(v)*), we immediately see that small jumps of the Lévy process have no influence on the tail behavior of $Y(0)$. Thus L_3 is negligible for the tail of $Y(0)$. We will also use that $\int_{-\infty}^{\infty} f^+(s)\, dL_1(s)$ as well as $\int_{-\infty}^{\infty} f^-(s)\, dL_2(s)$ are positive Poisson shot noise processes.

1.3.1 The Poisson shot noise process

As an important special case we calculate the tail behavior of the positive Poisson shot noise process given in (1.1.11), whose jump sizes are subexponential.

Proposition 1.3.1
Let Y be a positive Poisson shot noise process as given by (1.1.11) satisfying (L1) and (K1). Suppose U is an uniform r. v. on $(-1,1)$, independent of L. Then holds $Y(0) \in \mathcal{S}$ and there exists a $t_0 > 0$ such that for all $t \geq t_0$

$$\mathbb{P}(Y(0) > x) \sim 2t\mathbb{P}(f(tU)L(1) > x) \quad \text{for } x \to \infty.$$

The following Lemma gives an useful representation of a shot noise process. We include the proof here since (to our knowledge) it has not appeared in this generality elsewhere.

Lemma 1.3.2
Let L be a positive compound Poisson process given by (1.1.10), $f : \mathbb{R} \to \mathbb{R}$ be measurable, bounded and $\int_{-t}^{t} f(s)\, dL(s)$ be i. d. for some $t > 0$. Suppose $\{U_k\}_{k\in\mathbb{N}}$, U are i. i. d. uniform r. v. s on $(-1,1)$, independent of L. Then for any $t > 0$,

$$\int_{-t}^{t} f(s)\, dL(s) \stackrel{d}{=} \sum_{k=1}^{N(2t)} f(tU_k)Z_k.$$

Moreover, $\int_{-t}^{t} f(s)\, dL(s) \in \mathcal{S}$ if and only if $f(tU)Z_1 \in \mathcal{S}$. In this case

$$\mathbb{P}\left(\int_{-t}^{t} f(s)\, dL(s) > x\right) \sim 2\mu t\mathbb{P}(f(tU)Z_1 > x) \quad \text{for } x \to \infty.$$

Proof.
Define for $x > 0$ the set

$$M = \{(s,y) \in (-t,t] \times \mathbb{R}_+ \,:\, f(s)y \in (x,\infty)\}.$$

The generating triplet (ν_X, σ_X^2, m_X) of the i. d. r. v. $X = \int_{-t}^{t} f(s)\, L(s)$ is by (1.1.6) given as

$$\nu_X(x,\infty) = \int_{-t}^{t}\int_{-\infty}^{\infty} \mathbf{1}_M(s,y)\nu(dy)\, ds = \mu \int_{-t}^{t} \mathbb{P}(\{Z_1 \,:\, (s,Z_1) \in M\})\, ds,$$

$\sigma_X = 0$ and $m_X = \int_{-\infty}^{\infty} \kappa(x)\, \nu_X(dx)$. On the other hand, the generating triplet $(\nu_{\widetilde{X}}, \sigma^2_{\widetilde{X}}, m_{\widetilde{X}})$ of the i. d. r. v. $\widetilde{X} = \sum_{k=1}^{N(2t)} f(tU_k)Z_k$, which is a compound Poisson r. v., is $\nu_{\widetilde{X}}(x,\infty) = 2\mu t\mathbb{P}(f(tU)Z_1 > x)$ for $x > 0$, $m_{\widetilde{X}} = \int_{-\infty}^{\infty} \kappa(x)\, \nu_{\widetilde{X}}(dx)$ and $\sigma^2_{\widetilde{X}} = 0$. Recalling that an uniform r. v. on $(-t,t)$ has a density u satisfying $ds = 2tu(ds)$ yields

$$\begin{aligned}\nu_{\widetilde{X}}(x,\infty) &= 2\mu t\mathbb{P}(f(tU_1)Z_1 > x) = 2\mu t \int_{-t}^{t} \mathbb{P}(\{Z_1 : (s,Z_1) \in M\})u(ds) \\ &= \mu \int_{-t}^{t} \mathbb{P}(\{Z_1 : (s,Z_1) \in M\})ds = \nu_X(x,\infty).\end{aligned}$$

Thus the generating triplet of X and $\widetilde{X}$ are identical and the first statement holds. Taking into account that $f(tU)$ is a bounded r. v. and $Z_1 \in \mathcal{S}$, then we have $f(tU)Z_1 \in \mathcal{S}$ by Proposition 1.1.2 *(vi)*. Finally, by applying Proposition 1.1.2 *(v)* we obtain $X \in \mathcal{S}$ if and only if $f(tU_1)Z_1 \in \mathcal{S}$. In this case

$$\mathbb{P}(X > x) \sim \nu_X(x,\infty) = 2\mu t\mathbb{P}(f(tU)Z_1 > x) \quad \text{for } x \to \infty.$$

□

Proof of Proposition 1.3.1.

Note that $f : \mathbb{R} \to \mathbb{R}_+$ and Z_1 has only support on the positive real line. By condition (K1) f is bounded and the set D of discontinuities of f is at most countable. If $f(x) = 0$ for every $x \in D^c$, then $f(x) \neq 0$ only on a set with Lebesgue measure zero; hence $Y(0) = 0$ a. s.. Otherwise there exists an $x^* \in D^c$ with $f(x^*) > 0$. Since f is continuous in x^*, there exist a $x_1, x_2 \in \mathbb{R}$ with $x_1 < x^* < x_2$ such that

$$f(x) > f(x^*)/2 =: \widetilde{f} \quad \text{for } x \in (x_1,x_2). \tag{1.3.2}$$

Let define the positive sequence $c_n = \sup_{t\in[n-1,n)} f(t)$ for $n \in \mathbb{N}$. From (K1) we conclude

$$\sum_{k=-\infty}^{\infty} c_k^\delta < \infty. \tag{1.3.3}$$

Furthermore, define the kernel function

$$\overline{f}(t) = \sum_{k=-\infty}^{\infty} c_k\, \mathbf{1}_{[k-1,k)}(t) \quad \text{for } t \in \mathbb{R}.$$

Then $f(t) \le \overline{f}(t)$ for $t \in \mathbb{R}$. From (1.3.3) we know that $c_k \to 0$ for $k \to \pm\infty$. Thus there exists a $t_0 > \max\{|x_1|, |x_2|\}$ such that

$$f(t) \le \overline{f}(t) \le \widetilde{f}/4 \quad \text{for } |t| \ge t_0. \tag{1.3.4}$$

Using the notation $\int_a^b f(t-s)\,dL(s)$ for $\int_{(a,b]} f(t-s)\,dL(s)$ we define the following positive r. v. s with kernel function $\overline{f}$ and f, respectively,

$$X_0 := \int_{-\infty}^{-t_0} \overline{f}(s)\,dL(s) + \int_{t_0}^{\infty} \overline{f}(s)\,dL(s) = \sum_{|k|\ge t_0} c_k[L(k) - L(k-1)],$$

$$X_1 := \int_{-t_0}^{x_1} f(s)\,dL(s), \qquad X_2 := \int_{x_1}^{x_2} f(s)\,dL(s), \qquad X_3 := \int_{x_2}^{t_0} f(s)\,dL(s).$$

If we write c^* for $\max_{|k|>t_0} c_k$, which is by (1.3.4) less than $\widetilde{f}/4$, and

$$P^* = \#\{k : |k| \ge t_0, c_k = c^*\},$$

then we get using Proposition 1.3 in Davis and Resnick [55],

$$\mathbb{P}(X_0 > x) \sim P^*\mathbb{P}(c^*L(1) > x) \quad \text{for } x \to \infty.$$

Thus, by $L(1) \in \mathcal{R}_{-\infty}$ and (1.3.4)

$$\mathbb{P}(X_0 > x) = o(\mathbb{P}(L(1) > 2x/\widetilde{f})) \quad \text{for } x \to \infty. \tag{1.3.5}$$

Moreover, we obtain $X_1 + X_2 + X_3, X_2 \in \mathcal{S}$ and for $x \to \infty$

$$\mathbb{P}(X_1 + X_2 + X_3 > x) = \mathbb{P}\left(\int_{-t_0}^{t_0} f(s)\,dL(s) > x\right) \sim 2t_0\mathbb{P}(f(t_0U)L(1) > x), \tag{1.3.6}$$

$$\mathbb{P}(X_2 > x) \sim 2\frac{x_2 - x_1}{2}\mathbb{P}\left(f\left(\frac{x_1 + x_2}{2} + \frac{x_2 - x_1}{2}U\right)L(1) > x\right) \tag{1.3.7}$$

by Lemma 1.3.2. Taking (1.3.2), (1.3.7) and $\mathbb{P}(X_2 > x) \le \mathbb{P}(X_1 + X_2 + X_3 > x)$ for $x \in \mathbb{R}$ into account

$$\limsup_{x\to\infty} \frac{\mathbb{P}(L(1) > 2x/\widetilde{f})}{\mathbb{P}(X_1 + X_2 + X_3 > x)} \le 1. \tag{1.3.8}$$

Considering (1.3.5) and (1.3.8) we get

$$\mathbb{P}(X_0 > x) = o(\mathbb{P}(X_1 + X_2 + X_3 > x)) \quad \text{for } x \to \infty. \tag{1.3.9}$$

Applying (1.3.6), (1.3.9) and Proposition 1.1.2 *(iii)* gives $Y(0) \in \mathcal{S}$ and for $x \to \infty$

$$\mathbb{P}(Y(0) > x) \sim \mathbb{P}(X_1 + X_2 + X_3 > x) \sim 2t_0\mathbb{P}(f(t_0U)L(1) > x).$$

The result also holds if we replace t_0 by some $t \geq t_0$. □

1.3.2 The general MA process

Invoking the results for the positive Poisson shot noise process in the last section we can now derive results on the tail behavior of subexponential Lévy driven MA processes.

Theorem 1.3.3
Let Y be a stationary MA process as given in (1.0.1) satisfying conditions (K1) and (L1). If f is also negative we additionally assume (L2). Suppose U is a uniform r. v. on $(-1, 1)$ and independent of L. Then there exists a $t_0 > 0$ such that $Y(0) \in \mathcal{S}$ if and only if $f(tU)L(1) \in \mathcal{S}$ for every $t \geq t_0$. In this case

$$\mathbb{P}(Y(0) > x) \sim 2t\mathbb{P}(f(tU)L(1) > x) \quad \text{for } x \to \infty.$$

Note that $|f(t)| < f^+$ for $t \geq t_0$. For the proof we need the following Lemma.

Lemma 1.3.4
Let X_i, Z_i, $\widetilde{Z}_i$, be independent r. v. s, $i = 1, 2$, where X_i has bounded support on $\mathbb{R}_+$. Suppose $\mathbb{P}(Z_i > x)/\mathbb{P}(\widetilde{Z}_i > x) \stackrel{x\to\infty}{\longrightarrow} q \in [0, \infty]$. Then

$$\lim_{x\to\infty} \frac{\mathbb{P}(X_1Z_1 > x) + \mathbb{P}(X_2Z_2 > x)}{\mathbb{P}(X_1\widetilde{Z}_1 > x) + \mathbb{P}(X_2\widetilde{Z}_2 > x)} = q.$$

Proof.
Let $\epsilon > 0$ be arbitrary. Since $\mathbb{P}(Z_i > x)/\mathbb{P}(\widetilde{Z}_i > x) \stackrel{x\to\infty}{\longrightarrow} q$ there exists an $x_0 > 0$ such that $q - \epsilon < \mathbb{P}(Z_i > x)/\mathbb{P}(\widetilde{Z}_i > x) < q + \epsilon$ for $x > x_0$. Denote by F_{X_i} the d. f.

of X_i, which has support on $[0, c_i]$ for $c_i > 0$. Then for $x > \max\{c_1, c_2\}x_0$,

$$\begin{aligned}\mathbb{P}(X_1Z_1 > x) + \mathbb{P}(X_2Z_2 > x) &= \int_0^{c_1} \mathbb{P}\left(Z_1 > \frac{x}{u}\right) F_{X_1}(du) + \int_0^{c_2} \mathbb{P}\left(Z_2 > \frac{x}{u}\right) F_{X_2}(du) \\ &\leq (q+\epsilon)\int_0^{c_1} \mathbb{P}\left(\widetilde{Z}_1 > \frac{x}{u}\right) F_{X_1}(du) + (q+\epsilon)\int_0^{c_2} \mathbb{P}\left(\widetilde{Z}_2 > \frac{x}{u}\right) F_{X_2}(du) \\ &= (q+\epsilon)[\mathbb{P}(X_1\widetilde{Z}_1 > x) + \mathbb{P}(X_2\widetilde{Z}_2 > x)]\end{aligned}$$

and, similarly, $(q-\epsilon)[\mathbb{P}(X_1\widetilde{Z}_1 > x) + \mathbb{P}(X_2\widetilde{Z}_2 > x)] \leq \mathbb{P}(X_1Z_1 > x) + \mathbb{P}(X_2Z_2 > x)$ for $x > \max\{c_1, c_2\}x_0$. $\square$

Proof of Theorem 1.3.3.

We use the decompositions $L = L_1 - L_2 + L_3$ as given on p. 16. Let U_1, U_2 be independent uniform r. v. s on $(-1, 1)$. Then $L_1(1), L_2(1), f^+(tU_1)L_1(1), f^-(tU_2)L_2(1) \in \mathcal{S}$ by Proposition 1.3.1 *(vi)*. Taking Proposition 1.3.1 with some $t > t_0$, Proposition 1.1.2 *(v)* and Lemma 1.3.4 into account, we obtain $Y(0) \in \mathcal{S}$ if and only if for $x \to \infty$

$$\begin{aligned}\mathbb{P}(Y(0) > x) &\sim \nu_Y(x,\infty) = \nu_{Y_1}(x,\infty) + \nu_{Y_2}(x,\infty) \\ &\sim 2t\mathbb{P}(f^+(tU_1)L_1(1) > x) + 2t\mathbb{P}(f^-(tU_2)L_2(1) > x).\end{aligned} \tag{1.3.10}$$

Let $p > 0$. Since $L(1), -L(1) \in \mathcal{S}$, we get $\mathbb{P}(L^+(1) > x) \sim \mathbb{P}(L_1(1) > x)$ and $\mathbb{P}(L(1)^- > x) \sim \mathbb{P}(L_2(1) > x)$ for $x \to \infty$ by Proposition 1.1.2 *(v)*. Then with Lemma 1.3.4 we have for $x \to \infty$

$$\begin{aligned}\mathbb{P}(f(tU)L(1) > x) &= \mathbb{P}(f^+(tU)L^+(1) > x) + \mathbb{P}(f^-(tU)L^-(1) > x) \\ &\sim \mathbb{P}(f^+(tU)L_1(1) > x) + \mathbb{P}(f^-(tU)L_2(1) > x).\end{aligned} \tag{1.3.11}$$

Thus by (1.3.10)-(1.3.11) $Y(0) \in \mathcal{S}$ if and only if $\mathbb{P}(Y(0) > x) \sim 2t\mathbb{P}(f(tU)L(1) > x)$ for $x \to \infty$. In the case $p = 0$ the Lévy measure of $f^-(tU)L_2(1)$ has bounded support and hence also the d. f.. $\square$

We shall give some sufficient conditions for $f(tU)L(1) \in \mathcal{S}$ and hence $Y(0) \in \mathcal{S}$.

Remark 1.3.5

Let Y be a stationary MA process as given in (1.0.1) satisfying (K1), (L1) and has decomposition as given in (1.0.3). Taking $f^+(tU)L_1(1), f^-(tU)L_2(1) \in \mathcal{S}$, Proposition 1.1.2 *(iii)* and (1.3.11) into account a sufficient condition for $f(tU)L(1)$ to be subexponential is

$$\frac{\mathbb{P}(f^-(tU)L_2(1) > x)}{\mathbb{P}(f^+(tU)L_1(1) > x)} \stackrel{x\to\infty}{\longrightarrow} q \geq 0. \tag{1.3.12}$$

The following examples satisfy (1.3.12):

(a) $f(x) = -f(-x)$ for $x \in \mathbb{R}$ and condition (L2) is satisfied.

(b) $f \geq 0$.

(c) $L = L_1 + L_3$.

(d) f is right or left continuous in some α with $f(\alpha) = f^+$ and $f^- < f^+$.

(e) f is a step function and condition (L2) is satisfied.

□

Remark 1.3.6
In the case of a discrete-time MA process as given in Example 1.2.6, let $L(1) = \xi_1 \in \mathcal{S}$ satisfies the tail balance condition in (L2). This is sufficient for Y_n to be subexponential. Then

$$\mathbb{P}(Y_n > x) \sim \left(P^{(1)} + \frac{1-p}{p}P^{(2)}\right)\mathbb{P}(c^+L(1) > x) \quad \text{for } x \to \infty.$$

In this case the additional assumption that either $\nu(-\infty,-1) = 0$ or $-L(1) \in \mathcal{S}$ is not necessary since the tails of $\int_{-\infty}^{\infty} f^-(s)\,dL_2(s)$ and $\int_{-\infty}^{\infty} f^+(s)\,dL_1(s)$ are by

$$\lim_{x\to\infty} \frac{\mathbb{P}\left(\int_{-\infty}^{\infty} f^-(s)\,dL_2(s) > x\right)}{\mathbb{P}\left(\int_{-\infty}^{\infty} f^+(s)\,dL_1(s) > x\right)} = \frac{1}{p}\frac{P^{(2)}}{P^{(1)}}$$

comparable. □

Corollary 1.3.7
Let Y be a stationary MA process as given in (1.0.1) satisfying (K1) and (L1). If f is also negative assume additionally (L2). In addition, let $\widetilde{L}$ be a Lévy process satisfying for $x \to \infty$,

$$\mathbb{P}(\widetilde{L}(1) > x) \sim \mathbb{P}(L(1) > x) \quad \text{and} \quad \mathbb{P}(\widetilde{L}(1) < -x) \sim \mathbb{P}(L(1) < -x).$$

Define $\widetilde{Y}(t) = \int_{-\infty}^{\infty} f(t-s)\, d\widetilde{L}(s)$ *for* $t \in \mathbb{R}$. *Then* $Y(0) \in \mathcal{S}$ *if and only if* $\widetilde{Y}(0) \in \mathcal{S}$ *and in this case*

$$\mathbb{P}(Y(0) > x) \sim \mathbb{P}(\widetilde{Y}(0) > x) \quad \text{for } x \to \infty.$$

Proof.
By Theorem 1.3.3 holds $Y(0) \in \mathcal{S}$ if and only if $f(tU)L(1) \in \mathcal{S}$ and similarly holds $\widetilde{Y}(0) \in \mathcal{S}$ holds if and only if $f(tU)\widetilde{L}(1) \in \mathcal{S}$. If we apply Lemma 1.3.4 and (1.3.11) we get that $f(tU)L(1)$ and $f(tU)\widetilde{L}(1)$ are tail-equivalent. □

The next corollary shows that the assumption $Y(0) \in \mathcal{S}$ is not necessary for the right tail of $Y(0)$ to be lighter than the tail of $f^+L(1)$.

Corollary 1.3.8
Let Y *be a stationary MA process as given in (1.0.1) satisfying (K1), (K2) and (L1). If* f *is also negative assume additionally (L2). Then for* $x \to \infty$,

$$\mathbb{P}(Y(0) > x) = o(\mathbb{P}(f^+L(1) > x)) \quad \text{and} \quad \mathbb{P}(Y(0) < -x) = o(\mathbb{P}(f^+L(1) > x)).$$

Proof.
Let U be a uniform r. v. on $(-1, 1)$. Without loss of generality we assume $P^{(1)} = 1$ and $f(\alpha) = f^+$. Let $\epsilon > 0$ be arbitrary. Then

$$\widetilde{f} := \sup_{t \le \alpha - \epsilon} |f(t)| \vee \sup_{t \ge \alpha + \epsilon} |f(t)| < \sup_{t \in \mathbb{R}} |f(t)| = f^+.$$

On the one hand

$$\frac{\mathbb{P}(f(tU)L_1(1) > x \,|\, |tU - \alpha| \le \epsilon)}{\mathbb{P}(f^+L_1(1) > x)} \le 1$$

and on the other hand, taking into account that the tail of $L_1(1) \in \mathcal{R}_{-\infty}$, then

$$\frac{\mathbb{P}(f(tU)L_1(1) > x||tU - \alpha| > \epsilon)}{\mathbb{P}(f^+L_1(1) > x)} \le \frac{\mathbb{P}(\widetilde{f}L_1(1) > x)}{\mathbb{P}(f^+L_1(1) > x)} \overset{x\to\infty}{\longrightarrow} 0.$$

Thus

$$\begin{aligned}
&\frac{\mathbb{P}(f(tU)L_1(1) > x)}{\mathbb{P}(f^+L_1(1) > x)} \\
&\quad \le \frac{\epsilon}{t} + \left(1 - \frac{\epsilon}{t}\right) \frac{\mathbb{P}(f(tU)L_1(1) > x \,|\, |tU - \alpha| > \epsilon)}{\mathbb{P}(f^+L_1(1) > x)} \overset{x\to\infty}{\longrightarrow} \frac{\epsilon}{t}.
\end{aligned}$$

Regarding on the one hand $\mathbb{P}(f^+[L_1(1) + L_3(1)] > x) \sim \mathbb{P}(f^+L(1) > x)$ and on the other hand $\mathbb{P}(f(tU)[L_1(1) + L_3(1)] > x) \sim \mathbb{P}(f(tU)L_1(1) > x)$ for $x \to \infty$ also

$$\mathbb{P}(f(tU)[L_1(1) + L_3(1)] > x) = o(\mathbb{P}(f^+L(1) > x)) \quad \text{for } x \to \infty.$$

Applying Remark 1.3.5 *(d)* then $Y_1(0) + Y_3(0) \in \mathcal{S}$ and

$\mathbb{P}(Y_1(0) + Y_3(0) > x) \sim \mathbb{P}\left(f(tU)[L_1(1) + L_3(1)]\right) = o(\mathbb{P}(f^+L(1) > x))$ for $x \to \infty$.

Similarly $Y_2(0) \in \mathcal{S}$ and $\mathbb{P}(Y_2(0) > x) = o(\mathbb{P}(f^+L(1) > x))$ for $x \to \infty$. Taking Proposition 1.1.2 *(iii)* into account, then we get for $x \to \infty$,

$$\mathbb{P}(Y(0) > x) = \mathbb{P}([Y_1(0) + Y_3(0)] + Y_2(0) > x) = o(\mathbb{P}(f^+L(1) > x)).$$

We obtain the second statement of Corollary 1.3.8 by choosing $-f$ as kernel function and applying the first statement. The assumption $f^- \le f^+$ has no influence on the proof. □

1.3.3 Tail behavior of $M(h)$ and $M(\Gamma_k)$

We now investigate the extreme behavior of the MA process Y with a. s. sample paths in $\mathbb{D}(\mathbb{R})$. We first study the local maxima of the process, i. e. maxima over an interval of fixed or random length. Define $M(h) = \sup_{0 \le t \le h} Y(t)$ for $h > 0$. Our first result is in the spirit of Rosinski and Samorodnitsky [132], Theorem 2.1. They study the tail behavior of subexponential r. v. s, which are functionals of stochastic processes.

Theorem 1.3.9
Let Y be a stationary MA process as given in (1.0.1) with a. s. sample path in $\mathbb{D}(\mathbb{R})$ and $f \in \mathbb{D}(\mathbb{R})$. Define for $h > 0$, $t \in \mathbb{R}$,

$$\begin{aligned} f_h^+(s) &= \sup_{0\le t\le h} f^+(t+s), & E_h^+ &= \{t \in \mathbb{R} : f_h^+(t\pm) = f^+\}, \\ f_h^-(s) &= \sup_{0\le t\le h} f^-(t+s), & E_h^- &= \{t \in \mathbb{R} : f_h^-(t\pm) = f^+\}. \end{aligned}$$

(a) Let $h > 0$ be fixed. Define for $x > 1$

$$\nu_{\overline{Y}}(x,\infty) = \int_{f_h^+(s)>0} \nu\left(\frac{x}{f_h^+(s)}, \infty\right) ds + \int_{f_h^-(s)>0} \nu\left(-\infty, \frac{-x}{f_h^-(s)}\right) ds.$$

Suppose $\nu_{\overline{Y}}(\cdot,\infty)/\nu_{\overline{Y}}(1,\infty) \in \mathcal{S}$. Then $M(h) \in \mathcal{S}$ and

$$\mathbb{P}(M(h) > x) \sim \nu_{\overline{Y}}(x,\infty) \quad \text{for } x \to \infty.$$

Furthermore, if

$$\overline{Y}(t) = \int_{-\infty}^{\infty} f_h^+(t-s)\, dL_1(s) + \int_{-\infty}^{\infty} f_h^-(t-s)\, dL_2(s)$$

is a well-defined, stationary i. d. process, then $\mathbb{P}(M(h) > x) \sim \mathbb{P}(\overline{Y}(0) > x)$ *for* $x \to \infty$.

(b) Assume $L(1)$ *satisfies conditions (L1), (L2) and* f *satisfies conditions (K1), (K2). Denote by* λ *the Lebesgue measure on* $\mathbb{R}$.

(b1) Then $M(h) \in \mathcal{S}$ *and*

$$\mathbb{P}(M(h) > x) \sim \left(\lambda(E_h^+) + \frac{1-p}{p}\lambda(E_h^-)\right)\mathbb{P}(f^+L(1) > x) \quad \text{for } x \to \infty.$$

(b2) Let τ *be a positive r. v. on the same probability space than* Y. *Furthermore,* τ *is independent of* Y *with* $\mathbb{E}\tau < \infty$, *then also* $M(\tau) \in \mathcal{S}$ *and with* $K := \int_0^\infty \lambda(E_h^+) + \frac{1-p}{p}\lambda(E_h^-)\, F_\tau(dh)$ *holds*

$$\mathbb{P}(M(\tau) > x) \sim K\mathbb{P}(f^+L(1) > x) \quad \text{for } x \to \infty.$$

Remark 1.3.10
(i) Note that *(a)* links the tail behavior of $M(h)$ and $\nu_{\overline{Y}}$ under subexponentiality. No assumption is made concerning the maximum domain of attraction.

(ii) If h is less than the smallest distance between successive extrema, we obtain $\lambda(E_h^+) = P^{(1)}h$ and $\lambda(E_h^-) = P^{(2)}h$. □

For the proof of Theorem 1.3.9 we need the following Lemma.

Lemma 1.3.11
Let Y *be a separable stationary MA process as given in (1.0.1) satisfying* $f \in \mathbb{D}(\mathbb{R})$ *and*

$$\mathbb{P}(\sup_{0\le t\le h} |Y(t)| < \infty) = 1 \quad \text{for } h > 0.$$

Assume that the Lévy measure ν *of* L *has only support on* $[-c, c]$ *for some* $c > 0$. *Then, for every* $\epsilon > 0$ *there exists a* $C > 0$ *such that*

$$\mathbb{P}(M(h) > x) \le \mathbb{P}(\sup_{0\le t\le h} |Y(t)| > x) \le C\mathrm{e}^{-\epsilon x} \quad \text{for } x > 0.$$

Proof.
Let ν_h be the Lévy measure of the i. d. process $\{Y(t)\}_{0\le t\le h}$, i. e. ν_h is the Lévy measure corresponding to the finite dimensional distributions of $\{Y(t)\}_{0\le t\le h}$ (see

Maruyama [106]). Denote by $F = \lambda \times \nu$ the product measure of λ and ν on $\mathbb{R}^2$. Define also the function $V : \mathbb{R}^2 \to \mathbb{R}^{\mathbb{Q}}$ by

$$V(s,x) = \{xf(t-s)\}_{\substack{0\le t\le h\\ t\in\mathbb{Q}}}. \tag{1.3.13}$$

Then $\nu_h = F \circ V^{-1}$ and we calculate

$$\begin{aligned} &\nu_h\left(\left\{\alpha \in \mathbb{R}^{[0,h]} \,:\, \sup_{\substack{0\le t\le h\\ t\in\mathbb{Q}}} |\alpha(t)| > f^+c\right\}\right) \qquad (1.3.14)\\ &= F\circ V^{-1}\left(\left\{\alpha \in \mathbb{R}^{[0,h]} \,:\, \sup_{\substack{0\le t\le h\\ t\in\mathbb{Q}}} |\alpha(t)| > f^+c\right\}\right)\\ &= \int_{\sup_{0\le t\le h}|f(t-s)|\neq 0} \nu\left(\frac{f^+c}{\sup_{0\le t\le h}|f(t-s)|}, \infty\right) + \nu\left(-\infty, \frac{-f^+c}{\sup_{0\le t\le h}|f(t-s)|}\right) ds = 0. \end{aligned}$$

Thus the assumptions of Braverman and Samorodnitsky [33], Lemma 2.1, are satisfied and

$$C := \mathbb{E}\exp(\epsilon \sup_{0\le t\le h}|Y(t)|) = \mathbb{E}\exp(\epsilon \sup_{\substack{0\le t\le h\\ t\in\mathbb{Q}}}|Y(t)|) < \infty \quad \text{for every } \epsilon > 0.$$

Using the Markov inequality we obtain for every $x > 0$

$$\mathbb{P}(M(h) > x) \le \mathbb{P}(\sup_{0\le t\le h}|Y(t)| > x) \le \mathrm{e}^{-\epsilon x}\mathbb{E}\exp(\epsilon \sup_{0\le t\le h}|Y(t)|) = C\mathrm{e}^{-\epsilon x}. \tag{1.3.15}$$

□

Proof of Theorem 1.3.9.
(a) Step 1. Assume $L = L_1 - L_2$.
We have $\nu_{\overline{Y}}$ is the Lévy measure of $\overline{Y}(t)$. Applying Proposition 1.1.2 *(v)* the Lévy measure and the probability measure of $\overline{Y}(t)$ are tail-equivalent. Denote by $F = \lambda\times\nu$ the product measure of λ and ν on $\mathbb{R}^2$ and let the function V as be given in (1.3.13), such that the Lévy measure ν_h of $\{Y(t)\}_{0\le t\le h}$ is $F\circ V^{-1}$. Thus, similarly to (1.3.14),

$$\begin{aligned} H(x) &= \nu_h\left(\left\{\alpha \in \mathbb{R}^{[0,h]} \,:\, \sup_{\substack{0\le t\le h\\ t\in\mathbb{Q}}} \alpha(t) > x\right\}\right)\\ &= \int_{\sup_{0\le t\le h} f^+(t-s)>0} \nu\left(\frac{x}{\sup_{0\le t\le h} f^+(t-s)}, \infty\right) ds + \int_{\sup_{0\le t\le h} f^-(t-s)>0} \nu\left(-\infty, \frac{-x}{\sup_{0\le t\le h} f^-(t-s)}\right) ds\\ &= \nu_{\overline{Y}}(x,\infty) \in \mathcal{S}. \end{aligned}$$

Since Y has a. s. sample paths in $\mathbb{D}(\mathbb{R})$ we have $\mathbb{P}(\sup_{0\le t\le h}|Y(t)|<\infty)=1$ and $M(h)=\sup_{\substack{0\le t\le h\\ t\in\mathbb{Q}}} Y(t)$. Therefore the assumptions of Rosinski and Samorodnitsky [132], Theorem 2.1, are satisfied and hence $M(h)\in\mathcal{S}$ and

$$\mathbb{P}\left(M(h)>x\right)\sim\nu_{\overline{Y}}\left[x,\infty\right)\qquad\text{for } x\to\infty.$$

Step 2. Assume $L=L_1-L_2+L_3$.
We decompose L into two independent Lévy processes $\widetilde{L}$, L_3, such that $L=\widetilde{L}+L_3$, where $\widetilde{L}=L_1-L_2$. Since Y and $\int_{-\infty}^{\infty}f(t-s)\,d\widetilde{L}(s)$ has a. s. sample path in $\mathbb{D}(\mathbb{R})$ also $\int_{-\infty}^{\infty}f(t-s)\,dL_3(s)$ has a. s. sample path in $\mathbb{D}(\mathbb{R})$. Then by Lemma 1.3.11 and Proposition 1.1.2 *(i)*,

$$\mathbb{P}\left(\sup_{0\le t\le h}\left|\int_{-\infty}^{\infty}f(t-s)\,dL_3(s)\right|>x\right)=o(\nu_{\overline{Y}}(x,\infty))\qquad\text{for } x\to\infty.$$

Applying Step 1 and Proposition 1.1.2 *(iii)* we obtain for $x\to\infty$,

$$\begin{aligned}\mathbb{P}(M(h)>x) &\le \mathbb{P}\left(\sup_{0\le t\le h}\int_{-\infty}^{\infty}f(t-s)\,d\widetilde{L}(s)+\sup_{0\le t\le h}\left|\int_{-\infty}^{\infty}f(t-s)\,dL_3(s)\right|>x\right)\\ &\sim \mathbb{P}\left(\sup_{0\le t\le h}\int_{-\infty}^{\infty}f(t-s)\,d\widetilde{L}(s)>x\right)\sim\nu_{\overline{Y}}(x,\infty).\end{aligned}$$

Similarly, by symmetry we have for $x\to\infty$

$$\begin{aligned}\mathbb{P}(M(h)>x) &\ge \mathbb{P}\left(\sup_{0\le t\le h}\int_{-\infty}^{\infty}f(t-s)\,d\widetilde{L}(s)-\sup_{0\le t\le h}\left|\int_{-\infty}^{\infty}f(t-s)\,dL_3(s)\right|>x\right)\\ &\sim \nu_{\overline{Y}}(x,\infty).\end{aligned}$$

Thus, $\mathbb{P}(M(h)>x)\sim\nu_{\overline{Y}}(x,\infty)$ for $x\to\infty$ and hence we obtain $M(h)\in\mathcal{S}$.

(b1) Here $\overline{Y}$ is a well-defined stationary i. d. process. W. l. o. g. we assume $f\ge0$ and $\alpha_j^{(1)}-\alpha_{j-1}^{(1)}\ge1$ for $j=2,\ldots,P^{(1)}$. Define $g_h(t)=f_h^+(t)\,\mathbf{1}_{\mathbb{R}\setminus E_h^+}(t)$ for $t\in\mathbb{R}$,

$$\overline{Y}_1(t):=\int_{-\infty}^{\infty}g_h(t-s)\,dL(s),\qquad\overline{Y}_2(t):=\overline{Y}(t)-\overline{Y}_1(t).$$

The MA process $\overline{Y}_1=\{\overline{Y}_1(t)\}_{t\in\mathbb{R}}$ satisfies the assumptions of Corollary 1.3.8, thus

$$\mathbb{P}(\overline{Y}_1(0)>x)=o(\mathbb{P}(f^+L(1)>x))\qquad\text{for } x\to\infty.$$

Moreover, let $t_1=h$, $t_j=\min\{h,\alpha_j^{(1)}-\alpha_{j-1}^{(1)}\}$ for $j=2,\ldots,P^{(1)}$, such that $\sum_{j=1}^{P^{(1)}}t_j=\lambda(E_h^+)$ and $\overline{Y}_2(0)=f^+\sum_{j=1}^{P^{(1)}}[L(\alpha_j^{(1)})-L(\alpha_j^{(1)}-t_j)]$. Thus by Proposition 1.1.2 *(iii)*,

$$\mathbb{P}(\overline{Y}_2(0)>x)\sim\lambda(E_h^+)\mathbb{P}(f^+L(1)>x)\qquad\text{for } x\to\infty.$$

Applying *(a)* and Proposition 1.1.2 *(iii)* again yields for $x \to \infty$,

$$\begin{aligned} \mathbb{P}(M(h) > x) &\sim \mathbb{P}(\overline{Y}(0) > x) = \mathbb{P}(\overline{Y}_1(0) + \overline{Y}_2(0) > x) \\ &\sim \lambda(E_h^+)\mathbb{P}(f^+ L(1) > x). \end{aligned} \tag{1.3.16}$$

(b2) The sequence $\overline{M}_k = \sup_{k-1 \le t \le k} Y(t)$ is stationary such that for $x \ge 0$,

$$\mathbb{P}(M(s) > x) \le \mathbb{P}\left(\bigcup_{k=1}^{\lceil s \rceil} \{\overline{M}_k > x\}\right) \le (s+1)\mathbb{P}(M(1) > x).$$

Denote by F_τ the d. f. of τ. Then we have an uniform bound

$$\frac{\mathbb{P}(M(\tau) > x)}{\mathbb{P}(M(1) > x)} = \int_0^\infty \frac{\mathbb{P}(M(h) > x)}{\mathbb{P}(M(1) > x)} F_\tau(dh) \le \int_0^\infty (h+1) F_\tau(dh) = \mathbb{E}\tau + 1 \tag{1.3.17}$$

for any $x > 0$. Regarding $\alpha_j^{(1)} - \alpha_{j-1}^{(1)} \ge 1$ for $j = 2, \ldots, P^{(1)}$, (1.3.16) and Remark 1.3.10 *(ii)*, we get

$$\mathbb{P}(M(h) > x) \sim \frac{\lambda(E_h^+)}{hP^{(1)}} \mathbb{P}(M(1) > x) \quad \text{for } x \to \infty,$$

and thus we obtain by dominated convergence

$$\begin{aligned} \lim_{x\to\infty} \frac{\mathbb{P}(M(\tau) > x)}{\mathbb{P}(f^+ L(1) > x)} &= hP^{(1)} \lim_{x\to\infty} \frac{\mathbb{P}(M(\tau) > x)}{\mathbb{P}(M(1) > x)} \\ &= hP^{(1)} \lim_{x\to\infty} \int_0^\infty \frac{\mathbb{P}(M(h) > x)}{\mathbb{P}(M(1) > x)} F_\tau(dh) \\ &= \int_0^\infty \lambda(E_h^+)\, F_\tau(dh) < \infty. \end{aligned}$$

□

1.4 Extremal behavior

1.4.1 The marked point process at a discrete-time skeleton

We start with a short motivation why it makes sense to investigate the continuous-time process Y at a discrete-time skeleton $\{Y(t_n)\}_{n\in\mathbb{N}}$ where $\{t_n\}_{n\in\mathbb{N}}$ are chosen properly to capture the times where big jumps of the Lévy process and extremes of the kernel function occur. Consider the Poisson shot noise process given in (1.1.11) with

$$Y(\Gamma_k+t)=f(t)Z_k+\sum_{\substack{j=-\infty\\ j\neq k,0}}^{\infty} f(\Gamma_k-\Gamma_j+t)Z_j \quad \text{for } k\in\mathbb{N},\ t\in\mathbb{R},$$

and assume that f satisfies condition (K2). If Z_k is really large in comparison to $\{Z_j\}_{j\in\mathbb{Z}\setminus\{0,k\}}$, then $Y(\Gamma_k+t)$ behaves like $f(t)Z_k$, and hence $Y(\Gamma_k+t)/Y(\Gamma_k+\alpha_l^{(1)})$ behaves like $f(t)/f^+$. Since in the case of positive jumps, the process $\{f(t)Z_k\}_{t\in\mathbb{R}}$ achieves local suprema if and only if f achieves local suprema in $\alpha_1^{(1)},\dots,\alpha_{P^{(1)}}^{(1)}$, the points $Y(\Gamma_k+\alpha_l^{(1)})$, $k\in\mathbb{N}, l=1,\dots,P^{(1)}$ are significant. In contrast to the negative jumps the points, where $Y(\Gamma_k+\alpha_l^{(2)})$, $k\in\mathbb{N}, l=1,\dots,P^{(2)}$ are significant.

We come back to the decomposition of $L=L_1-L_2+L_3$, where L_i, $i=1,2$, are positive compound Poisson processes with jumps larger than 1 and generating triplet $(0,0,\nu_i)$. Hence, for $i=1,2$, they have the representations

$$L_i(t)=\sum_{j=1}^{N_i(t)} Z_j^{(i)} \quad \text{for } t\geq 0 \quad \text{and} \quad L_i(t)=\sum_{j=1}^{-N_i(t-)} Z_{-j}^{(i)} \quad \text{for } t\leq 0, \tag{1.4.1}$$

where $N_i=\{N_i(t)\}_{t\in\mathbb{R}}$ is a Poisson process with intensities $\mu_i=\nu_i(\mathbb{R})$, and jump times $\Gamma^{(i)}=\{\Gamma_k^{(i)}\}_{k\in\mathbb{Z}\setminus\{0\}}$, $\Gamma_{-1}^{(i)}\leq 0<\Gamma_1^{(i)}$, $\Gamma_k^{(i)}<\Gamma_{k+1}^{(i)}$, $k\in\mathbb{Z}\setminus\{-1,0\}$. The sequences $Z^{(i)}=\{Z_k^{(i)}\}_{k\in\mathbb{Z}}$ consist of i. i. d. r. v. s with d. f. s $\nu_i\,(-\infty,x]\,/\mu_i$ for $x\in\mathbb{R}$. Furthermore, N_1, N_2, $Z^{(1)}$ and $Z^{(2)}$ are independent.

In the setup of a subexponential Lévy driven MA process we consider on the one hand large positive jumps of the Lévy process in cooperation with suprema of f, i. e. the sequence $Y(\Gamma_k^{(1)}+\alpha_l^{(1)})$, $k\in\mathbb{N}$, $l=1,\dots,P^{(1)}$, and on the other hand large negative jumps in cooperation with local infima of f, i. e. $Y(\Gamma_k^{(2)}+\alpha_l^{(2)})$, $k\in\mathbb{N}$, $l=1,\dots,P^{(2)}$.

First we compute the point process behavior of the discrete-time skeleton $\{Y(t_n)\}_{n\in\mathbb{N}}$ with $t_n\in\{\Gamma_k^{(i)}+\alpha_l^{(i)}: k\in\mathbb{N}, l=1,\dots,P^{(i)}, i=1,2\}$. After that we study the behavior of the continuous-time process Y, if $Y(t_n)$ exceeds a large threshold, i. e. we consider marked point processes.

Theorem 1.4.1
Let Y be a stationary MA process as given by (1.0.1) satisfying conditions (K1) and (K2), where L has the decomposition given in (1.4.1). Suppose $L(1)\in\mathcal{S}\cap\mathrm{MDA}(\Lambda)$ with $a_T>0$, $b_T\in\mathbb{R}$, $u_T=a_Tx+b_T$ for $x\in\mathbb{R}$ such that

$$\lim_{T\to\infty} T\mathbb{P}(f^+L(1)>u_T)=\exp(-x).$$

If f is also negative, we additionally assume (L2). For $i = 1, 2$, $l = 1, \dots, P^{(i)}$ define the corresponding point processes in S by

$$\kappa_T^{(i,l)} = \sum_{k=0}^{\infty} \varepsilon_{\left((\Gamma_k^{(i)}+\alpha_l^{(i)})/T, a_T^{-1}(Y(\Gamma_k^{(i)}+\alpha_l^{(i)})-b_T)\right)}, \qquad \widetilde{\kappa}_T^{(i)} = \sum_{k=0}^{\infty} \varepsilon_{\left(k/(T\mu_i), a_T^{-1}(f^+ Z_k^{(i)}-b_T)\right)}.$$

Let $\kappa^{(i)}$ be a $\mathrm{PRM}(\vartheta_i)$, *$i = 1, 2$, with mean measure $\vartheta_1(dt \times dx) = dt \times \mathrm{e}^{-x}\, dx$ and $\vartheta_2(dt \times dx) = dt \times (1-p)/p\, \mathrm{e}^{-x}\, dx$, respectively. Suppose $\kappa^{(1)}$ and $\kappa^{(2)}$ are independent. Furthermore, if $p > 0$, define the point processes*

$$\kappa_T = \sum_{l=1}^{P^{(1)}} \kappa_T^{(1,l)} + \sum_{l=1}^{P^{(2)}} \kappa_T^{(2,l)}, \quad \widetilde{\kappa}_T = P^{(1)} \widetilde{\kappa}_T^{(1)} + P^{(2)} \widetilde{\kappa}_T^{(2)} \text{ and } \kappa = P^{(1)} \kappa^{(1)} + P^{(2)} \kappa^{(2)}.$$

For $p = 0$, define $\kappa_T = \sum_{l=1}^{P^{(1)}} \kappa_T^{(1,l)}$, $\widetilde{\kappa}_T = P^{(1)} \widetilde{\kappa}_T^{(1)}$ and $\kappa = P^{(1)} \kappa^{(1)}$. Suppose $I = [s, t) \times (x, \infty) \subseteq S$. Then for $i = 1, 2$, $l = 1, \dots, P^{(i)}$, it holds

$$\lim_{T\to\infty} \mathbb{P}(\kappa_T^{(i,l)}(I) \neq \widetilde{\kappa}_T^{(i)}(I)) = 0$$

and $\kappa_T \overset{T\to\infty}{\Longrightarrow} \kappa$.

Theorem 1.4.1 states that exceedances of $\{Y(\Gamma_k^{(i)} + \alpha_l^{(i)})\}_{k\in\mathbb{N}}$ above a high threshold behave like $\{f^+ Z_k^{(i)}\}_{k\in\mathbb{Z}}$, this are the extremes of f times the i. i. d. sequence of jump sizes. Hence, the influence of small jumps of the Lévy process is negligible. We notice that the limit process of the *point process of exceedances* $\kappa_T(\cdot \times (x, \infty))$, $x > 0$ be fixed, is the sum of two independent compound Poisson processes with constant cluster sizes $P^{(1)}$, $P^{(2)}$. The following corollary describes the behavior of the marked point process and the excess over threshold distribution function.

Corollary 1.4.2
Assume that the conditions of Theorem 1.4.1 holds and Y has a. s. sample paths in $\mathbb{D}(\mathbb{R})$. Then the following statements hold:

(a) *Let $X_{k,T}$, $k \in \mathbb{N}$, $T > 0$, be random elements in $\mathbb{D}(\mathbb{R})$ with distributions given by*

$$\begin{aligned} \mathbb{P}\left(X_{2k-1,T} \in D\right) &= \mathbb{P}\left(Y(\Gamma_k^{(1)} + \cdot) \in D \,\middle|\, Y(\Gamma_k^{(1)} + \alpha_1^{(1)}) > u_T\right), \\ \mathbb{P}\left(X_{2k,T} \in D\right) &= \mathbb{P}\left(Y(\Gamma_k^{(2)} + \cdot) \in D \,\middle|\, Y(\Gamma_k^{(2)} + \alpha_1^{(2)}) > u_T\right) \end{aligned}$$

for $D \in \mathcal{D}$. The point processes κ_T, κ are defined as in Theorem 1.4.1,

$$\begin{aligned}\eta_T &:= \left(\kappa_T, \frac{X_{1,T}}{b_T}, \frac{X_{1,T}}{X_{1,T}(\alpha_1^{(1)})}, \frac{X_{2,T}}{b_T}, \frac{X_{2,T}}{X_{2,T}(\alpha_1^{(2)})}, \dots\right) \quad \text{and}\\ \eta &:= \left(\kappa, \frac{f}{f^+}, \frac{f}{f^+}, \frac{-f}{f^+}, \frac{-f}{f^+}, \dots\right)\end{aligned}$$

are marked point processes in E. Then

$$\eta_T \stackrel{T\to\infty}{\Longrightarrow} \eta \quad \text{in } E.$$

(b) *Let $i \in \{1,2\}$ be fixed. Define $P = P^{(i)}$, $\alpha_l = \Gamma_k^{(i)} + \alpha_l^{(i)}$, $l = 1, \dots, P$ and $\alpha = \Gamma_k^{(i)} + \alpha_1^{(i)}$. For $y_1, \dots, y_P \in \mathbb{R}$, and $y = \max\{0, y_1, \dots, y_P\}$ we have*

$$\lim_{T\to\infty} \mathbb{P}\left(Y(\alpha_1) > u_T + a_T y_1, \dots, Y(\alpha_P) > u_T + a_T y_P \middle| Y(\alpha) > u_T\right) = \exp(-y).$$

(c) *Let $i \in \{1,2\}$ be fixed, $t \notin O_i$ and $y \in \mathbb{R}$. Then*

$$\lim_{T\to\infty} \mathbb{P}\left(Y(\Gamma_k^{(i)} + t) > u_T + a_T y \middle| Y(\Gamma_k^{(i)} + \alpha_1^{(i)}) > u_T\right) = 0.$$

(d) *Let*

$$\begin{aligned}K_T &= \sum_{k=0}^{\infty} \varepsilon_{\left((\Gamma_k^{(1)}+\alpha_1^{(1)})/T, a_T^{-1}(Y(\Gamma_k^{(1)}+\alpha_l^{(1)})-b_T), \{a_T^{-1}(Y(\Gamma_k^{(1)}+t_i)-b_T)\}_{i=1,\dots m}\right)} \quad \text{and}\\ K &= \sum_{k=0}^{\infty} \varepsilon_{(s_k, P_k, \{P_k \mathbf{1}_{\{f(t_i)=f^+\}}\}_{i=1,\dots,m})}\end{aligned}$$

be point processes in $M_P([0,\infty) \times \mathbb{R}^{m+1})$ for any $t_1, \dots, t_m \in \mathbb{R}$. Then $K_T \stackrel{T\to\infty}{\Longrightarrow} K$.

(e) *Let*

$$\begin{aligned}K_T &= \sum_{k=0}^{\infty} \varepsilon_{\left((\Gamma_k^{(1)}+\alpha_1^{(1)})/T, a_T^{-1}(Y(\Gamma_k^{(1)}+\alpha_l^{(1)})-b_T), \{Y(\Gamma_k^{(1)}+t_i)/b_T\}_{i=1,\dots m}\right)} \quad \text{and}\\ K &= \sum_{k=0}^{\infty} \varepsilon_{(s_k, P_k, \{f(t_i)/f^+\}_{i=1,\dots,m})}\end{aligned}$$

be point processes in $M_P([0,\infty) \times \mathbb{R}^{m+1})$ for any $t_1, \dots, t_m \in \mathbb{R}$. Then $K_T \stackrel{T\to\infty}{\Longrightarrow} K$.

Remark 1.4.3

(i) Marks can be chosen at any extremum of the kernel function f in combination with a large jump of the Lévy process, i. e. it is possible to take the conditional probability on any of the events $\{Y(\Gamma_k^{(i)} + \alpha_l^{(i)}) > u_T\}$ for $\alpha_l^{(i)} \in O_i$. It is also possible to take the conditional probability under $f^+ Z_k^{(i)}$. Then the normalized process $X_{k,T}/b_T \in \mathbb{D}(\mathbb{R})$, where

$$\mathbb{P}(X_{k,T} \in D) = \mathbb{P}(Y(\Gamma_k^{(i)} + \cdot) \in D | f^+ Z_k^{(i)} > u_T) \quad \text{for } D \in \mathcal{D},$$

converges weakly to the deterministic function f/f^+ or $-f/f^+$, respectively. Thus we see, that large jumps of the Lévy process cause extremes of Y if and only if the time is properly chosen with an extreme of f. Hence *(b)* is no surprise, since

$$\lim_{T\to\infty} \mathbb{P}\left(f^+ Z_k^{(i)} > u_T + a_T y_1, \ldots, f^+ Z_k^{(i)} > u_T + a_T y_P \middle| f^+ Z_k^{(i)} > u_T \right) = \exp(-y).$$

For $P^{(1)} = 1$ we obtain the generalized Pareto distribution. This result is also confirmed by the point process convergence.

(ii) Note the similarities of the extreme behavior of the continuous-time MA process and the discrete-time MA process, see Example 1.2.6, 1.2.9. In the case of a Poisson shot noise process this is no surprise since a shot noise process is a MA process with random coefficients.

(iii) Assume Y has a. s. sample paths in $\mathbb{D}(\mathbb{R})$. In *(d)* we can also replace the mark $\left\{a_T^{-1}(Y(\Gamma_k^{(1)} + t_i) - b_T)\right\}_{i=1,\ldots m}$ in $\mathbb{R}^m$ by $\{a_T^{-1}(Y(\Gamma_k^{(1)} + t) - b_T)\}_{t\in I}$ in $\mathbb{D}(I)$ for any compact set $I \subseteq \mathbb{R}$. Instead of weak convergence, we have to use $\widehat{\omega}$-convergence, introduced by Daley and Vere-Jones [50], Section A.2.6, since $\mathbb{D}(I)$ is not locally compact. The same holds for *(e)*. □

We shall use an important standard result of probability theory. If τ is a r. v. with d. f. F_τ independent of the stationary process X, then we get for $x \in \mathbb{R}$,

$$\begin{aligned} \mathbb{P}(X(\tau) > x) &= \int \mathbb{P}(X(t) > x)\, F_\tau(dt) = \int \mathbb{P}(X(0) > x)\, F_\tau(dt) \\ &= \mathbb{P}(X(0) > x). \end{aligned} \tag{1.4.2}$$

Since Y, Y_i are not independent of $\Gamma^{(i)}$, it is no surprise, that the distributions of $Y(\Gamma_k^{(i)} + t)$ and $Y_i(\Gamma_k^{(i)} + t)$, $t \geq 0$, differ with those of $Y(0)$ and $Y_i(0)$, respectively.

But in contrast to this, Y_{3-i}, Y_3 are independent of $\Gamma^{(i)}$, $i = 1, 2$, and thus for $x \in \mathbb{R}$,

$$\mathbb{P}(Y_j(\Gamma_k^{(i)} + t) > x) = \mathbb{P}(Y_j(0) > x) \quad \text{for } x \in \mathbb{R},\ j = 3 - i, 3.$$

In our situation, however, we are confronted with jump times $\Gamma^{(i)}$ which influence the process Y directly. We shall decouple this dependence by using properties of the kernel function f in combination with the decomposition $Y = Y_1 - Y_2 + Y_3$ in independent components. Most important for the extremal behavior of $\{Y(\Gamma_k^{(1)} + \alpha_l^{(1)})\}_{k\in\mathbb{N}}$ will be the Poisson shot noise process Y_1; the remaining part $-Y_2 + Y_3$ will be negligible.

For the proofs of Theorem 1.4.1 and Corollary 1.4.2 we first show the following Lemma.

Lemma 1.4.4
Suppose the assumptions of Theorem 1.4.1 hold. Then for $i = 1, 2$, $\alpha \in O_i$ and $x \to \infty$, we can find r. v. s $\Theta^{(i)}$, $\widetilde{\Theta}^{(i)}$ such that for $x \in \mathbb{R}$

$$\begin{aligned} \mathbb{P}(Y(\Gamma_k^{(i)} + \alpha) - f^+ Z_k^{(i)} > x) &\le \mathbb{P}(\Theta^{(i)} > x), \\ \mathbb{P}(Y(\Gamma_k^{(i)} + \alpha) - f^+ Z_k^{(i)} < -x) &\le \mathbb{P}(\widetilde{\Theta}^{(i)} > x) \end{aligned} \tag{1.4.3}$$

and

$$\mathbb{P}(\Theta^{(i)} > x) = o(\mathbb{P}(f^+ Z_k^{(i)} > x)) \text{ and } \mathbb{P}(\widetilde{\Theta}^{(i)} > x) = o(\mathbb{P}(f^+ Z_k^{(i)} > x)) \quad \text{for } x \to \infty.$$

Moreover, the r. v. $Y(\Gamma_k^{(i)} + \alpha) - f^+ Z_k^{(i)}$ is independent of $Z_k^{(i)}$. If $Y(0) \in \mathcal{S}$ and $t \notin O_i$, then

$$\mathbb{P}(Y(\Gamma_k^{(i)} + t) > x) = o(\mathbb{P}(f^+ Z_1^{(i)} > x)) \quad \text{for } x \to \infty$$

holds.

Proof.
Step 1. Assume L is a compound Poisson process as given in (1.1.10).
Choose $k > 0$ be fixed and define a shifted compound Poisson process $\widetilde{L} = \{\widetilde{L}(t)\}_{t\in\mathbb{R}}$ with jump times

$$\widetilde{\Gamma}_j = \begin{cases} \Gamma_k & \text{for } j = k, \\ \Gamma_k - \Gamma_{k-j} & \text{for } j \neq k, \end{cases} \qquad \text{for } j \in \mathbb{Z},$$

corresponding jump sizes Z_{k-j} at $\widetilde{\Gamma}_j$ and intensity μ. Define $\widetilde{\Gamma}_0 := 0$. Without loss of generality the support of f can be set to $\mathbb{R}$. Then

$$\begin{aligned} Y(\Gamma_k + t) &= \int_{-\infty}^{\infty} f(\Gamma_k + t - s)\, dL(s) = \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} f(t + \widetilde{\Gamma}_{k-m}) Z_m \\ &= \sum_{j=-\infty}^{\infty} f(t+\widetilde{\Gamma}_j) Z_{k-j} - f(t+\Gamma_k) Z_0 \text{ for all } t \in \mathbb{R} \text{ a.s..} \end{aligned}$$

As this equality is a. s., we obtain

$$\begin{aligned} Y(\Gamma_k + t) + f(t+\Gamma_k) Z_0 &= \sum_{\substack{j=-\infty \\ j\neq 0}}^{\infty} f(t+\widetilde{\Gamma}_j) Z_{k-j} + f(t) Z_k \qquad (1.4.4) \\ &= \int_{-\infty}^{\infty} f(t+s)\, d\widetilde{L}(s) + f(t) Z_k = \widetilde{Y}(t) + f(t) Z_k \end{aligned}$$

a. s., where $\widetilde{Y}(t) = \int_{-\infty}^{\infty} f(t+s)\, d\widetilde{L}(s) \stackrel{d}{=} Y(t)$.

Step 2. Assume L is a Lévy process.
For $j = 1, 2, 3$, denote by $Y_j^+ = \{Y_j^+(t)\}_{t\in\mathbb{R}}$ and $Y_j^- = \{Y_j^-(t)\}_{t\in\mathbb{R}}$ processes with representation

$$Y_j^+(t) = \int_{-\infty}^{\infty} f^+(t-s)\, dL_j(s), \quad Y_j^-(t) = \int_{-\infty}^{\infty} f^-(t-s)\, dL_j(s) \quad \text{for } t \in \mathbb{R},$$

which are well-defined by Proposition 1.1.7. We prove the case $i = 1$.

Step 2.1. Determination of $\Theta^{(1)}$.
By (1.4.4) we have

$$f^+ Z_k^{(1)} \le Y_1^+(\Gamma_k^{(1)} + \alpha) \le f^+ Z_k^{(1)} + \widetilde{Y}_1^+(\alpha) \quad \text{a.s.},$$

where $\widetilde{Y}_1^+(\alpha) \stackrel{d}{=} Y_1^+(0)$. We can estimate $Y(\Gamma_k^{(1)} + \alpha)$ from above by

$$\begin{aligned} Y(\Gamma_k^{(1)} + \alpha) &\le Y_1^+(\Gamma_k^{(1)} + \alpha) + Y_2^-(\Gamma_k^{(1)} + \alpha) + Y_3(\Gamma_k^{(1)} + \alpha) \\ &\le f^+ Z_k^{(1)} + \widetilde{Y}_1^+(\alpha) + Y_2^-(\Gamma_k^{(1)} + \alpha) + Y_3(\Gamma_k^{(1)} + \alpha) \quad \text{a.s..} \qquad (1.4.5) \end{aligned}$$

Choose

$$\Theta^{(1)} := Y_1^+(0) + Y_2^-(0) + Y_3(0).$$

By the independence of $\Gamma^{(1)}$, Y_2^- and Y_3 as well as the stationarity of $\widetilde{Y}_1^+, Y_2^-$ and Y_3 we obtain, similarly to (1.4.2), for $x \in \mathbb{R}$,

$$\mathbb{P}(\widetilde{Y}_1^+(\alpha) + Y_2^-(\Gamma_k^{(1)} + \alpha) + Y_3(\Gamma_k^{(1)} + \alpha) > x) = \mathbb{P}(\Theta^{(1)} > x). \qquad (1.4.6)$$

Thus by (1.4.5) and (1.4.6) for $x \in \mathbb{R}$,

$$\mathbb{P}(Y(\Gamma_k^{(1)} + \alpha) - f^+ Z_k^{(1)} > x) \leq \mathbb{P}(\Theta^{(1)} > x).$$

Taking Remark 1.3.5 into account, we have $Y_2^-(0)$, $Y_1^+(0) + Y_3(0) \in \mathcal{S}$ and, by Corollary 1.3.8, for $x \to \infty$,

$$\begin{aligned} \mathbb{P}(Y_1^+(0) + Y_3(0) > x) &= o(\mathbb{P}(f^+ Z_1^{(1)} > x)), \\ \mathbb{P}(Y_2^-(0) > x) &= o(\mathbb{P}(f^- Z_1^{(2)} > x)). \end{aligned}$$

Hence with Proposition 1.1.2 *(iii)*, the subexponentiality and (L2) we obtain

$$\mathbb{P}(\Theta^{(1)} > x) = o(\mathbb{P}(f^+ Z_1^{(1)} > x)) \quad \text{for } x \to \infty. \tag{1.4.7}$$

Step 2.2. Determination of $\widetilde{\Theta}^{(1)}$.
A lower bound of $Y(\Gamma_k^{(1)} + \alpha) - f^+ Z_k^{(1)}$ can be obtained as follows:

$$f^+ Z_k^{(1)} - Y_1^-(\Gamma_k^{(1)} + \alpha) - Y_2^+(\Gamma_k^{(1)} + \alpha) + Y_3(\Gamma_k^{(1)} + \alpha) \leq Y(\Gamma_k^{(1)} + \alpha).$$

We choose

$$\widetilde{\Theta}^{(1)} := Y_1^-(0) + Y_2^+(0) - Y_3(0).$$

Then we estimate similarly to Step 2.1

$$\begin{aligned} &\mathbb{P}(Y(\Gamma_k^{(1)} + \alpha) - f^+ Z_k^{(1)} < -x) \\ &\quad \leq \mathbb{P}(-Y_1^-(\Gamma_k^{(1)} + \alpha) - Y_2^+(\Gamma_k^{(1)} + \alpha) + Y_3(\Gamma_k^{(1)} + \alpha) < -x) \\ &\quad = \mathbb{P}(\widetilde{\Theta}^{(1)} > x) = o(\mathbb{P}(f^+ Z_1^{(1)} > x)) \quad \text{for } x \to \infty. \end{aligned}$$

Step 2.3. We show: if $Y(0) \in \mathcal{S}$ and $t \notin O_1$ then

$$\mathbb{P}(Y(\Gamma_k^{(1)} + t) > x) = o(\mathbb{P}(f^+ Z_1^{(1)} > x)) \quad \text{for } x \to \infty.$$

$Y(0) \in \mathcal{S}$ if and only if $\Theta^{(1)} = Y_1^+(0) + Y_2^-(0) + Y_3(0) \in \mathcal{S}$ since by (1.3.1) the Lévy measures of $Y(0)$ and $\Theta^{(1)}$ coincide on $[f^+, \infty)$. Consequently, by (1.4.5), (1.4.7), $\mathbb{P}(f(t) Z_k^{(1)} > x) = o(\mathbb{P}(f^+ Z_1^{(1)} > x))$ for $x \to \infty$ and Proposition 1.1.2 *(iii)* it follows that

$$\mathbb{P}(Y(\Gamma_k^{(1)} + t) > x) \leq \mathbb{P}(f(t) Z_k^{(1)} + \Theta^{(1)} > x) = o(\mathbb{P}(f^+ Z_1^{(1)} > x)) \quad \text{for } x \to \infty. \quad \square$$

Proof of Theorem 1.4.1.
Recall that by (1.1.18) the normalizing constants of $Z_k^{(1)}$ are a_{T/μ_1}, b_{T/μ_1} and the normalizing constants of $Z_k^{(2)}$ are a_{T/μ_2}, b_{T/μ_2}. On the basis of Lemma 1.4.4 the conditions of Corollary 1.2.2 are satisfied and thus

$$\lim_{T\to\infty} \mathbb{P}(\kappa_{T/\mu_1}^{(1,l)}(I) \neq \widetilde{\kappa}_{T/\mu_1}^{(1)}(I)) = \lim_{T\to\infty} \mathbb{P}(\kappa_{T/\mu_2}^{(2,l)}(I) \neq \widetilde{\kappa}_{T/\mu_2}^{(2)}(I)) = 0.$$

Hence

$$\begin{aligned} &\mathbb{P}(\kappa_T(I) \neq \widetilde{\kappa}_T(I)) \\ &\leq \sum_{l=1}^{P^{(1)}} \mathbb{P}(\kappa_T^{(1,l)}(I) \neq \widetilde{\kappa}_T^{(1)}(I)) + \sum_{l=1}^{P^{(2)}} \mathbb{P}(\kappa_T^{(2,l)}(I) \neq \widetilde{\kappa}_T^{(2)}(I)) \stackrel{T\to\infty}{\longrightarrow} 0. \end{aligned}$$

By the independence of $\kappa_T^{(1)}$, $\kappa_T^{(2)}$ and $\kappa^{(1)}$, $\kappa^{(2)}$ also $\widetilde{\kappa}_T \stackrel{T\to\infty}{\Longrightarrow} \kappa$. A conclusion of Rootzén [131], Lemma 3.3 is $\kappa_T \stackrel{T\to\infty}{\Longrightarrow} \kappa$. □

Proof of Corollary 1.4.2.
Define for $k \in \mathbb{N}$,

$$\begin{aligned} Y_{2k-1}(t) &:= Y(\Gamma_k^{(1)} + t) = f_{2k-1}(t) Z_{2k-1} + \widetilde{Y}_{2k-1}(t), \qquad & \tau_{2k-1} &:= Y(\Gamma_k^{(1)} + \alpha_1^{(1)}), \\ Y_{2k}(t) &:= Y(\Gamma_k^{(2)} + t) = f_{2k}(t) Z_{2k} + \widetilde{Y}_{2k}(t), & \tau_{2k} &:= Y(\Gamma_k^{(2)} + \alpha_1^{(2)}), \end{aligned}$$

where $f_{2k-1}(t) := f(t)$, $f_{2k}(t) := -f(t)$, $Z_{2k-1} := Z_k^{(1)}$, $Z_{2k} := Z_k^{(2)}$ and

$$\widetilde{Y}_{2k-1}(t) := Y(\Gamma_k^{(1)} + t) - f(t) Z_k^{(1)}, \qquad \widetilde{Y}_{2k}(t) := Y(\Gamma_k^{(2)} + t) - f(t) Z_k^{(2)}.$$

Let α_{2k-1} be $\alpha_1^{(1)}$ and α_{2k} be $\alpha_1^{(2)}$ for $k \in \mathbb{N}$. The proof of *(a)-(c)* follows then by Lemma 1.2.7, Theorem 1.2.11, Theorem 1.4.1 and Lemma 1.4.4. Claims *(d)-(e)* are conclusions of *(a)-(c)*, so we shall only sketch the proof.

Without loss of generality we assume $m = 1$. Let $I = [s,t) \times (x,\infty) \times (y_1, y_2]$ and $I_\epsilon = [s,t) \times (x,\infty) \times (y_1 - \epsilon, y_1 + \epsilon] \cup [s,t) \times (x,\infty) \times (y_2 - \epsilon, y_2 + \epsilon]$. We define

$$\widetilde{K}_T = \sum_{k=0}^{\infty} \varepsilon_{\left((\Gamma_k^{(1)}+\alpha_1^{(1)})/T,\, a_T^{-1}(Y(\Gamma_k^{(1)}+\alpha_l^{(1)})-b_T),\, a_T^{-1}(Y(\Gamma_k^{(1)}+\alpha_l^{(1)})-b_T)\,\mathbf{1}_{\{f(t_1)=f^+\}}\right)}.$$

By *(a)* we obtain $\widetilde{K}_T \stackrel{T\to\infty}{\Longrightarrow} K$. But

$$\begin{aligned} &\mathbb{P}(\widetilde{K}_T(I) \neq K_T(I)) \\ &\leq \mathbb{P}(\widetilde{K}_T(I_\epsilon) > 0) + \sum_{k\in[Ts,Tt)} \mathbb{P}(a_T^{-1}(Y(\Gamma_k^{(1)} + \alpha_l^{(1)}) - b_T) > x, \\ &\qquad |a_T^{-1}(Y(\Gamma_k^{(1)} + \alpha_l^{(1)}) - b_T)\,\mathbf{1}_{\{f(t_1)=f^+\}} - a_T^{-1}(Y(\Gamma_k^{(1)} + t_1) - b_T)| > \epsilon). \end{aligned}$$

If $\{(Y(\Gamma_k^{(1)} + \alpha_l^{(1)}), Y(\Gamma_k + t_i))\}_{k\in\mathbb{Z}}$ had been stationary, then the result would have followed in the case $f(t_1) = f^+$ by (1.2.29) and in the case $f(t_1) \neq f^+$ by *(c)*. Analogously to Lemma 1.4.4 we can find an upper bound for the last inequality, which converges to zero as $T \to \infty$. □

1.4.2 The point process of local maxima

In this section we restrict ourselves to MA processes, where $P^{(1)} = 1$, $P^{(2)} = 0$, i. e. f has an unique extremum with $f^+ = f(\alpha)$. Consider for instance the positive Poisson shot noise process as given in Example 1.1.9 with non-increasing kernel function f, whose support is $\mathbb{R}_+$. Then Y is non-increasing between consecutive jumps. Thus the process has a local supremum at point t if and only if $t \in \Gamma$. In some sense, this property is also valid for the subexponential Lévy driven MA processes of this section. Given a high threshold u_T, the asymptotic behavior of the supremum of Y in some neighborhood of $\Gamma_k + \alpha$ being larger than u_T is caused by $Y(\Gamma_k + \alpha)$ being above u_T. In the last section we have only shown the converse (cf. Corollary 1.4.2 and Remark 1.4.3): if $Y(\Gamma_k + \alpha)$ is asymptotically larger than u_T, then it is also a local supremum of Y. In this simpler model we can also prove the necessity of this condition.

Theorem 1.4.5
Let Y be the stationary MA process as given in (1.0.1) satisfying conditions (K1), (K2) with $P^{(1)} = 1$, $P^{(2)} = 0$, $O_1 = \{\alpha\}$ and a. s. sample paths in $\mathbb{D}(\mathbb{R})$, where L has the decomposition as given in (1.4.1). Further assume $\overline{f} : \mathbb{R} \to \mathbb{R}_+$ with

$$\overline{f}(t) = \sup_{-\infty<s<t} f^+(s)\,\mathbf{1}_{(-\infty,\alpha)}(t) + \sup_{t\le s<\infty} f^+(s)\,\mathbf{1}_{[\alpha,\infty)}(t)\,, \quad t \in \mathbb{R}$$

satisfies (K1). Suppose $L(1) \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$ with $a_T > 0$, $b_T \in \mathbb{R}$, $u_T = a_T x + b_T$ for $x \in \mathbb{R}$ such that

$$\lim_{T\to\infty} T\mathbb{P}(f^+ L(1) > u_T) = \exp(-x).$$

If f takes also negative values, we assume additionally (L2). Define

$$I_k = \left[\frac{\Gamma_{k-1}^{(1)} + \Gamma_k^{(1)}}{2} + \alpha, \frac{\Gamma_k^{(1)} + \Gamma_{k+1}^{(1)}}{2} + \alpha\right)$$

and $M_k = \sup_{t\in I_k} Y(t)$ *for* $k \in \mathbb{N}$. *For* $T > 0$ *denote the corresponding point processes by*

$$\kappa_T^M = \sum_{k=0}^{\infty} \varepsilon_{\left(s_k/T, a_T^{-1}(M_k - b_T)\right)}, \qquad \kappa_T = \sum_{k=0}^{\infty} \varepsilon_{\left((\Gamma_k^{(1)}+\alpha)/T, a_T^{-1}(Y(\Gamma_k^{(1)}+\alpha)-b_T)\right)},$$

where s_k *is a point in* $I_k \cup \{(\Gamma_k^{(1)} + \Gamma_{k+1}^{(1)})/2 + \alpha\}$ *satisfying* $Y(s_k) = M_k$. *Denote by* κ *a* PRM(ϑ) *with mean measure* $\vartheta(dt \times dx) = dt \times \mathrm{e}^{-x}\, dx$. *Let* $I = [s,t) \times (x, \infty) \subseteq S$. *Then* $\lim_{T\to\infty} \mathbb{P}(\kappa_T^M(I) \neq \kappa_T(I)) = 0$ *and*

$$\kappa_T^M \overset{T\to\infty}{\Longrightarrow} \kappa.$$

The marked point processes exhibit the same behavior as described in Corollary 1.4.2. Exceedances of Y are only caused by large positive jumps of the Lévy process. The size of an excess behaves like f^+ times the jump size of the driving Lévy process. Consider a positive Poisson shot noise process as given in Example 1.1.3 with non-increasing kernel function. Then $M_k = Y(\Gamma_k)$ and thus Theorem 1.4.1 and Theorem 1.4.5 coincide. The proof of an analogous result for $P^{(1)} > 1$ is much more involved as it is not possible to choose disjoint intervals I_k, whose length are identically distributed, as easily (see also Corollary 1.4.11).

Corollary 1.4.6
Let the assumptions and notations of Theorem 1.4.5 hold. Define $M(T) = \sup_{0\le t\le T} Y(t)$ *for* $T > 0$. *Then*

$$\lim_{T\to\infty} \mathbb{P}\left(a_T^{-1}(M(T) - b_T) \le x\right) = \exp(-\mathrm{e}^{-x}) \quad \text{for } x \in \mathbb{R}.$$

Proof.
Applying Theorem 1.4.5 yields

$$\begin{aligned} \mathbb{P}\left(a_T^{-1}(M(T) - b_T) \le x\right) &= \mathbb{P}(\kappa_T^M((0,1) \times (x,\infty)) = 0) \\ &\overset{T\to\infty}{\longrightarrow} \mathbb{P}(\kappa((0,1) \times (x,\infty)) = 0) \\ &= \exp(\vartheta((0,1) \times (x,\infty))) \\ &= \exp(-\mathrm{e}^{-x}). \end{aligned}$$

□

Essential for the proof of Theorem 1.4.5 is the following Lemma.

Lemma 1.4.7
Let the assumptions of Theorem 1.4.5 hold. Then we can find r. v. s Θ, $\widetilde{\Theta}$ *such that for* $x \in \mathbb{R}$

$$\begin{aligned} \mathbb{P}(M_k - f^+ Z_k^{(1)} > x) \le \mathbb{P}(\Theta > x), \\ \mathbb{P}(M_k - f^+ Z_k^{(1)} < -x) \le \mathbb{P}(\widetilde{\Theta} > x). \end{aligned} \tag{1.4.8}$$

Furthermore,

$$\mathbb{P}(\Theta > x) = o(\mathbb{P}(f^+ Z_1^{(1)} > x)) \quad \textit{and} \quad \mathbb{P}(\widetilde{\Theta} > x) = o(\mathbb{P}(f^+ Z_1^{(1)} > x)) \quad \textit{for } x \to \infty.$$

Moreover, the r. v. $M_k - f^+ Z_k^{(1)}$ *is independent of* $Z_k^{(1)}$.

Proof.
W. l. o. g. we assume f has support on $\mathbb{R}$.

Step 1. We first show that the Lemma is valid for the MA process Y_1.
Let $\{e_k\}_{k\in\mathbb{Z}}$ be an i. i. d. sequence of exponential r. v. s such that the jump times of L_1 are $\Gamma_k^{(1)} = \sum_{j=1}^k e_j$, $\Gamma_{-k}^{(1)} = \sum_{j=1}^k e_{-j}$ for $k \in \mathbb{N}$; recall $\mathbb{E}e_1 = 1/\mu_1$. Note that $\sup_{t\in\mathbb{R}} \overline{f}(t) = f(\alpha) = f^+$ and $\overline{f}$ is non-decreasing on $(-\infty, \alpha)$ and non-increasing on (α, ∞). Hence, by (K1) for $\overline{f}$ and Proposition 1.1.7, $\overline{Y}(t) = \int_{-\infty}^{\infty} \overline{f}(t-s)\, dL_1(s)$, $t \in \mathbb{R}$, is a well-defined stationary i. d. process. Define now the kernel functions

$$f_1(t) = \overline{f}(t)\, \mathbf{1}_{[\alpha,\infty)}(t), \qquad f_2(t) = \overline{f}(t)\, \mathbf{1}_{(-\infty,\alpha)}(t) \quad \text{for } t \in \mathbb{R} \tag{1.4.9}$$

and the corresponding MA processes

$$\begin{aligned} \overline{Y}_1(t) &= \int_{-\infty}^{\infty} f_1(t-s)\, dL_1(s) = \sum_{\Gamma_j^{(1)} \le t-\alpha} \overline{f}(t-\Gamma_j^{(1)}) Z_j^{(1)}, \\ \overline{Y}_2(t) &= \int_{-\infty}^{\infty} f_2(t-s)\, dL_1(s) = \sum_{t-\alpha < \Gamma_j^{(1)}} \overline{f}(t-\Gamma_j^{(1)}) Z_j^{(1)}, \quad \forall t \in \mathbb{R} \text{ a. s..} \end{aligned}$$

If $f(t) = 0$ for $t < \alpha$ then $\overline{Y}_2 \equiv 0$. Let $k > 0$ be fixed. Then for $t \in I_k$ we have $t - \alpha \in \left[\Gamma_{k-1}^{(1)}, \Gamma_{k+1}^{(1)}\right)$. Moreover,

$$\begin{aligned} \overline{Y}(t) &= \sum_{j\le k-2} \overline{f}(t-\Gamma_j^{(1)}) Z_j^{(1)} + \overline{f}(t-\Gamma_{k-1}^{(1)}) Z_{k-1}^{(1)} \\ &\quad + \overline{f}(t-\Gamma_k^{(1)}) Z_k^{(1)} + \overline{f}(t-\Gamma_{k+1}^{(1)}) Z_{k+1}^{(1)} + \sum_{k+2\le j} \overline{f}(t-\Gamma_j^{(1)}) Z_j^{(1)} \quad \forall t \in \mathbb{R} \text{ a. s..} \end{aligned} \tag{1.4.10}$$

Now, we find an upper bound for all terms on the right hand side. At first,

$$\overline{f}(t-\Gamma_k^{(1)})Z_k^{(1)} \leq f^+ Z_k^{(1)}. \tag{1.4.11}$$

The function $\overline{f} : [\alpha, \infty) \to \mathbb{R}_+$ is non-increasing and $\overline{Y}_1$ has a. s. sample paths in $\mathbb{D}(\mathbb{R})$, so we have for $t \in I_k$,

$$\begin{aligned}\sum_{j\leq k-2} \overline{f}(t-\Gamma_j^{(1)})Z_j^{(1)} &\leq \sum_{j<k-1} \overline{f}(\alpha+\Gamma_{k-1}^{(1)}-\Gamma_j^{(1)})Z_j^{(1)} = \overline{Y}_1(\alpha+\Gamma_{k-1}^{(1)}-) \\ &= \sum_{j\leq k-2} \overline{f}\left(\alpha+\sum_{l=j+1}^{k-1} e_l\right) Z_j^{(1)}. \end{aligned} \tag{1.4.12}$$

Since $\overline{f} : (-\infty, \alpha) \to \mathbb{R}_+$ is non-decreasing, it holds for $t \in I_k$,

$$\begin{aligned}\sum_{k+2\leq j} \overline{f}(t-\Gamma_j^{(1)})Z_j^{(1)} &\leq \sum_{k+2\leq j} \overline{f}\left(\alpha+\Gamma_{k+1}^{(1)}-\Gamma_j^{(1)}\right) Z_j^{(1)} = \overline{Y}_2(\alpha+\Gamma_{k+1}^{(1)}) \\ &= \sum_{k+2\leq j} \overline{f}\left(\alpha-\sum_{l=k+2}^{j} e_l\right) Z_j^{(1)}. \end{aligned} \tag{1.4.13}$$

For the two remaining terms of (1.4.10) we need a finer upper bound than those derived in (1.4.11)-(1.4.13). By monotonicity of $\overline{f}$ for $t \in I_k$ we estimate

$$\begin{aligned}\overline{f}(t-\Gamma_{k-1}^{(1)})Z_{k-1}^{(1)} &\leq \overline{f}\left(\alpha+\frac{\Gamma_{k-1}^{(1)}+\Gamma_k^{(1)}}{2}-\Gamma_{k-1}^{(1)}\right) Z_{k-1}^{(1)} = \overline{f}\left(\alpha+\tfrac{e_k}{2}\right) Z_{k-1}^{(1)}, \\ \overline{f}(t-\Gamma_{k+1}^{(1)})Z_{k+1}^{(1)} &\leq \overline{f}\left(\alpha+\frac{\Gamma_k^{(1)}+\Gamma_{k+1}^{(1)}}{2}-\Gamma_{k+1}^{(1)}\right) Z_{k+1}^{(1)} = \overline{f}\left(\alpha-\tfrac{e_{k+1}}{2}\right) Z_{k+1}^{(1)}. \end{aligned} \tag{1.4.14}$$

If we now define

$$\theta_k^{(1)} := \overline{Y}_1(\alpha+\Gamma_{k-1}^{(1)}-) + \overline{f}\left(\alpha+\frac{e_k}{2}\right) Z_{k-1}^{(1)} + \overline{f}\left(\alpha-\frac{e_{k+1}}{2}\right) Z_{k+1}^{(1)} + \overline{Y}_2(\alpha+\Gamma_{k+1}^{(1)}),$$

then we conclude by (1.4.10)-(1.4.14) that

$$M_k^{(1)} := \sup_{t\in I_k} Y_1(t) \leq \sup_{t\in I_k} \overline{Y}(t) \leq f^+ Z_k^{(1)} + \theta_k^{(1)}, \tag{1.4.15}$$

and the two r. v. s on the right hand side are independent. We calculate now an upper bound in distribution of $\theta_k^{(1)}$, which is independent of k. Observing (1.4.9) and using (1.4.4) we obtain

$$\begin{aligned}\overline{Y}_2(\alpha+\Gamma_{k+1}^{(1)}) &= \overline{Y}_2(\alpha+\Gamma_{k+1}^{(1)}) + f_2(\alpha+\Gamma_{k+1}^{(1)})Z_0^{(1)} \stackrel{d}{=} \overline{Y}_2(0) + f_2(\alpha)Z_k^{(1)} = \overline{Y}_2(0), \\ \overline{Y}_1(\alpha+\Gamma_{k-1}^{(1)}-) &+ f_1(\alpha+\Gamma_{k-1}^{(1)}-)Z_0^{(1)} \stackrel{d}{=} \overline{Y}_1(0) + f_1(\alpha-)Z_k^{(1)} = \overline{Y}_1(0). \end{aligned} \tag{1.4.16}$$

Define $\widetilde{Y}(t)=\int_{-\infty}^{\infty}\overline{f}(\alpha+t-s)\,d\widetilde{L}(s)$, where $\widetilde{L}$ is a compound Poisson process with Lévy measure $2\mu_1\mathbb{P}(Z_1^{(1)}>x)$ for $x\in\mathbb{R}$ such that

$$\overline{f}\left(\alpha+\frac{e_k}{2}\right)Z_{k-1}^{(1)}+\overline{f}\left(\alpha-\frac{e_{k+1}}{2}\right)Z_{k+1}^{(1)}\leq\widetilde{Y}(0) \tag{1.4.17}$$

and $\widetilde{Y}(0)$ is independent of $\overline{Y}_1(\alpha+\Gamma_{k-1}^{(1)}-)$ and $\overline{Y}_2(\alpha+\Gamma_{k+1}^{(1)})$. From Proposition 1.3.1 we know that $\widetilde{Y}(0)\in\mathcal{S}$ with

$$\mathbb{P}(\widetilde{Y}(0)>x)\sim 2\mathbb{P}(\overline{Y}(0)>x)\quad\text{ for } x\to\infty.$$

Define $\Theta_1=\overline{\overline{Y}}+\widetilde{\widetilde{Y}}$, where $\overline{\overline{Y}}$, $\widetilde{\widetilde{Y}}$ are independent with $\overline{\overline{Y}}\stackrel{d}{=}\overline{Y}(0)$ and $\widetilde{\widetilde{Y}}\stackrel{d}{=}\widetilde{Y}(0)$. By Proposition 1.1.2 *(iii)* we have $\Theta_1\in\mathcal{S}$ with tail behavior $\mathbb{P}(\Theta_1>x)\sim 3\mathbb{P}(\overline{Y}(0)>x)$ for $x\to\infty$. Taking Corollary 1.3.8 into account we conclude

$$\mathbb{P}(\Theta_1>x)=o(\mathbb{P}(f^+Z_1^{(1)}>x))\quad\text{ for } x\to\infty. \tag{1.4.18}$$

Having (1.4.16) and (1.4.17) in mind we obtain

$$\mathbb{P}(\theta_k^{(1)}>x)\leq\mathbb{P}(\overline{Y}_1(0)+\overline{Y}_2(0)+\widetilde{\widetilde{Y}}>x)=\mathbb{P}(\Theta_1>x)\text{ for } x\in\mathbb{R}.$$

Hence by (1.4.15) and (1.4.18) we get for $x\to\infty$,

$$\mathbb{P}(M_k^{(1)}-f^+Z_k^{(1)}>x)\leq\mathbb{P}(\theta_k^{(1)}>x)\leq\mathbb{P}(\Theta_1>x)=o(\mathbb{P}(f^+Z_1^{(1)}>x)). \tag{1.4.19}$$

Step 2. For arbitrary Y we use the decomposition (1.0.3).
Step 2.1. Determination of Θ.
By (1.4.15) we have

$$M_k\leq M_k^{(1)}+\sup_{t\in I_k}[Y_2(t)+Y_3(t)]\leq f^+Z_k^{(1)}+\theta_k^{(1)}+\sup_{t\in I_k}[Y_2(t)+Y_3(t)]. \tag{1.4.20}$$

Further, we set $\theta_k:=\theta_k^{(1)}+\sup_{t\in I_k}[Y_2(t)+Y_3(t)]$ for $k\in\mathbb{N}$. Writing the interval I_k as $\Big[\alpha+\Gamma_{k-1}^{(1)}+1/2e_k\,,\,\alpha+\Gamma_{k-1}^{(1)}+e_k+1/2e_{k+1}\Big)$, we see that the dependence of the two quantities on the right hand side is only given by e_k and e_{k+1}. As Y_2+Y_3 is a stationary process independent of $\Gamma^{(1)}$, we conclude that

$$\theta_k\stackrel{d}{=}\theta_k^{(1)}+\sup_{0\leq t<\frac{e_k+e_{k+1}}{2}}[Y_2(t)+Y_3(t)].$$

For $k\in\mathbb{N}$ define the r. v. s

$$W_k=\theta_k^{(1)}-\overline{f}\left(\alpha+\frac{e_k}{2}\right)Z_{k-1}^{(1)}-\overline{f}\left(\alpha-\frac{e_{k+1}}{2}\right)Z_{k+1}^{(1)}$$

independent of e_k and e_{k+1}. Let $\epsilon > 0$ be arbitrary. We shall show that

$$\begin{aligned}\mathbb{P}(\theta_k > x) &= \mathbb{P}(\theta_k > x, e_k \le \epsilon,\, e_{k+1} \le \epsilon) + \mathbb{P}(\theta_k > x, e_k > \epsilon,\, e_{k+1} \le \epsilon) \\ &\quad + \mathbb{P}(\theta_k > x, e_k \le \epsilon,\, e_{k+1} > \epsilon) + \mathbb{P}(\theta_k > x, e_k > \epsilon,\, e_{k+1} > \epsilon) \\ &\le \mathbb{P}(\Theta > x) = o(\mathbb{P}(f^+ Z_1^{(1)} > x)) \quad \text{for } x \to \infty. \end{aligned} \tag{1.4.21}$$

To this end, note that the first term on the right hand side satisfies the inequality

$$\begin{aligned}\mathbb{P}(\theta_k > x, e_k \le \epsilon,\, e_{k+1} \le \epsilon) &\le \mathbb{P}(\theta_k^{(1)} + \sup_{0\le t\le\epsilon} [Y_2(t) + Y_3(t)] \ge x) \\ &\le \mathbb{P}(\Theta_1 + \sup_{0\le t\le\epsilon} [Y_2(t) + Y_3(t)] \ge x), \end{aligned} \tag{1.4.22}$$

where (1.4.19) was used. Using Theorem 1.3.9 *(b)*, if $\nu(-\infty, -1] > 0$, we conclude that $\sup_{0\le t\le\epsilon}[Y_2(t) + Y_3(t)] \in \mathcal{S}$, and as $f^- < f^+$ for $x \to \infty$,

$$\mathbb{P}(\sup_{0\le t\le\epsilon} [Y_2(t) + Y_3(t)] > x) = O(\mathbb{P}(f^- L_2(1) > x)) = o(\mathbb{P}(f^+ Z_1^{(1)} > x)).$$

If $\nu(-\infty, -1] = 0$ this holds by Lemma 1.3.11. Since (1.4.18) holds, as a consequence of (1.4.22) and Proposition 1.1.2 *(iii)* we obtain

$$\mathbb{P}(\theta_k > x, e_k \le \epsilon,\, e_{k+1} \le \epsilon) = o(\mathbb{P}(f^+ Z_1^{(1)} > x)) \quad \text{for } x \to \infty. \tag{1.4.23}$$

Next we estimate the second term of (1.4.21):

$$\begin{aligned}&\mathbb{P}(\theta_k > x, e_k > \epsilon,\, e_{k+1} \le \epsilon) \\ &\le \mathbb{P}\left(W_k + \overline{f}\left(\alpha + \frac{\epsilon}{2}\right) Z_{k-1}^{(1)} + \overline{f}\left(\alpha - \frac{e_{k+1}}{2}\right) Z_{k+1}^{(1)} + \sup_{0\le t\le(\epsilon+e_k)/2} [Y_2(t) + Y_3(t)] > x\right).\end{aligned}$$

Using again Theorem 1.3.9 *(b)* in the case $\nu(-\infty, -1] > 0$ we have $\sup_{0\le t\le(\epsilon+e_k)/2}[Y_2(t) + Y_3(t)] \in \mathcal{S}$ and

$$\mathbb{P}\left(\sup_{0\le t\le(\epsilon+e_k)/2} [Y_2(t) + Y_3(t)] > x\right) = O(\mathbb{P}(f^- L_2(1) > x)) \quad \text{for } x \to \infty.$$

Hence by Proposition 1.1.2 *(iii)* $\overline{f}\left(\alpha + \frac{\epsilon}{2}\right) Z_{k-1}^{(1)} + \sup_{0\le t\le(\epsilon+e_k)/2}[Y_2(t) + Y_3(t)] \in \mathcal{S}$ and taking $f^+ > \max\{f(\alpha-\epsilon/2), f(\alpha+\epsilon/2), f^-\}$ into account we obtain for $x \to \infty$,

$$\mathbb{P}\left(\overline{f}\left(\alpha + \frac{\epsilon}{2}\right) Z_{k-1}^{(1)} + \sup_{0\le t\le(\epsilon+e_k)/2} [Y_2(t) + Y_3(t)] > x\right) = o(\mathbb{P}(f^+ Z_1^{(1)} > x)).$$

Note, that $\mathbb{P}(W_k + \overline{f}\left(\alpha - \frac{e_{k+1}}{2}\right) > x) \le \mathbb{P}(\Theta_1 > x)$, which is also independent of k. Thus again by Proposition 1.1.2 *(iii)* we obtain for $x \to \infty$,

$$\mathbb{P}(\theta_k > x, e_k > \epsilon,\ e_{k+1} \le \epsilon) = o(\mathbb{P}(f^+ Z_1^{(1)} > x)). \tag{1.4.24}$$

By symmetry also the third term of (1.4.21) satisfies for $x \to \infty$

$$\mathbb{P}(\theta_k > x, e_k \le \epsilon,\ e_{k+1} > \epsilon) = o(\mathbb{P}(f^+ Z_1^{(1)} > x)). \tag{1.4.25}$$

In order to compute an upper bound for the last term of (1.4.21) we decouple the dependence using the monotonicity of $\overline{f}$ and making this part independent of e_k and e_{k+1}. Then we use $(e_k + e_{k+1})/2 \overset{d}{=} e_k$ and the same argumentation as above to show

$$\begin{aligned}
&\mathbb{P}(\theta_k > x, e_k > \epsilon,\ e_{k+1} > \epsilon) \\
&\le \mathbb{P}\left(W_k + \overline{f}\left(\alpha + \frac{\epsilon}{2}\right) Z_{k-1}^{(1)} + \overline{f}\left(\alpha - \frac{\epsilon}{2}\right) Z_{k+1}^{(1)} + \sup_{0 \le t \le (e_k + e_{k+1})/2} [Y_2(t) + Y_3(t)] > x\right) \\
&= o(\mathbb{P}(f^+ Z_1^{(1)} > x)) \quad \text{for } x \to \infty.
\end{aligned} \tag{1.4.26}$$

Combining the results (1.4.23)-(1.4.26) proves (1.4.21). Then, by (1.4.20) and (1.4.21) there exists a r. v. Θ, independent of k, such that for $x \to \infty$

$$\mathbb{P}(M_k - f^+ Z_k^{(1)} > x) \le \mathbb{P}(\theta_k > x) \le \mathbb{P}(\Theta > x) = o(\mathbb{P}(f^+ Z_1^{(1)} > x)).$$

Step 2.2. Determination of $\widetilde{\Theta}$.
Using $\Gamma_k^{(1)} + \alpha \in I_k$, Lemma 1.4.4 and the notation there we get

$$\mathbb{P}(M_k - f^+ Z_k^{(1)} < -x) \le \mathbb{P}(Y(\Gamma_k^{(1)} + \alpha) - f^+ Z_k^{(1)} < -x) \le \mathbb{P}(\widetilde{\Theta}^{(1)} > x),$$

where $\mathbb{P}(\widetilde{\Theta}^{(1)} > x) = o(\mathbb{P}(f^+ Z_1^{(1)} > x))$ for $x \to \infty$. Thus we choose $\widetilde{\Theta} := \widetilde{\Theta}^{(1)}$. □

Proof of Theorem 1.4.5.
Denote by $\widetilde{\kappa}_T^{(1)}$ the point process of Theorem 1.4.1. The result $\kappa_T^M \overset{T\to\infty}{\Longrightarrow} \kappa$ and $\lim_{T\to\infty} \mathbb{P}(\kappa_T^M(I) \ne \widetilde{\kappa}_T^{(1)}(I)) = 0$ is a conclusion of Lemma 1.4.7 and Corollary 1.2.2. Thus by Theorem 1.4.1

$$\mathbb{P}(\kappa_T^M(I) \ne \kappa_T(I)) \le \mathbb{P}(\kappa_T^M(I) \ne \widetilde{\kappa}_T^{(1)}(I)) + \mathbb{P}(\widetilde{\kappa}_T^{(1)}(I) \ne \kappa_T(I)) \overset{T\to\infty}{\longrightarrow} 0.$$

□

1.4.3 Normalizing constants of running maxima

In this section we shall calculate the normalizing constants of running maxima of a subexponential Lévy driven MA process in MDA(Λ). For a Poisson shot noise process with non-negative, non-increasing kernel function, they have been calculated by Lebedev [98] (see also Corollary 1.4.6).

Theorem 1.4.8
Let Y be the stationary MA process given in (1.0.1) satisfying the conditions (K1) and (K2) with a. s. sample paths in $\mathbb{D}(\mathbb{R})$. Suppose $L(1) \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$ with $a_T > 0$, $b_T \in \mathbb{R}$, $u_T = a_T x + b_T$ for $x \in \mathbb{R}$ such that

$$\lim_{T\to\infty} T\mathbb{P}(f^+ L(1) > u_T) = \exp(-x).$$

If f takes negative values, we assume additionally (L2). For $T > 0$ let be $M(T) = \sup_{0\le t\le T} Y(t)$. Then

$$\lim_{T\to\infty} \mathbb{P}\left(a_T^{-1}(M(T) - b_T) \le x\right) = \exp[-(1 + (1-p)/p\, \mathbf{1}\{f^- = f^+\})\, \mathrm{e}^{-x}] \quad \textit{for } x \in \mathbb{R}.$$

Proof.
Let $c_n = \sup_{t\in[n-1,n+1)} f^+(t)$ and $d_n = \sup_{t\in[n-1,n+1)} f^-(t)$ for $n \in \mathbb{Z}$. From (K1) we conclude $\sum_{n=-\infty}^{\infty} c_n^\delta < \infty$ and $\sum_{n=-\infty}^{\infty} d_n^\delta < \infty$. Define the kernel functions

$$\overline{f}(t) = \sum_{k=-\infty}^{\infty} \sup_{t\in[k-1,k)} f^+(t)\, \mathbf{1}_{[k-1,k)}(t), \quad \underline{f}(t) = \sum_{k=-\infty}^{\infty} \sup_{t\in[k-1,k)} f^-(t)\, \mathbf{1}_{[k-1,k)}(t)$$

for $t \in \mathbb{R}$ and the discrete-time MA processes

$$\begin{aligned} X_n^{(1)} &= \sum_{k=-\infty}^{\infty} c_{n-k}[L_1(k) - L_1(k-1)] = \sum_{k=-\infty}^{\infty} c_{n-k}\xi_k^{(1)}, \\ X_n^{(2)} &= \sum_{k=-\infty}^{\infty} d_{n-k}[L_2(k) - L_2(k-1)] = \sum_{k=-\infty}^{\infty} d_{n-k}\xi_k^{(2)} \end{aligned}$$

for $n \in \mathbb{N}$, where $\xi_k^{(i)} = L_i(k) - L_i(k-1)$ for $k \in \mathbb{Z}$, $i = 1, 2$. Let

$$X_n^{(3)} = \sup_{t\in[n-1,n)} \int_{-\infty}^{\infty} f(t-s)\, dL_3(s),$$

and $M_n = \sup_{t\in[n-1,n)} Y(t)$ for $n \in \mathbb{N}$. Both $X_n^{(1)}, X_n^{(2)}, X_n^{(3)}$ and M_n are finite a. s., since Y has a. s. sample path in $\mathbb{D}(\mathbb{R})$ and Proposition 1.1.7. As L_1, L_2 are increasing,

$$f(t) \le \overline{f}(t) \quad \text{and} \quad -f(t) \le \underline{f}(t) \quad \text{for all } t \in \mathbb{R},$$

we have for $t \in [n-1, n)$

$$\begin{aligned} Y(t) &\leq \int_{-\infty}^{\infty} \overline{f}(t-s)\, dL_1(s) + \int_{\infty}^{\infty} \underline{f}(t-s)\, dL_2(s) + X_n^{(3)} \\ &= \sum_{k=-\infty}^{\infty} \sup_{s \in [k-1,k)} f^+(s)[L_1(t-(k-1)) - L_1(t-k)] \\ &\quad + \sum_{k=-\infty}^{\infty} \sup_{s \in [k-1,k)} f^-(s)[L_2(t-(k-1)) - L_2(t-k)] + X_n^{(3)} \\ &\leq X_n^{(1)} + X_n^{(2)} + X_n^{(3)} =: X_n, \end{aligned} \tag{1.4.27}$$

i.e. $M_n \leq X_n$. Since $\{\xi_k^{(i)}\}_{k\in\mathbb{N}}$ is an i.i.d. sequence with $\xi_k^{(i)} \stackrel{d}{=} L_i(1)$ and $X^{(i)} = \{X_n^{(i)}\}_{n\in\mathbb{N}}$ is a discrete-time MA process, which satisfies the assumptions of Example 1.2.6, we obtain for $i = 1, 2$

$$\kappa_T^{(i)} = \sum_{k=1}^{\infty} \varepsilon_{\left(k/T, a_T^{-1}(X_k^{(i)} - b_T)\right)} \stackrel{T\to\infty}{\Longrightarrow} \widetilde{P}^{(i)} \kappa^{(i)}$$

with $\kappa^{(i)}$ as given in Theorem 1.4.1. Furthermore, $\widetilde{P}^{(1)} = \text{card}\{k : c_k = f^+\} < \infty$ and $\widetilde{P}^{(2)} = \text{card}\{k : d_k = f^+\} < \infty$. The processes $X^{(1)}$, $X^{(2)}$ are independent. The fact that $\mathbb{P}(X_n^{(i)} > x) \sim \widetilde{P}^{(i)}\mathbb{P}(f^+ L_i(1) > x)$ for $x \to \infty$, Proposition 1.2.5 (if $f^- = f^+$) and Theorem 1.2.1 (if $f^- < f^+$) we conclude

$$\sum_{k=1}^{\infty} \varepsilon_{\left(k/T, a_T^{-1}(X_k^{(1)} + X_k^{(2)} - b_T)\right)} \stackrel{T\to\infty}{\Longrightarrow} \widetilde{P}^{(1)} \kappa^{(1)} + \widetilde{P}^{(2)} \kappa^{(2)}.$$

Using Lemma 1.3.11 and Proposition 1.1.2 *(i)* we also have for $x \to \infty$

$$\mathbb{P}(X_k^{(3)} > x) = o(\mathbb{P}(f^+ L(1) > x)) \quad \text{and} \quad \mathbb{P}(X_k^{(3)} < -x) = o(\mathbb{P}(f^+ L(1) > x)).$$

Applying Theorem 1.2.1 yields

$$\sum_{k=1}^{\infty} \varepsilon_{\left(k/T, a_T^{-1}(X_k^{(1)} + X_k^{(2)} + X_k^{(3)} - b_T)\right)} \stackrel{T\to\infty}{\Longrightarrow} \widetilde{P}^{(1)} \kappa^{(1)} + \widetilde{P}^{(2)} \kappa^{(2)}.$$

Thus, for $I = (0, 1] \times (x, \infty)$ we have on the one hand with (1.4.27)

$$\begin{aligned} \lim_{T\to\infty} \mathbb{P}(a_T^{-1}(M(T) - b_T) \leq x) &\geq \lim_{T\to\infty} \mathbb{P}\left(\sum_{k=1}^{\infty} \varepsilon_{\left(k/T, a_T^{-1}(X_k^{(1)} + X_k^{(2)} + X_k^{(3)} - b_T)\right)}(I) = 0\right) \\ &= \mathbb{P}(\widetilde{P}^{(1)} \kappa^{(1)}(I) + \widetilde{P}^{(2)} \kappa^{(2)}(I) = 0) \qquad (1.4.28) \\ &= \mathbb{P}(\kappa^{(1)}(I) = 0)[\mathbf{1}_{\{f^- < f^+\}} + \mathbf{1}_{\{f^- = f^+\}}\, \mathbb{P}(\kappa^{(2)}(I) = 0)]. \end{aligned}$$

On the other hand, Theorem 1.4.1 gives

$$\begin{aligned}\lim_{T\to\infty}\mathbb{P}(a_T^{-1}(M(T)-b_T)\le x) &\le \lim_{T\to\infty}\mathbb{P}(\kappa_T(I)=0)\\ &=\mathbb{P}(P^{(1)}\kappa^{(1)}(I)+P^{(2)}\kappa^{(2)}(I)=0)\\ &=\mathbb{P}(\kappa^{(1)}(I)=0)[\mathbf{1}_{\{f^-<f^+\}}+\mathbf{1}_{\{f^-=f^+\}}\,\mathbb{P}(\kappa^{(2)}(I)=0)].\end{aligned}\tag{1.4.29}$$

Taking

$$\begin{aligned}&\mathbb{P}(\kappa^{(1)}(I)=0)[\mathbf{1}_{\{f^-<f^+\}}+\mathbf{1}_{\{f^-=f^+\}}\,\mathbb{P}(\kappa^{(2)}(I)=0)]\\ &\quad=\exp[-(1+(1-p)/p\,\mathbf{1}\{f^-=f^+\})\,\mathrm{e}^{-x}]\end{aligned}$$

into account we obtain by (1.4.28) and (1.4.29) the result. □

Remark 1.4.9

If f is flat in its maximum and either $f^- < f^+$ or f is also flat in its minimum $-f^+$, the convergence of running maxima of Y is also ensured. Following the proof of Theorem 1.4.8 line by line and replacing the suprema in $X_n^{(i)}$ by the infima, a lower bound of M_n can be found, without using Theorem 1.4.1. □

1.4.4 Extremal index function

The question arises, what influence the dependence structure of a stochastic process has to its extremal behavior. In the case of an i. i. d. sequence $\{Y_k\}_{k\in\mathbb{Z}}$ there is the well known result about Poisson approximation (cf. Leadbetter et al. [94], Theorem 1.5.1), i. e. for some $0\le\tau\le\infty$ and a sequence of constants $\{u_n\}_{n\in\mathbb{N}}$ in $\mathbb{R}$ we have $\lim_{n\to\infty} n\overline{F}(u_n)=\tau$ if and only if

$$\lim_{n\to\infty}\mathbb{P}(\max_{k=1,\dots,n} Y_k\le u_n)=\exp(-\tau).$$

This also holds for weakly dependent data. Weak dependence is then expressed in the $D(u_n)$ and $D'(u_n)$ condition (Definition A.3.1). Strong dependence may result in clustering within short time intervals. A measure for the dependence in the extremes of a discrete-time process is the *extremal index* θ (Definition A.1.9, a detailed discussion about the extremal index can be found in Leadbetter et al. [94], pp. 67). In the case of an i. i. d. sequence or when the process is weakly dependent the extremal index is one (Leadbetter et al. [94], Theorem 3.5.2). These holds for Gaussian processes whose covariance function γ satisfies the *Berman's condition* $\lim_{h\to\infty}\log(h)\gamma(h) = 0$ (Leadbetter et al. [95], Theorem 4.3.3). The extremal index can only take values in $[0,1]$. If $\theta \in (0,1]$, then the normalized maxima of the stationary sequence and the normalized maxima of the associated i. i. d. sequence have the same limiting distribution. The case $\theta = 0$ is degenerate, where both normalized maxima can not have the same limiting distribution. The extremal index is a quantity, which allows us to characterize the relationship between the dependence structure of the data and their extremal behavior. The extremal index can be interpreted as the reciprocal of the mean cluster size of the limit process of point processes of exceedances. The value $\theta < 1$ indicates that the limit is a compound Poisson process, i. e. exceedances over high thresholds tend to occur in clusters; $\theta = 1$ meets a Poisson process. A summary about the extremal index and its estimation can be found in Embrechts et al. [60], Section 8.1.

Continuous-time processes are highly dependent in small time intervals by the continuity of the process. Thus it is not adequate to adopt the notation of an extremal index of a discrete-time process. As a more appropriate object to capture the extreme dependence structure of a continuous-time process we define a function with similar properties.

Definition 1.4.10
Let $\{Y(t)\}_{t\geq 0}$ be a stationary process. Define the r. v. s $M_k^{(h)} = \sup_{(k-1)h\leq t\leq kh} Y(t)$ for $k \in \mathbb{N}$, $h > 0$. Let $\theta(h)$ be the extremal index of the sequence $\{M_k^{(h)}\}_{k\in\mathbb{N}}$. Then we call the function $\theta : \mathbb{R}_+ \to [0,1]$ extremal index function.

Dividing the positive real line into blocks of length h, the extremal index function is a measure for the expected cluster sizes of such blocks. By building these blocks, on the one hand the natural dependence of a continuous-time process in small intervals is neglected. On the other hand we can choose the blocks arbitrarily small. Gaussian processes, whose covariance function satisfies the Berman's condition have $\theta(h) \equiv 1$ (Leadbetter et al. [95], Theorem 12.2.9 and Theorem 12.3.4), e. g. Gaussian CARMA processes.

Corollary 1.4.11
Let Y be a stationary MA process given by (1.0.1) satisfying the assumptions of Theorem 1.4.8. Denote by $d_i = \alpha^{(i)}_{P^{(i)}} - \alpha^{(i)}_1$ for $i = 1, 2$. Then the extremal index

function is

$$\theta(h) = h\frac{1 + (1-p)/p\,\mathbf{1}\{f^- = f^+\}}{[h+d_1] + (1-p)/p\,\mathbf{1}\{f^- = f^+\}[h+d_2]} \qquad \textit{for } h > \max\{d_1, d_2\},$$

where $\lim_{h\to\infty}\theta(h) = 1$.

Proof.
For $h > 0$ let $M_k^{(h)} = \sup_{(k-1)h \le t \le kh} Y(t)$ for $k \in \mathbb{N}$. On the one hand Theorem 1.3.9 *(b1)* and Remark 1.3.10 yields for $n \to \infty$

$$n\mathbb{P}(M(h) > u_{nh}) \sim \exp(-x)\left[[h+d_1] + (1-p)/p\,\mathbf{1}\{f^- = f^+\}[h+d_2]\right]/h.$$

On the other hand we get by Theorem 1.4.8

$$\begin{aligned}\lim_{n\to\infty}\mathbb{P}(\max_{k=1,\dots,n} M_k^{(h)} \le u_{nh}) &= \lim_{n\to\infty}\mathbb{P}(a_{nh}^{-1}(M(nh) - b_{nh}) \le x)\\ &= \exp[-\mathrm{e}^{-x}(1 + (1-p)/p\,\mathbf{1}\{f^- = f^+\})]. \qquad \square\end{aligned}$$

If the kernel function has only one maximum and at most one minimum with value f^+, then the extremal index function is constant one. Dividing the positive real line into blocks of the same length, results in no clusters of exceedances of these blocks. In contrast to this, if there is more than one maximum or minimum, then $\theta(h) < 1$ for every $h > \max\{d_1, d_2\}$. Thus the dependence of the process results in clusters of exceedances of blocks, where the mean cluster size tends to 1, as h tends to ∞. This is obvious as cluster sizes will be smaller, because more data are condensed into one block. This result is also in analogy to discrete-time MA processes, where clusters only occur, when the kernel function has more than one maximum or minimum with value f^+ (Example 1.2.6).

We shall give an intuitive explanation. A large positive jump of the Lévy process in $\Gamma_k^{(1)}$ results in large values of $Y(\Gamma_k^{(1)} + \alpha_i^{(1)})$, $i = 1, \dots, P^{(1)}$. Thus the process $\{\sup_{0\le s\le h} Y(t+s)\}_{t\ge 0}$ achieves in $J = \bigcup_{i=1}^{P^{(1)}}[\Gamma_k^{(1)} + \alpha_i^{(1)} - h, \Gamma_k^{(1)} + \alpha_i^{(1)}]$ a large value. If $h > d_1$, then the length of the interval J is $h + d_1$. Since the length of a block h is less than $h + d_1$, the interval J intersects at least two blocks. Thus we obtain clusters of exceedances of the stationary sequence $\{M_k^{(h)}\}$. If there is only one maximum of the kernel function ($d_1 = 0$), then the length of J is equal to the length of a block, which induces no clusters. Analogously is the explanation for negative jumps in combination with minima. Positive and negative jumps of the

Lévy process are independent, so that there may be at most one maximum and minimum to obtain still no clusters among the blocks.

Exceedances over thresholds are the natural events to be studied in discrete-time. A similar role are in continuous-time *upcrossings* above thresholds. For an a. s. continuous-time process Y an upcrossing of level u is a point t_0 for which $Y(t) < u$ when $t \in (t_0 - \epsilon, t_0)$ and $Y(t) \geq u$ when $t \in (t_0, t_0 + \epsilon)$ for some $\epsilon > 0$. Note, however, that in a finite interval there may be infinitely many upcrossings. Examples include not mean square differentiable Gaussian processes (Leadbetter et al. [95], p. 152). A related concept are point processes of ϵ-*upcrossings*. Poisson convergence of the point process of ϵ-upcrossings may be obtained (Leadbetter et al. [95], Theorem 12.4.2). For an a. s. continuous process Y an ϵ-*upcrossing* of level u is a point t_0 with $Y(t) < u$ when $t \in (t_0 - \epsilon, t_0)$ and $Y(t_0) = u$. Every ϵ-upcrossing is an upcrossing, while obviously an upcrossing need not to be an ϵ-upcrossing.

If the assumption of Theorem 1.4.5 holds, so that $P^{(1)} = 1$ and $P^{(2)} = 0$, the kernel function f is non-increasing, and the driving Lévy process is a positive compound Poisson process, then also the point process of upcrossings and ϵ-upcrossings converges to a Poisson process since upcrossings and ϵ-upcrossings occur only at positive jump times of the Lévy process in combination with the supremum of the kernel function (Theorem 1.4.5, Corollary 1.4.2). In contrast, if the kernel function has more than one extremum, then the point process of upcrossings and ϵ-upcrossings converges to a cluster Poisson process.

Chapter 2

Extremes of regularly varying Lévy driven mixed MA processes

In this chapter we investigate the extremal behavior of a stationary continuous-time *mixed moving average (MA) process* of the form

$$Y(t) = \int_{\mathbb{R}_+ \times \mathbb{R}} f(r, t-s)\, d\Lambda(r,s) \quad \text{for } t \in \mathbb{R}\,, \tag{2.0.1}$$

where $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$, called *kernel function*, is measurable and Λ is an *infinitely divisible independently scattered random measure* (i. d. i. s. r. m.). We recall the definition of an i. d. i. s. r. m. on $\mathbb{R}_+ \times \mathbb{R}$: let $\mathcal{A}$ be a δ-ring (i. e. a ring which is closed under countable intersections) of $\mathbb{R}_+ \times \mathbb{R}$ such that there exists an increasing sequence $\{S_n\}_{n\in\mathbb{N}}$ of sets in $\mathcal{A}$ with $\bigcup_{n=1}^{\infty} S_n = \mathbb{R}_+ \times \mathbb{R}$. Moreover, let $\Lambda = \{\Lambda(A) \,:\, A \in \mathcal{A}\}$ be a real valued stochastic process defined on some probability space. We call Λ an *independently scattered random measure*, if for every sequence $\{A_n\}_{n\in\mathbb{N}}$ of disjoint sets in $\mathcal{A}$, the random variables (r. v. s) $\Lambda(A_n)$, $n \in \mathbb{N}$, are independent, and, if $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$, then we also have

$$\Lambda\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \Lambda(A_n) \quad \text{a. s.},$$

where the sum on the right hand side is assumed to converge almost surely (a. s.). In addition, if $\Lambda(A)$ is a symmetric random variable (r. v.) for every $A \in \mathcal{A}$, then we call Λ a *symmetric random measure*. We call a random measure *infinitely divisible* (i. d.), if $\Lambda(A)$ is i. d. for every $A \in \mathcal{A}$. The reader is referred to Rajput and Rosinski [121], Urbanik [140, 141] and Kwapieǹ and Woyczyzǹski [92] for more details on i. d. i. s. r. m. and integrals as given in (2.0.1).

In the following we consider only i. d. i. s. r. m., where the characteristic function of $\Lambda(A)$ has the representation $\mathbb{E}[\exp(iu\Lambda(A))] = \exp(\psi_A(u))$ for $u \in \mathbb{R}$, $A \in \mathcal{A}$ with

$$\psi_A(u) = iuM(A) - \frac{1}{2}u^2\Sigma^2(A) + \int_{\mathbb{R}} \left(\mathrm{e}^{iux} - 1 - iu\kappa(x)\right) Q_A(dx)\,,$$

$\kappa(x) = \mathbf{1}_{[-1,1]}(x)$ and $\mathbf{1}_A$ denotes the indicator function of the set A. The quantities (M, Σ^2, Q) are called *generating triplet*, where $M : \mathcal{A} \to \mathbb{R}$ is a signed measure, $\Sigma^2 : \mathcal{A} \to [0, \infty)$ is a positive measure and Q_A is a Lévy measure on $\mathbb{R}$ for every $A \in \mathcal{A}$.

Here we consider i. d. i. s. r. m. with cumulant generating function of the form $\psi_A(u) = \lambda(A)\psi(u)$ where ψ is the cumulant generating function of a Lévy process

$$\psi(u) = ium - \frac{1}{2}u^2\sigma^2 + \int_{\mathbb{R}} \left(\mathrm{e}^{iux} - 1 - iu\kappa(x)\right) \nu(dx) \quad \text{for } u \in \mathbb{R} \tag{2.0.2}$$

with a Lévy measure ν on $\mathbb{R}$, $\lambda(d\omega) = \pi(dr) \times dt$ for $\omega = (r, t) \in \mathbb{R}_+ \times \mathbb{R}$ and a probability measure π on $\mathbb{R}_+$. We denote the *generating quadruple* by (m, σ^2, ν, π) and by L the Lévy process corresponding to the generating triplet (m, σ^2, ν). Typical examples for mixed MA processes are *superpositions of Ornstein-Uhlenbeck* (supOU) processes (Example 2.2.10) or *Lévy driven MA processes*, which are often used for stochastic volatility models (Chapter 3). If $f(r, s)$ is independent of r, i. e. $f(r, s) = \widetilde{f}(s)$ for every $r \in \mathbb{R}_+, s \in \mathbb{R}$ and $\widetilde{f} : \mathbb{R} \to \mathbb{R}$ measurable, then Y given by (2.0.1) is the classical Lévy driven MA process

$$Y(t) = \int_{\mathbb{R}} f(t - s)\, dL(s) \quad \text{for } t \in \mathbb{R}, \tag{2.0.3}$$

where we used the same symbol f for the kernel function $\widetilde{f}$.

We shall decompose Λ into two independent i. d. i. s. r. m. according to the jump sizes of the underlying Lévy process L, which are represented by ν:

$$\Lambda = \Lambda_1 + \Lambda_2 \quad \text{and} \quad \Lambda_1(A) = \int_A \int_{\mathbb{R}} x\, d\widetilde{N}_1(\omega, x) \quad \text{for } A \in \mathcal{A} \tag{2.0.4}$$

where $\widetilde{N}_1$ is a *Poisson random measure* with intensity

$$\vartheta(dr \times dt \times dx) = \pi(dr) \times dt \times \nu_1(dx),$$

denoted by PRM(ϑ), and ν_1 is the Lévy measure

$$\nu_1(A) = \nu(A \cap (1,\infty)) + \nu(A \cap (-\infty,-1)) \quad \text{for } A \in \mathcal{B}(\mathbb{R}).$$

The generating quadruple of Λ_1 is $(0, 0, \nu_1, \pi)$. Furthermore, Λ_1 is called *compound Poisson random measure*. The i. d. i. s. r. m. Λ_2 has the generating quadruple $(m, \sigma^2, \nu_2, \pi)$ with Lévy measure

$$\nu_2(A) = \nu(A \cap [-1,1]) \quad \text{for } A \in \mathcal{B}(\mathbb{R}),$$

i. e. it has finite support. We refer to Pedersen [117] for the Lévy-Ito decomposition of i. d. i. s. r. m. s.

Hence, the Lévy process L_1 with generating triplet $(0, 0, \nu_1)$ has jumps with absolute value larger than one, and the Lévy process L_2 with generating triplet (m, σ^2, ν_2) has jumps with modulus strictly smaller than one. Furthermore $\widetilde{N}_1$ has the representation

$$\widetilde{N}_1 = \sum_{k=-\infty}^{\infty} \varepsilon_{(R_k, \Gamma_k, Z_k)}, \tag{2.0.5}$$

where $-\infty < \ldots < \Gamma_{-1} < \Gamma_0 \leq 0 < \Gamma_1 < \ldots < \infty$ are the jump times of a Poisson process $N = \{N(t)\}_{t\in\mathbb{R}}$ with intensity $\mu = \nu_1(\mathbb{R}) > 0$, $Z = \{Z_k\}_{k\in\mathbb{Z}}$ is an i. i. d. sequence with distribution function (d. f.) $\mathbb{P}(Z_1 \leq x) = \nu_1(-\infty, x]/\mu$ for $x \in \mathbb{R}$ and $R = \{R_k\}_{k\in\mathbb{Z}}$ is an i. i. d. sequence with d. f. π. The processes N, Z and R are independent. It is also possible to choose a different decomposition in (2.0.4) by a Poisson random measure and an i. d. i. s. r. m., whose underlying Lévy process has bounded support in an environment of the origin.

This decomposition of Λ induces a decomposition of Y as given in (2.0.1) by $Y = Y_1 + Y_2$, where, for $i = 1, 2$,

$$Y_i(t) = \int_{\mathbb{R}_+\times\mathbb{R}} f(r, t-s)\, d\Lambda_i(r, s)\,, \quad t \in \mathbb{R} \tag{2.0.6}$$

are independent mixed MA processes. The extremal behavior of a mixed MA process Y driven by a subexponential Lévy process is completely determined by extremes of the *mixed Poisson shot noise process*

$$Y_1(t) = \sum_{k=-\infty}^{\infty} f(R_k, t-\Gamma_k) Z_k, \tag{2.0.7}$$

which plays a crucial role in this chapter. A mixed MA process driven by a compound Poisson random measure is called a mixed Poisson shot noise process. Subexponential models are typical models for situations, where extremely large values are likely to occur in comparison to the mean size of the data. The large jumps affect the mixed MA process as large jumps of the Lévy process are carried on in time by the kernel function causing long high level excursions of the process Y. The mixed MA process Y_2 has no influence on the extremal behavior. Extremes of subexponential Lévy driven MA processes, which are in the maximum domain of attraction of the *Gumbel* distribution have been studied in Chapter 1. In this chapter we investigate subexponential Lévy driven mixed MA processes in the maximum domain of attraction of the *Fréchet* distribution, i. e. with *regularly varying tails* (Definition 2.1.1). They include, in particular, stable, Pareto, lognormal and Burr distribution. We present the precise conditions for the Lévy process L below. First, we give some notations: $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$, $\stackrel{w}{\Longrightarrow}$ denotes weak convergence and $\stackrel{v}{\Longrightarrow}$ denotes vague convergence.

Condition (L1).
The marginal distribution $L(1)$ of the Lévy process L is regularly varying of index $-\alpha$ for some $\alpha > 0$, i.e there exists a sequence $0 < a_n \uparrow \infty$ of constants such that

$$n\mathbb{P}(a_n^{-1}L(1) \in \cdot) \stackrel{v}{\Longrightarrow} \sigma(\cdot) \quad \text{on } \mathcal{B}(\overline{\mathbb{R}}\backslash\{0\}) \text{ for } n \to \infty, \tag{2.0.8}$$

where

$$\sigma(dx) = p\alpha x^{-\alpha-1}\, \mathbf{1}_{(0,\infty)}(x)\, dx + q\alpha(-x)^{-\alpha-1}\, \mathbf{1}_{(-\infty,0)}(x)\, dx \tag{2.0.9}$$

for some $p \in [0,1]$ and $q = 1 - p$.

Applying Karamata's Theorem (Bingham et al. [29], Theorem 1.11), we get

$$\int_{|x|>1} |x|^\delta\, \nu(dx) < \infty \text{ for } \delta < \alpha \quad \text{and} \quad \int_{|x|>1} |x|^\delta\, \nu(dx) = \infty \text{ for } \delta > \alpha. \tag{2.0.10}$$

Thus, by Sato [138], Corollary 25.8, we obtain

$$\mathbb{E}|L(1)|^\delta \begin{cases} < \infty & \text{for } \delta < \alpha \\ = \infty & \text{for } \delta > \alpha \end{cases}. \tag{2.0.11}$$

Extreme value theory for stable MA processes was derived in Rootzén [130]. To our knowledge Rootzén's work on stable MA processes in [130] is the only investigation on the extremal behavior of regularly varying MA processes in continuous-time. We extend Rootzén's results for the larger class of regularly varying mixed MA processes. Furthermore, we weaken his assumptions on the kernel function. This includes also heavy tailed, long memory processes like FICARMA processes, Examples 2.2.9, and supOU processes, Example 2.2.10.

Throughout the chapter we shall assume the following condition on the mixed MA process. Therefore we define

$$\mathbb{L}^\delta(\pi) := \left\{ f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R} \text{ measurable}, \int_{\mathbb{R}_+} \int_{\mathbb{R}} |f(r,s)|^\delta \, ds \, \pi(dr) < \infty \right\}$$

for $\delta > 0$. If $f(r,s)$ is independent of r we write $f \in \mathbb{L}^\delta$ instead of $f \in \mathbb{L}^\delta(\pi)$.

Condition (M1).
Let Y be a mixed MA process as given in (2.0.1). We assume the Lévy process L satisfies (L1). Furthermore, the kernel function $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ is in $\mathbb{L}^\delta(\pi)$ for some $\delta < \alpha$ or $L(1)$ is α-stable and $f : \mathbb{R} \to \mathbb{R}$ belongs to $\mathbb{L}^\alpha(\pi)$. Suppose $f^+ = \sup_{(r,t)\in\mathbb{R}_+\times\mathbb{R}} f^+(r,t) < \infty$ and $f^- = \sup_{(r,t)\in\mathbb{R}_+\times\mathbb{R}} f^-(r,t) \leq f^+$ with $f^+(r,t) = \max\{f(r,t), 0\}$, $f^-(r,t) = \max\{-f(r,t), 0\}$. Furthermore, we assume Y, Y_1 and Y_2 are stationary i. d. processes.

We will give sufficient conditions for Y to be a stationary i. d. process and also regularly varying of index $-\alpha$ (Proposition 2.2.3, Proposition 2.2.7). For later reference we formulate an additional assumption.

Condition (M2).
Let Y be a mixed MA process given by (2.0.1) satisfying (M1) with a. s. sample paths in $\mathbb{D}(\mathbb{R})$. We assume that there exists an $\eta^{(1)} \in \mathbb{R}$ with $f(r, \eta^{(1)}) = f^+ \geq f^-$ for every $r \in \operatorname{supp}(\pi)$. Define

$$g(r,t) := \sup_{s\in(-\infty,t]} |f(r,s)| \, \mathbf{1}_{\left(-\infty,\eta^{(1)}\right]}(t) + \sup_{s\in(t,\infty)} |f(r,s)| \, \mathbf{1}_{\left(\eta^{(1)},\infty\right)}(t), \qquad (2.0.12)$$

then we suppose $g \in \mathbb{L}^\delta(\pi)$ and $\left\{\sum_{k=-\infty}^{\infty} g(R_k, t - \Gamma_k)|Z_k|\right\}_{t\in\mathbb{R}}$ is a stationary i. d. process.

Our investigation on the extremal behavior of Y is based on a partition of the real line into intervals I_k, each containing exactly one time point $\{\Gamma_k + \eta^{(1)}\}_{k\in\mathbb{N}}$. We consider the *marked point process* formed by the coordinates

$$(\sup_{h\in I_k} Y(h), Y(\Gamma_k + \eta^{(1)}), Y(\Gamma_k + \eta^{(2)}), Y(\Gamma_k + t_1), \ldots, Y(\Gamma_k + t_d))$$

for $k \in \mathbb{Z}$, $f(\eta^{(2)}) = -f^-$ and any $t_1, \ldots, t_d \in \mathbb{R}$. We investigate its limit behavior, which is a marked cluster Poisson random measure. We will see that exceedances of Y are caused by extreme jumps of the Lévy process, such that extremes of Y above high levels occur at large jump times of the Lévy process in combination with extremes of the kernel function. In the neighborhood of such an extreme the behavior of the process in relation to this extreme is like the kernel function. Moreover, exceedances over high thresholds are carried on in time by the kernel function. In contrast to Lévy processes in $\mathcal{S} \cap \mathrm{MDA}(\Lambda)$ (Chapter 1) cluster sizes are randomly distributed. The results are applied in particular to supOU processes. As an important subclass of mixed MA processes we obtain not only the extremal behavior of heavy tailed Poisson shot noise processes, but also of *stationary renewal shot noise processes*.

This chapter is organized as follows: we start with a short introduction of multivariate regular variation in Section 2.1 followed in Section 2.2 by an investigation of heavy tailed mixed MA processes. This includes, on the one hand sufficient conditions for (M1) in Section 2.2.1, on the other hand the tail behavior of a mixed MA process Y satisfying (M1) as well as the tail behavior of $M(h) = \sup_{t \in [0,h]} Y(t)$ for $h > 0$ in Section 2.2.2. Finally, Section 2.2.3 gives, with FICARMA and supOU processes, examples for heavy tailed mixed MA processes which exhibit long range dependence. In Section 2.3 we obtain analogous results as for regularly varying mixed MA processes for regularly varying renewal shot noise processes.

Section 2.4 is concerned with the point process behavior of multivariate regularly varying stationary sequences. Moreover, Davis and Mikosch [53] generalize results of Davis and Hsing [52] on the point process behavior of stationary processes with regularly varying tails to a multivariate setting. We give an overview of their results, which are then applied in Section 2.5. First, we study the asymptotic behavior of the embedded marked point process of local maxima of Y in Section 2.5.1. Afterwards, in Section 2.5.2, we present under less restrictive assumptions than in Section 2.5.1 a marked point process result, which also includes the behavior of large infima of Y. Moreover, we obtain the limit distribution of running maxima of the mixed MA process Y in Section 2.5.3, and compute the extremal index function in Section 2.5.4. The mixed MA process exhibits clusters in extremes. The results are in particular valid for stationary renewal shot noise processes.

Throughout the chapter we use the following notation:
Let $\eta^{(1)}, \eta^{(2)} \in \mathbb{R}$ with $f(r, \eta^{(1)}) = f^+$, $f(r, \eta^{(2)}) = -f^-$ for every $r \in \mathrm{supp}(\pi)$,

$t_1, \ldots, t_d \in \mathbb{R}$ for $d \in \mathbb{N}_0$, then

$$\mathbf{f}(r,t) = (f(r,t+t_1), \ldots, f(r,t+t_d), f(r,t+\eta^{(1)}), f(r,t+\eta^{(2)})) \qquad (2.0.13)$$

and

$$\mathbf{Y}(t) = (Y(t+t_1), \ldots, Y(t+t_d), Y(t+\eta^{(1)}), Y(t+\eta^{(2)})) \text{ for } t \in \mathbb{R}. \qquad (2.0.14)$$

In the case that the assumption $f(r,\eta^{(2)}) = -f^-$ is not needed, we always assume $\eta^{(2)} \in \mathbb{R}$.

We write $\overline{F} = 1 - F$ for the right tail of the d. f. F, F^{2*} for the convolution $F * F$ and $\overline{F^{2*}} = 1 - F^{2*}$. For any r. v. Z on $\mathbb{R}$ we write $Z^+ = Z \vee 0$ and $Z^- = -Z \vee 0$; $X \stackrel{d}{=} Y$, if the distributions of the r. v. s X and Y coincide. For real functions g and h we abbreviate $g(t) \sim h(t)$ for $t \to \infty$, if $g(t)/h(t) \stackrel{t\to\infty}{\longrightarrow} 1$, $g^+(t) = \max\{0, g(t)\}$, $g^-(t) = \max\{0, -g(t)\}$, $g^+ = \sup_{t\in\mathbb{R}} g^+(t)$ and $g^- = \sup_{t\in\mathbb{R}} g^-(t)$. For a vector $\mathbf{x} \in \mathbb{R}^r$ we denote by $\mathbf{x}^t$ the transposed of $\mathbf{x}$ and by $|\mathbf{x}| = \max\{|x_1|, \ldots, |x_r|\}$ the maximum norm. For a matrix $\mathbf{A} \in \mathbb{R}^{d\times r}$ we denote by $\|A\|$ the row-sum-norm. Further, $\sum_{k=1}^{0} := 0$ and $\bigvee_{k=1}^{0} := 0$.

2.1 Multivariate regular variation

By considering the multidimensional stationary process $\{\mathbf{Y}(t)\}_{t\in\mathbb{R}}$, whose marginal is a multivariate distribution, we need the definition of regular variation for a multivariate distribution:

Definition 2.1.1 (Multivariate regular variation)
A random vector $\mathbf{X} = (X_1, \ldots, X_d)$ *on* $\mathbb{R}^d$ *is said to be regularly varying with index* $-\alpha$, $\alpha > 0$, *if there exists a random vector* Θ *with values on the unit sphere* $\mathbb{S}^{d-1} = \{\mathbf{x} \in \mathbb{R}^d : |\mathbf{x}| = 1\}$ *such that for every* $x > 0$

$$\frac{\mathbb{P}(|\mathbf{X}| > ux, \mathbf{X}/|\mathbf{X}| \in \cdot)}{\mathbb{P}(|\mathbf{X}| > u)} \stackrel{w}{\Longrightarrow} x^{-\alpha}\mathbb{P}(\Theta \in \cdot) \quad \text{on } \mathcal{B}(\mathbb{S}^{d-1}) \text{ for } u \to \infty. \qquad (2.1.1)$$

The distribution of Θ is referred to as the *spectral measure* of $\mathbf{X}$. It describes in which direction we are likely to find extreme realizations of $\mathbf{X}$. This definition of regular variation is equivalent to the following (Lindskøg [99], Theorem 1.15):

There exists a Radon measure $\sigma(\cdot)$ on $\overline{\mathbb{R}}^d\backslash\{\mathbf{0}\}$, which is defined to be finite on compact sets, with $\sigma(\overline{\mathbb{R}}^d\backslash\mathbb{R}^d) = 0$ and $\sigma(E) > 0$ for at least one relative compact set

$E \subseteq \overline{\mathbb{R}}^d \backslash \{\mathbf{0}\}$, where $\mathbf{0} = (0, \ldots, 0) \in \mathbb{R}^d$, and a sequence $0 < a_n \uparrow \infty$ of constants such that

$$n\mathbb{P}(a_n^{-1}\mathbf{X} \in \cdot) \stackrel{v}{\Longrightarrow} \sigma(\cdot) \quad \text{on } \mathcal{B}(\overline{\mathbb{R}}^d \backslash \{\mathbf{0}\}) \text{ for } n \to \infty. \tag{2.1.2}$$

More about multivariate regular variation can be found in Resnick [125], Chapter 5, Basrak et al. [19], Lindskøg [99], Mikosch [110] and in the references of Bingham et al. [29].

The following Lemma is a multivariate extension of Breiman's [34] classical result on regular variation of products.

Lemma 2.1.2 (Basrak et al. [20], Proposition A.1)
Let $\mathbf{X}$ be a regularly varying random vector of index $-\alpha$ on $\mathbb{R}^r$, $r \in \mathbb{N}$, in the sense of (2.1.2) and $\mathbf{A}$ be a random $d \times r$-matrix, independent of $\mathbf{X}$. If $0 < \mathbb{E}\|\mathbf{A}\|^\gamma < \infty$ for some $\gamma > \alpha$, then $\mathbf{AX}$ is regularly varying of index $-\alpha$ and

$$n\mathbb{P}(a_n^{-1}\mathbf{AX} \in \cdot) \stackrel{v}{\Longrightarrow} \mathbb{E}\left(\sigma \circ \mathbf{A}^{-1}(\cdot)\right) \quad \text{on } \mathcal{B}(\overline{\mathbb{R}}^d \backslash \{\mathbf{0}\}) \text{ for } n \to \infty, \tag{2.1.3}$$

where $\mathbf{A}^{-1}$ is the inverse image of $\mathbf{A}$.

We need the following special case of Lemma 2.1.2.

Lemma 2.1.3
Let $\mathbf{Z} = (Z_1, \ldots, Z_r)$ be a vector of independent r. v. s, which are regularly varying in the sense of (2.1.2) such that for $j = 1, \ldots, r$,

$$n\mathbb{P}(a_n^{-1}Z_j \in \cdot) \stackrel{v}{\Longrightarrow} \sigma_j(\cdot) \quad \text{on } \mathcal{B}(\overline{\mathbb{R}} \backslash \{0\}) \text{ for } n \to \infty,$$

where

$$\sigma_j(dx) = p_j \alpha x^{-\alpha-1} \mathbf{1}_{(0,\infty)}(x)\, dx + q_j \alpha (-x)^{-\alpha-1} \mathbf{1}_{(-\infty,0)}(x)\, dx \tag{2.1.4}$$

with $p_j, q_j \geq 0$, $p_j + q_j > 0$. Furthermore, let $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_r)$ be a random $d \times r$-matrix, independent of $\mathbf{Z}$. If $0 < \mathbb{E}\|\mathbf{A}\|^\gamma < \infty$ for some $\gamma > \alpha$, then $\mathbf{Y} = \mathbf{AZ}$ is regularly varying of index $-\alpha$ and has spectral measure with respect to a_n

$$\mathbb{P}(\Theta \in \cdot) = \frac{\sum_{j=1}^r \left[p_j \mathbb{E}(|\mathbf{a}_j|^\alpha \mathbf{1}_{\{\mathbf{a}_j/|\mathbf{a}_j| \in \cdot\}}) + q_j \mathbb{E}(|\mathbf{a}_j|^\alpha \mathbf{1}_{\{-\mathbf{a}_j/|\mathbf{a}_j| \in \cdot\}})\right]}{\sum_{j=1}^r (p_j + q_j)\mathbb{E}|\mathbf{a}_j|^\alpha}. \tag{2.1.5}$$

For $x > 0$

$$\lim_{n\to\infty} n\mathbb{P}(|\mathbf{Y}| > a_n x) = x^{-\alpha} \sum_{j=1}^r (p_j + q_j)\mathbb{E}|\mathbf{a}_j|^\alpha, \tag{2.1.6}$$

and in the case $d = 1$,

$$\lim_{n\to\infty} n\mathbb{P}(Y > a_n x) = x^{-\alpha} \sum_{j=1}^{r} [p_j \mathbb{E} a_j^{+\,\alpha} + q_j \mathbb{E} a_j^{-\,\alpha}]. \tag{2.1.7}$$

Proof.
Write $\mathbf{e}_j = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^r$, where in the j^{th} coordinate is one, and $E_j = \{c\mathbf{e}_j : c \in \mathbb{R}\}$. The map $\mathrm{pr}_j : \mathbb{R}^r \to \mathbb{R}$ is the projection on the j^{th} coordinate, $\mathbf{x} = (x_1, \ldots, x_r) \mapsto x_j$, $j = 1, \ldots, r$. We apply Lemma 2.1.2 and obtain that $\mathbf{Y}$ is multivariate regularly varying of index $-\alpha$ with

$$n\mathbb{P}(a_n^{-1}\mathbf{Y} \in \cdot) \overset{v}{\Longrightarrow} \mathbb{E}[\sigma \circ \mathbf{A}^{-1}(\cdot)] \quad \text{on } \mathcal{B}(\overline{\mathbb{R}}^d \backslash \{\mathbf{0}\}) \text{ for } n \to \infty, \tag{2.1.8}$$

where

$$\sigma(\cdot) = \sum_{j=1}^{r} \sigma_j \circ \mathrm{pr}_j(E_j \cap \cdot). \tag{2.1.9}$$

For a set $\widetilde{S} \in \mathcal{B}(\mathbb{S}^{d-1})$ we define the set

$$B_x(\widetilde{S}) := \left\{ \mathbf{x} \in \mathbb{R}^d : |\mathbf{x}| > x, \mathbf{x}/|\mathbf{x}| \in \widetilde{S} \right\} \tag{2.1.10}$$

for $x > 0$. Note, that $\mathbf{a}_j = \mathbf{A}\mathbf{e}_j$ for $j = 1, \ldots, r$. Then

$$\begin{aligned} &E_j \cap \mathbf{A}^{-1} B_x(\widetilde{S}) \\ &\quad = \{c\mathbf{e}_j \in \mathbb{R}^r : c|\mathbf{a}_j| > x\}\, \mathbf{1}_{\{\mathbf{a}_j/|\mathbf{a}_j| \in \widetilde{S}\}} + \{c\mathbf{e}_j \in \mathbb{R}^r : -c|\mathbf{a}_j| > x\}\, \mathbf{1}_{\{-\mathbf{a}_j/|\mathbf{a}_j| \in \widetilde{S}\}} \end{aligned}$$

and, in particular, by (2.1.4) we obtain

$$\begin{aligned} &\mathbb{E}[\sigma_j \circ \mathrm{pr}_j(E_j \cap \mathbf{A}^{-1} B_x(\widetilde{S}))] \\ &\quad = \mathbb{E}[\sigma_j(x/|\mathbf{a}_j|, \infty)\, \mathbf{1}_{\{\mathbf{a}_j/|\mathbf{a}_j| \in \widetilde{S}\}}] + \mathbb{E}[\sigma_j(-\infty, -x/|\mathbf{a}_j|)\, \mathbf{1}_{\{-\mathbf{a}_j/|\mathbf{a}_j| \in \widetilde{S}\}}] \\ &\quad = x^{-\alpha}(p_j \mathbb{E}[|\mathbf{a}_j|^\alpha\, \mathbf{1}_{\{\mathbf{a}_j/|\mathbf{a}_j| \in \widetilde{S}\}}] + q_j \mathbb{E}[|\mathbf{a}_j|^\alpha\, \mathbf{1}_{\{-\mathbf{a}_j/|\mathbf{a}_j| \in \widetilde{S}\}}]). \end{aligned} \tag{2.1.11}$$

Taking (2.1.8)-(2.1.11) into account, we obtain

$$\begin{aligned} \mathbb{P}(\Theta \in \cdot) &= \frac{\mathbb{E}[\sigma \circ \mathbf{A}^{-1}(B_1(\cdot))]}{\mathbb{E}[\sigma \circ \mathbf{A}^{-1}(B_1(\mathbb{S}^{d-1}))]} \qquad (2.1.12) \\ &= \sum_{j=1}^{r} \left[p_j \mathbb{E}(|\mathbf{a}_j|^\alpha\, \mathbf{1}_{\{\mathbf{a}_j/|\mathbf{a}_j| \in \cdot\}}) + q_j \mathbb{E}(|\mathbf{a}_j|^\alpha\, \mathbf{1}_{\{-\mathbf{a}_j/|\mathbf{a}_j| \in \cdot\}}) \right] \Big/ \sum_{j=1}^{r} (p_j + q_j) \mathbb{E}|\mathbf{a}_j|^\alpha \end{aligned}$$

and

$$\lim_{n\to\infty} n\mathbb{P}(|\mathbf{Y}| > a_n x) = \mathbb{E}[\sigma \circ \mathbf{A}^{-1}(B_x(\mathbb{S}^{d-1}))] = x^{-\alpha}\sum_{j=1}^{r}(p_j + q_j)\mathbb{E}|\mathbf{a}_j|^{\alpha}. \quad (2.1.13)$$

In the case $d = 1$, $a_j/|a_j| \in \{-1, 1\}$ so that by (2.1.12) we have

$$\lim_{n\to\infty} n\mathbb{P}(Y > a_n x) = x^{-\alpha}\mathbb{P}(\Theta = 1) = x^{-\alpha}\sum_{j=1}^{r}[p_j\mathbb{E}a_j^{+\,\alpha} + q_j\mathbb{E}a_j^{-\,\alpha}]. \qquad \square$$

Remark 2.1.4
Let $\mathbf{A}$ be a deterministic matrix and $\rho = \sum_{j=1}^{r}(p_j + q_j)\mathbb{E}|\mathbf{a}_j|^{\alpha}$. An interpretation of Lemma 2.1.3 is that the spectral measure Θ reaches the value $\mathbf{a}_j/|\mathbf{a}_j|$ with probability $p_j|\mathbf{a}_j|^{\alpha}/\rho$ and $-\mathbf{a}_j/|\mathbf{a}_j|$ with probability $q_j|\mathbf{a}_j|^{\alpha}/\rho$. Thus only in the directions $\mathbf{a}_j/|\mathbf{a}_j|$, $-\mathbf{a}_j/|\mathbf{a}_j|$, $j = 1, \ldots, r$, extremes are likely to occur. $\square$

2.2 Regularly varying mixed MA processes

This chapter is concerned with extremes of regularly varying mixed MA processes Y as given in (2.0.1). This means that the underlying Lévy process of the driving i. d. i. s. r. m. Λ as given in (2.0.2) is regularly varying. The question arises on the existence of such mixed MA processes. On the one hand we will give in Section 2.2.1 sufficient conditions on the existence of stationary mixed MA processes and on the other hand, in Section 2.2.2, we show that also the stationary distribution is regularly varying. Further we compute the tail behavior of the stationary distribution explicitly.

2.2.1 Existence of heavy tailed mixed MA processes

Let Y be a mixed MA process as given in (2.0.1). Under certain conditions $Y(t)$ is well-defined as a limit in probability of integrals of step functions approximating f and moreover, Y is stationary. This has been shown by Rajput and Rosinski [121] (see also Kwapień and Woyczyǹski [92]), who also give conditions for $Y(0)$ to be i. d.. The conditions are formulated in terms of the kernel function f and the generating quadruple (m, σ^2, ν, π) of the i. d. i. s. r. m. Λ.

Proposition 2.2.1 (Rajput and Rosinski [121], Theorem 2.7)
Let Λ be an i. d. i. s. r. m. with generating quadruple (m, σ^2, ν, π). Then $Y(0)$ given in (2.0.1) is well-defined and i. d. if and only if

$$\begin{aligned}
&\int_{\mathbb{R}_+} \left[\int_{\mathbb{R}} \left| m f(r,s) + \int_{\mathbb{R}} \left(\kappa(x f(r,s)) - f(r,s)\kappa(x) \right) \nu(dx) \right| ds \right] \pi(dr) < \infty, \\
&\int_{\mathbb{R}_+} \left[\int_{\mathbb{R}} |\sigma^2 f(r,s)|^2 \, ds \right] \pi(dr) < \infty, \qquad (2.2.1)\\
&\int_{\mathbb{R}_+} \left[\int_{\mathbb{R}} \left[\int_{\mathbb{R}} \min\{1, |f(r,s)x|^2\} \, \nu(dx) \right] ds \right] \pi(dr) < \infty,
\end{aligned}$$

where $\kappa(x) = x\, \mathbf{1}_{[-1,1]}(x)$. The generating triplet of Y is (m_Y, σ_Y^2, ν_Y), where

$$\begin{aligned}
m_Y &= \int_{\mathbb{R}_+} \left[\int_{\mathbb{R}} m f(r,s) + \int_{-\infty}^{\infty} \left(\kappa(x f(r,s)) - f(r,s)\kappa(x) \right) \nu(dx) \right] ds \, \pi(dr), \\
\sigma_Y^2 &= \sigma^2 \int_{\mathbb{R}_+} \left[\int_{\mathbb{R}} f^2(r,s)\, ds \right] \pi(dr), \qquad (2.2.2)\\
\nu_Y\, [x, \infty) &= \int_{f(r,s)>0} \nu \left[\frac{x}{f(r,s)}, \infty \right) ds \, \pi(dr) + \int_{f(r,s)<0} \nu \left(-\infty, \frac{x}{f(r,s)} \right] ds \, \pi(dr)
\end{aligned}$$

for $x > 0$.

Typical examples for regularly varying mixed MA processes are mixed MA processes driven by a *stable Lévy process*. They are very well known and have been thoroughly investigated by Samorodnitsky and Taqqu [137].

Example 2.2.2
Let $L(1)$ be α-stable with $\alpha \in (0,2)$, $\alpha \neq 1$, $c \geq 0$, $\beta \in [-1,1]$, $\tau \in \mathbb{R}$, (we write $L(1) \sim S_\alpha(c^{1/\alpha}, \beta, \tau)$). The cumulant generating function (2.0.2) is

$$\psi(u) = \exp\{-c|u|^\alpha [1 - i\beta \tan(\pi\alpha/2)\, \text{sign}(u)] + iu\tau\}.$$

We define $C_\alpha := [\cos(\pi\alpha/2)\Gamma(1-\alpha)]^{-1}$ and take

$$\begin{aligned}
\textstyle\int_0^\infty x^{-\alpha} \sin(ux)\, dx &= \cos(\pi\alpha/2)\, \Gamma(1-u) u^{\alpha-1} && \text{for } \alpha \in (0,2), \\
\textstyle\int_0^\infty x^{-\alpha+1} \sin(ux)\, dx &= u^{\alpha-2} \tan(\pi\alpha/2)\, (1-\alpha) C_\alpha && \text{for } \alpha \in (1,2)
\end{aligned}$$

into account to obtain the generating triplet (m, σ^2, ν) of L as

$$m = \alpha \frac{c_1 - c_2}{1-\alpha} + \tau, \quad \sigma^2 = 0 \text{ and } \nu(dx) = \alpha \left(c_1\, \mathbf{1}_{(-\infty,0)}(x) + c_2\, \mathbf{1}_{(0,\infty)}(x) \right) |x|^{-(\alpha+1)}\, dx,$$

with $c_1 = cC_\alpha(1-\beta)/2$ and $c_2 = cC_\alpha(1+\beta)/2$. Then

$$\begin{aligned}
&\int_{\mathbb{R}_+}\left[\int_{\mathbb{R}}\left|mf(r,s)+\int_{\mathbb{R}}(\kappa(xf(r,s))-f(r,s)\kappa(x))\,\nu(dx)\right|\,ds\right]\pi(dr)\\
&\qquad = \alpha\frac{|c_1-c_2|}{1-\alpha}\int_{\mathbb{R}_+}\int_{\mathbb{R}}|f(r,s)|^\alpha\,ds\pi(dr),\\
&\int_{\mathbb{R}_+}\left[\int_{\mathbb{R}}\left[\int_{\mathbb{R}}\min\{1,|f(r,s)x|^2\}\,\nu(dx)\right]\,ds\right]\,\pi(dr)\\
&\qquad = \alpha(c_1+c_2)\left(\frac{1}{2-\alpha}+\frac{1}{\alpha}\right)\int_{\mathbb{R}_+}\left[\int_{\mathbb{R}}|f(r,s)|^\alpha\,ds\right]\pi(dr).
\end{aligned}$$

This means, by Proposition 2.2.1, that the stable mixed MA process Y is stationary and the marginal distribution is i. d. if and only if $f \in \mathbb{L}^\alpha(\pi)$. □

In the following we give sufficient conditions for (2.2.1) in the case of mixed MA processes with a regularly varying Lévy measure. Apart from the stable case and some special regularly varying Lévy measures there seem to be no simple equivalent conditions to (2.2.1).

Proposition 2.2.3

Let Λ be an i. d. i. s. r. m. with generating quadruple $(m, 0, \nu, \pi)$ and ν be a regularly varying function of index $-\alpha$, $\alpha > 0$. Assume $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ to be bounded. Then $Y(0)$ given by (2.0.1) is well-defined, i. d. and Y is stationary, if one of the following conditions is satisfied:

(a) $L(1)$ *is α-stable,* $\alpha \neq 1$, *and* $f \in \mathbb{L}^\alpha(\pi)$.

(b) $f \in \mathbb{L}^\delta(\pi)$ *for some* $\delta < \alpha$, $\delta \leq 1$.

(c) $\mathbb{E}L(1) = 0$, $\alpha > 1$ *and* $f \in \mathbb{L}^\delta(\pi)$ *for some* $\delta < \alpha$, $\delta \leq 2$.

Proof.

For the proof of *(a)* see Example 2.2.2. W. l. o. g. $\sup_{(r,s)\in\mathbb{R}_+\times\mathbb{R}} |f(r,s)| = f^+ \leq 1$. The conditions (2.2.1) of Proposition 2.2.1 are equivalent to

$$I_1 = \int_{\mathbb{R}_+} \left[\int_{\mathbb{R}} \left| mf(r,s) + \int_{|x|>1} f(r,s)x\, \mathbf{1}_{\{|f(r,s)x| \le 1\}}\, \nu(dx) \right| ds \right] \pi(dr) < \infty, \quad (2.2.3)$$

$$I_2 = \int_{\mathbb{R}_+} \left[\int_{\mathbb{R}} \nu\left(\frac{1}{|f(r,s)|}, \infty \right) + \nu\left(-\infty, -\frac{1}{|f(r,s)|} \right) ds \right] \pi(dr) < \infty, \quad (2.2.4)$$

$$I_3 = \int_{\mathbb{R}_+} \left[\int_{\mathbb{R}} \int_{|x|>1} f(r,s)^2 x^2\, \mathbf{1}_{\{|f(r,s)x| \le 1\}}\, \nu(dx)\, ds \right] \pi(dr) < \infty. \quad (2.2.5)$$

First we show that $f \in \mathbb{L}^\delta(\pi)$ for some $\delta < \alpha$, $\delta \le 2$ is a sufficient condition for (2.2.4) and (2.2.5). For the proof of (2.2.3) we need the additional assumptions *(b)-(c)*.

On the one hand, we obtain (2.2.5) by

$$I_3 \le \int_{\mathbb{R}_+} \left[\int_{\mathbb{R}} |f(r,s)|^\delta\, ds \right] \pi(dr) \int_{|x|>1} |x|^\delta\, \nu(dx) < \infty,$$

where (2.0.10) was used.

On the other hand, $\nu(\cdot, \infty) : [1/f^+, \infty) \to (0, \nu(1/f^+, \infty)]$ is non-increasing and bounded away from 0 and ∞ on every compact subset of $[1/f^+, \infty)$. Applying Potter's Theorem A.1.10 *(iv)*, we find a $K > 0$ such that

$$I_2 \le K \int_{\mathbb{R}_+} \int_{\mathbb{R}} \left[\nu\left(\frac{1}{f^+}, \infty \right) \left| \frac{f(r,s)}{f^+} \right|^\delta + \nu\left(-\infty, -\frac{1}{f^+} \right) \left| \frac{f(r,s)}{f^+} \right|^\delta \right] ds\, \pi(dr) \quad (2.2.6)$$

is finite for $f \in \mathbb{L}^\delta(\pi)$. This is (2.2.4).

We prove (2.2.3) under the different assumptions *(b)-(c)*:

(b) The integral in (2.2.3) can be bounded above by

$$I_1 \le \int_{\mathbb{R}_+} \int_{\mathbb{R}} \left[|f(r,s)|^\delta\, ds \right] \pi(dr) \left[m f^{+\,(1-\delta)} + \int_{|x|>1} |x|^\delta\, \nu(dx) \right] < \infty,$$

where we used (2.0.10).

(c) Since $\mathbb{E}L(1) = 0$ by Sato [138], Example 25.11, we have $m = -\int_{|x|>1} x\, \nu(dx)$. Thus (2.2.3) is equivalent to

$$\int_{\mathbb{R}_+} \left[\int_{\mathbb{R}} \left| \int_{|x|>1} f(r,s)x\, \mathbf{1}_{\{|f(r,s)x| > 1\}}\, \nu(dx) \right| ds \right] \pi(dr) < \infty. \quad (2.2.7)$$

W. l. o. g. we assume that $\nu|_{(-\infty,-1)\cup(1,\infty)}$ is a probability measure and Z has d. f. $\nu|_{(-\infty,-1)\cup(1,\infty)}$. Since $\alpha > 1$ we have $\mathbb{E}|Z| < \infty$ by (2.0.11). Furthermore,

$$\begin{aligned}
&\int_{\mathbb{R}_+} \left[\int_{\mathbb{R}} \left| \int_{|x|>1} f(r,s)x\,\mathbf{1}_{\{|f(r,s)x|>1\}}\,\nu(dx) \right| ds \right] \pi(dr) \\
&\quad = \int_{\mathbb{R}_+} \left[\int_{\mathbb{R}} \left| \mathbb{E}(f(r,s)Z\,\mathbf{1}_{\{|f(r,s)Z|>1\}}) \right| ds \right] \pi(dr) \\
&\quad \le \int_{\mathbb{R}_+} \int_{\mathbb{R}} \left[\mathbb{P}(|f(r,s)Z| > 1) + \int_1^\infty \mathbb{P}(|f(r,s)Z| > x)\,dx \right] ds\,\pi(dr).
\end{aligned}$$

Again, by Potter's Theorem A.1.10, analog to (2.2.6), the r. h. s. is bounded by

$$\begin{aligned}
&K \int_{\mathbb{R}_+} \int_{\mathbb{R}} \left[\mathbb{P}(|Z| > 1)|f(r,s)|^\delta + |f(r,s)|^\delta \int_1^\infty \mathbb{P}(|Z| > x)\,dx \right] ds\,\pi(dr) \\
&\qquad = K\mathbb{E}[|Z\,\mathbf{1}_{\{|Z|>1\}}|] \int_{\mathbb{R}_+} \left[\int_{\mathbb{R}} |f(r,s)|^\delta\,ds \right] \pi(dr), \qquad (2.2.8)
\end{aligned}$$

which is finite for $f \in \mathbb{L}^\delta(\pi)$ and $\mathbb{E}|Z| < \infty$.

The proof of the stationarity of Y is analog to the proof in Proposition 1.1.7. □

Remark 2.2.4

For a Lévy driven MA process given by (2.0.3), Proposition 2.2.3 provides sufficient conditions to be stationary and the marginal distribution to be i. d.. Then $\mathbb{L}^\delta(\pi)$ can be replaced by $\mathbb{L}^\delta$. Typical examples for functions in $\mathbb{L}^\delta$ are bounded functions f with $f(t) \sim K_1 t^{-\delta+\epsilon}$, $f(-t) \sim K_2 t^{-\delta+\epsilon}$ for $t \to \infty$ and for some $\epsilon \in (0,\delta)$, K_1, $K_2 \in \mathbb{R}$. □

The conditions in Proposition 2.2.3 are nearly necessary, which is shown in the following Lemma.

Lemma 2.2.5

Let Y be a stationary mixed MA process given by (2.0.1) with kernel function f and generating quadruple $(m, 0, \nu, \pi)$ of the i. d. i. s. r. m. Λ. Then $f \in \mathbb{L}^{\alpha+\epsilon}(\pi)$ for any $\epsilon > 0$.

Proof.
By Proposition 2.2.1 we get I_2 given in (2.2.4) is finite. Using analog techniques as in (2.2.6) we apply Potter's Theorem A.1.10 on $\nu(1/f^+, \infty)/\nu(1/f(r,s), \infty)$ and obtain that there exists, for all $\epsilon > 0$, a $K(\epsilon) > 0$ such that

$$I_2 \geq K(\epsilon) \int_{\mathbb{R}_+} \int_{\mathbb{R}} \left[\nu\left(\frac{1}{f^+}, \infty\right) \left|\frac{f(r,s)}{f^+}\right|^{\alpha+\epsilon} + \nu\left(-\infty, -\frac{1}{f^+}\right) \left|\frac{f(r,s)}{f^+}\right|^{\alpha+\epsilon} \right] ds\, \pi(dr).$$

The result follows by the finiteness of I_2. □

Lemma 2.2.5 says that a MA processes driven by a regularly varying Lévy process with index $-\alpha$, $\alpha \leq 1$, and bounded function f satisfies $f \in \mathbb{L}$.

Let Y given by (2.0.1) be stationary, i. d. and $\mathbb{E}|Y(0)|^2 < \infty$ with $\mathbb{E}|L(1)|^2 < \infty$. Analog to Remark 1.1.8 we obtain

$$\gamma(h) = \mathrm{Cov}(Y(0), Y(h)) = \int_{\mathbb{R}_+ \times \mathbb{R}} f(r,s) f(r, s+h)\, ds\, \pi(dr). \tag{2.2.9}$$

The standard definition of a *long range dependent* (long memory) process Y is $\int_{\mathbb{R}} \gamma(h)\, dh = \infty$. The process Y has short memory, if $\int_{\mathbb{R}} \gamma(h)\, dh < \infty$. Then $\int_{\mathbb{R}} \gamma(h)\, dh = \int_{\mathbb{R}_+} \left[\int_{\mathbb{R}} f(r,s)\, ds\right]^2 \pi(dr)$. For regularly varying distributions of index $-\alpha$, $\alpha < 2$, the definition of long memory is meaningless, since the second moments are infinite. For stable MA processes long memory was introduced in Samorodnitsky [133, 134]. The class of processes we are studying are, in his definition, short memory processes. But for $\alpha \geq 2$, Proposition 2.2.3 gives sufficient conditions for the existence of mixed MA processes having long memory. Note, that we assume $\mathbb{E}L(1) = 0$, which is not offhand negligible. This shows the following Remark.

Remark 2.2.6
Let Y be a stationary mixed MA process given by (2.0.1) and $L(1)$ be regularly varying of index $-\alpha$, $\alpha \geq 1$, and $f \in \mathbb{L}^\delta(\pi)$ positive with $\delta < \alpha$. The additional assumption $\mathbb{E}L(1) = 0$ in Proposition 2.2.3 *(c)* can not be dispensed.

This is manifested by the following example:

We assume that $L(1)$ is a subordinator with cumulant generating function given by (2.0.2) of the form $\psi(u) = \exp(\int_0^\infty (\mathrm{e}^{iux} - 1)\, \nu(dx))$, i. e. $\int_{(-\infty,0)} \nu(dx) = 0$, $\int_{(0,1]} x\, \nu(dx) < \infty$. Then $m = \int_0^1 x\nu(dx)$. We show that $f \in \mathbb{L}(\pi)$ is necessary for Y to be stationary and i. d.. We treat the two cases $m > 0$ and $m = 0$ separately.

If $m = 0$, then $\nu(0,1] = 0$ and L is a compound Poisson process. Taking (2.2.3) into account we get

$$\int_{\mathbb{R}_+}\int_{\mathbb{R}}\left[\int_0^\infty xf(r,s)\,\mathbf{1}_{\{|xf(r,s)|\le 1\}}\;\nu(dx)\right]ds\,\pi(dr)$$
$$=\int_{\mathbb{R}_+}\int_{\mathbb{R}}\left[\int_1^\infty xf(r,s)\,\mathbf{1}_{\{|xf(r,s)|\le 1\}}\;\nu(dx)\right]ds\,\pi(dr)<\infty. \qquad (2.2.10)$$

On the other hand, with (2.2.8),

$$\int_{\mathbb{R}_+}\left[\int_{\mathbb{R}}\left|\int_0^\infty f(r,s)x\,\mathbf{1}_{\{|f(s)x|>1\}}\;\nu(dx)\right|ds\right]\pi(dr)<\infty. \qquad (2.2.11)$$

By (2.2.10)-(2.2.11) and $\int_0^\infty x\,\nu(dx)<\infty$ we obtain $f\in\mathbb{L}(\pi)$.

If $m > 0$, then by (2.2.1),

$$m\left[\int_{\mathbb{R}_+}\int_{-\infty}^{\infty}|f(r,s)|\,ds\right]\pi(dr)$$
$$\le\int_{\mathbb{R}_+}\left[\int_{\mathbb{R}}\left|mf(r,s)+\int_0^\infty(\kappa(xf(r,s))-f(r,s)\kappa(x))\,\nu(dx)\right|ds\right]\pi(dr)<\infty$$

and thus $f\in\mathbb{L}(\pi)$. □

2.2.2 Tail behavior of heavy tailed mixed MA processes

Let Y be a stationary mixed MA process given by (2.0.1). We assume that the underlying Lévy process L of the i.d.i.s.r.m. Λ is regularly varying of index $-\alpha$, $\alpha > 0$ such that for the sequence $0 < a_n \uparrow \infty$ of constants, $p, q \in [0,1]$, $p + q = 1$, the following holds:

$$\lim_{n\to\infty} n\mathbb{P}(L(1) > a_n x) = px^{-\alpha} \quad\text{and}\quad \lim_{n\to\infty} n\mathbb{P}(L(1) < -a_n x) = qx^{-\alpha} \quad\text{for } x > 0.$$

Proposition 2.2.7
Let Y be a stationary mixed MA process given by (2.0.1) satisfying (M1) and $x > 0$. If $pf^+ + qf^- > 0$, then

$$\lim_{n\to\infty} n\mathbb{P}(Y(0) > a_n x) = \left[\int_{\mathbb{R}_+}\int_{\mathbb{R}} p(f^+(r,s))^\alpha + q(f^-(r,s))^\alpha\,ds\,\pi(dr)\right]. \qquad (2.2.12)$$

Similarly, if $pf^- + qf^+ > 0$ we obtain

$$\lim_{n\to\infty} n\mathbb{P}(Y(0) < -a_n x) = \left[\int_{\mathbb{R}_+}\int_{\mathbb{R}} q(f^+(r,s))^\alpha + p(f^-(r,s))^\alpha \, ds\, \pi(dr)\right]. \quad (2.2.13)$$

Then Y is regularly varying of index $-\alpha$ in the sense of (2.1.2). Especially for $t_i \in \mathbb{R}$, $i = 1, \dots, k$,

$$\lim_{n\to\infty} n\mathbb{P}(\max_{i=1,\dots,k} |Y(t_i)| > a_n x) = \int_{\mathbb{R}_+}\int_{\mathbb{R}} \max_{i=1,\dots,k} |f(r, t_i - s)|^\alpha \, ds\, \pi(dr). \quad (2.2.14)$$

Furthermore, let Y has a. s. sample paths in $\mathbb{D}(\mathbb{R})$ and define $M(h) = \sup_{t\in[0,h]} Y(t)$. Then

$$\lim_{n\to\infty} n\mathbb{P}(M(h) > a_n x) \quad (2.2.15)$$

$$\sim \left[\int_{\mathbb{R}_+}\int_{\mathbb{R}} p \sup_{t\in[0,h]} (f^+(r,t+s))^\alpha + q \sup_{t\in[0,h]} (f^-(r,t+s))^\alpha \, ds\, \pi(dr)\right].$$

Proof.
By Proposition 2.2.1 the Lévy measure of Y is

$$\nu_Y(x,\infty) = \int_{f(r,s)>0} \nu\left(\frac{x}{f(r,s)}, \infty\right) ds\, \pi(dr) + \int_{f(r,s)<0} \nu\left(-\infty, \frac{x}{f(r,s)}\right) ds\, \pi(dr)$$

for $x > 0$. Using Potter's Theorem A.1.10 there exists for every $x > 0$, $K > 1$ an $n_0(x) \in \mathbb{N}$ such that $\nu(a_n x y, \infty)/\nu(a_n x, \infty) \le K y^{-\delta}$ for $y \ge 1, n \ge n_0$. Taking $f \in \mathbb{L}^\delta(\pi)$ into account, dominated convergence and the boundedness of f yields for $n \to \infty$

$$\begin{aligned}\frac{\nu_Y(a_n x,\infty)}{\nu(a_n x,\infty)} &= \int_{f^+(r,s)>0} \frac{\nu(a_n x/f^+(r,s),\infty)}{\nu(a_n x,\infty)} ds\, \pi(dr) \\ &\quad + \int_{f^-(r,s)<0} \frac{\nu(-\infty, a_n x/f^-(r,s))}{\nu(a_n x,\infty)} ds\, \pi(dr) \\ &\sim \left[\int_{\mathbb{R}_+}\int_{\mathbb{R}} p(f^+(r,s))^\alpha + q(f^-(r,s))^\alpha \, ds\, \pi(dr)\right].\end{aligned}$$

The result (2.2.12) follows then by Proposition 1.1.2 *(v)*. The proof of (2.2.13) is analog. Statement (2.2.14) follows by Rosinski and Samorodnitsky [132], Theorem 3.1 and similar arguments as above. An application of Theorem 1.3.9 *(a)*, which also holds for mixed MA processes, and (2.2.12) yields to (2.2.15). □

From (2.2.12)-(2.2.13) we see that $Y(t)$ is again regularly varying in the sense of (2.1.2). The following Lemma is an application of Lemma 2.1.3 and Lemma A.4.1.

Lemma 2.2.8
Let Y be a stationary mixed Poisson shot noise process with representation (2.0.7) satisfying (M1). Suppose $I \subseteq \mathbb{Z}$ is finite, or I^c is finite, and define the sequence of random vectors $\mathbf{Y}_k := \sum_{k\in I} \mathbf{f}(R_k,\Gamma_k)Z_k := (Y_1,\dots,Y_d)$ on $\mathbb{R}^d$. Then

$$\lim_{n\to\infty} n\mathbb{P}\left(|\mathbf{Y}_k| > a_n x\right) = x^{-\alpha}/\mu \sum_{k\in I} \mathbb{E}|\mathbf{f}(R_k,\Gamma_k)|^\alpha, \tag{2.2.16}$$

and

$$\begin{aligned}\lim_{n\to\infty} n\mathbb{P}\left(\bigvee_{i=1}^d Y_i > a_n x\right) &= x^{-\alpha}/\mu \sum_{k\in I} p\mathbb{E}(|\mathbf{f}^+(R_k,\Gamma_k)|^\alpha \mathbf{1}_{\{|\mathbf{f}^+(R_k,\Gamma_k)|\geq|\mathbf{f}^-(R_k,\Gamma_k)|\}}) \\ &\quad + q\mathbb{E}(|\mathbf{f}^-(R_k,\Gamma_k)|^\alpha \mathbf{1}_{\{|\mathbf{f}^+(R_k,\Gamma_k)|\leq|\mathbf{f}^-(R_k,\Gamma_k)|\}}).\end{aligned} \tag{2.2.17}$$

Note, that for the mixed Poisson shot noise process Y given by (2.0.7) and a_n given by Lemma 2.2.8 we obtain from (2.2.17)

$$\begin{aligned}\lim_{n\to\infty} n\mathbb{P}\left(Y(t) > a_n x\right) &= x^{-\alpha}/\mu \sum_{k=-\infty}^{\infty} p\mathbb{E}(f^+(R_k,\Gamma_k))^\alpha + q\mathbb{E}(f^-(R_k,\Gamma_k))^\alpha \\ &= x^{-\alpha}/\mu\mathbb{E}\left[\int p(f^+(r,s))^\alpha + q(f^-(r,s))^\alpha \, d\Lambda(s,r)\right] \\ &= x^{-\alpha}\left[\int_{\mathbb{R}_+}\int_{\mathbb{R}} p(f^+(r,s))^\alpha + q(f^-(r,s))^\alpha \, ds\, \pi(dr)\right],\end{aligned}$$

which is (2.2.12), since $\mathbb{P}(L(1) > x) \sim \mu\mathbb{P}(Z_1 > x)$ for $x \to \infty$ by Proposition 1.1.2 *(v)*.

2.2.3 Examples

We will give some typical examples for regularly varying mixed MA processes, which are applied as stochastic volatility models.

Example 2.2.9 (CARMA, FICARMA process)
Let Y be a CARMA(p,q) process with kernel function f as given in Example 1.1.11. We obtain $f \in \mathbb{L}^\delta$ for every $\delta > 0$, as it is bounded and eventually exponential decreasing in $\pm\infty$. By Proposition 2.2.3 and (2.2.12) for every $\alpha > 0$ there exists an CARMA process, which is regularly varying of index $-\alpha$.

The *fractionally integrated CARMA(p,q)* (FICARMA(p,d,q)) process with $d \in (0, 0.5)$ was investigated by Brockwell and Marquardt [39]. Let f be the kernel function of a CARMA(p,q) process with a, b given by (1.1.13). Then the kernel function f_d of the FICARMA(p,d,q) process is defined by the convolution of f with $h(t) = t^{d-1}/\Gamma(d)\,\mathbf{1}_{(0,\infty)}(t)$, $t \in \mathbb{R}$, which results in

$$f_d(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{it\omega}(i\omega)^{-d}\frac{b(i\omega)}{a(i\omega)}\,d\omega \quad \text{for } t \in \mathbb{R}.$$

Note, that f_d is bounded and $f_d(t) \sim t^{d-1}/\Gamma(d)b(0)/a(0)$ as $t \to \infty$, implying $f_d \in \mathbb{L}^\delta$ for $\delta > (1-d)^{-1}$. Taking $(1-d)^{-1} \in (1,2)$ into account, by Lemma 2.2.5 we obtain that there is no FICARMA process with regularly varying tails of index $-\alpha$, $\alpha \leq 1$. □

Example 2.2.10 (supOU processes)
We consider the mixed MA process as given in (2.0.1), where the i. d. i. s. r. m. Λ has generating quadruple (m, σ^2, ν, π) and the kernel function is $f(r,s) = \mathbf{1}_{[0,\infty)}(s)\mathrm{e}^{-rs}$ for $r \in \mathbb{R}_+$, $s \in \mathbb{R}$. Then

$$Y(t) = \int_{\mathbb{R}_+\times\mathbb{R}} \mathbf{1}_{[0,\infty)}(t-s)\mathrm{e}^{-r(t-s)}\,d\Lambda(s,r) \quad \text{for } t \in \mathbb{R}. \tag{2.2.18}$$

An important special case of (2.2.18) is the so called *Ornstein-Uhlenbeck* (OU) process, for which π has only support in some $\lambda > 0$, i. e. $\pi(\lambda) = 1$. For some $\alpha > 0$ we have $f \in \mathbb{L}^\delta(\pi)$ if and only if

$$\int_0^\infty \int_0^\infty \mathrm{e}^{-rs\delta}\,ds\pi(dr) = \int_0^\infty r^{-\delta}\,\pi(dr) < \infty.$$

We assume in the following that $\int_{\mathbb{R}_+} r^{-1}\,\pi(dr) < \infty$ and write $1/\lambda = \int_{\mathbb{R}_+} r^{-1}\,\pi(dr)$. Hence, $f \in \mathbb{L}^\delta(\pi)$ for $\delta \leq 1$. Then $Y(0)$ is i. d. and Y is stationary if and only if $\int_{|x|>2} \log|x|\,\nu(dx) < \infty$. We sketch a short proof: Necessary and sufficient conditions for $Y(0)$ to be i. d. are given by Proposition 2.2.1. Inserting in (2.2.1) the kernel function $f(r,s) = \mathbf{1}_{[0,\infty)}(s)\mathrm{e}^{-rs}$ and substituting rs by u leads to the necessary and sufficient conditions of an OU process with parameter λ to be i. d.. By Sato [138], Theorem 17.5, an OU process exists if and only if $\int_{|x|>2} \log|x|\,\nu(dx) < \infty$. □
The generating triplet is then

$$\begin{aligned} m_Y &= \frac{1}{\lambda}\left[m + \int_{|y|>1} \frac{y}{|y|}\,\nu(dy)\right], \\ \sigma_Y^2 &= \sigma^2/(2\lambda), \\ \nu_Y\,[x,\infty) &= \int_x^\infty \frac{\nu\,[y,\infty)}{\lambda y}\,dy\,, \quad x > 0. \end{aligned} \tag{2.2.19}$$

Note, that the marginal distribution of $Y(0)$ coincides with the marginal distribution of an OU process with parameter λ. Furthermore, for any regularly varying Lévy processes Y is stationary.

Define the probability measure $\overline{\pi}(dr) := \lambda/r\pi(dr)$ and the i. d. i. s. r. m. $\overline{\Lambda}$ with generating quadruple $(m/\lambda, \sigma^2/\lambda, \nu/\lambda, \overline{\pi})$. Then the finite dimensional distributions of the stochastic process

$$X(t) = \int_{-\infty}^{\infty} \mathrm{e}^{-rt} \int_{-\infty}^{rt} \mathrm{e}^{s}\, d\overline{\Lambda}(r,s) \quad \text{for } t \in \mathbb{R} \tag{2.2.20}$$

coincide with those of Y, i. e. $X \overset{d}{=} Y$ (Barndorff-Nielsen [13], Theorem 3.1). X is called *superposition of Ornstein-Uhlenbeck type* (supOU) processes, since

$$dX(t) = \int_{\mathbb{R}_+} \{-rX(t,dr)\,dt + d\overline{\Lambda}(t,r)\}$$

with $X(t,B) = \int_B \mathrm{e}^{rt} \int_{-\infty}^{rt} \mathrm{e}^{s}\, d\overline{\Lambda}(r,s)$ for $t \in \mathbb{R}$, $B \in \mathcal{B}(\mathbb{R})$ and analog for Y. By (2.2.9) they have the covariance function

$$\gamma(h) = \mathrm{Cov}(Y(0),Y(h)) = \int_{\mathbb{R}_+} \frac{1}{2r}\mathrm{e}^{-rh}\,\pi(dr) = \frac{1}{2\lambda}\int_{\mathbb{R}_+} \mathrm{e}^{-rh}\,\overline{\pi}(dr) \quad \text{for } h \in \mathbb{R}$$

and the correlation function

$$\rho(h) = \lambda \int_{\mathbb{R}_+} \frac{1}{r}\mathrm{e}^{-rh}\,\pi(dr) = \int_{\mathbb{R}_+} \mathrm{e}^{-rh}\,\overline{\pi}(dr) \quad \text{for } h \in \mathbb{R}. \tag{2.2.21}$$

The class of achievable correlation functions is very large and includes covariance functions of long memory processes, e. g. π is a $\Gamma(2H+1,1)$ d. f. with $H > 0$. Recall that the density of a $\Gamma(\mu,\gamma)$ distribution, $\mu > 1$, $\gamma > 0$ is $p_{\gamma,\mu}(r) = \gamma^{\mu}/\Gamma(\mu) r^{\mu-1}\mathrm{e}^{-\gamma r}$ for $r \geq 0$. Then $\overline{\pi}$ is $\Gamma(2H,1)$. Thus the correlation function is

$$\begin{aligned}\rho(h) &= \int_0^{\infty} \mathrm{e}^{-rh} p_{2H,1}(r)\,dr = \int_0^{\infty} \frac{r^{2H-1}}{\Gamma(2H)}\mathrm{e}^{-r(h+1)}\,dr \\ &= (h+1)^{-2H} \int_0^{\infty} p_{2H,h+1}(r)\,dr = (h+1)^{-2H} \quad \text{for } h \in \mathbb{R},\end{aligned}$$

which implies that Y is long range dependent.

The class of *selfdecomposable distributions* plays, in this context, a very important role. A review about selfdecomposable distributions is given in the monograph of Sato [138]. A probability distribution μ with characteristic function φ_μ is called selfdecomposable, if for every $b > 1$, there is a probability distribution ρ_b with characteristic function ϕ_b such that $\varphi_\mu(u) = \varphi_\mu(b^{-1}u)\phi_b(u)$ for $u \in \mathbb{R}$. The distribution ρ_b is unique and i. d.. Moreover, μ is i. d. with generating triplet $(m_\mu, \sigma_\mu^2, \nu_\mu)$. Lévy processes corresponding to selfdecomposable distributions are called *selfdecomposable processes*. Typical examples for selfdecomposable distributions are stable, exponential, Pareto, lognormal, logistic and GIG distributions (see Chapter 3).

If the r. v. V with distribution μ is selfdecomposable, then there exists a Lévy process L with generating triplet (m, σ^2, ν) such that for some $\lambda > 0$

$$V \overset{d}{=} \int_0^\infty \mathrm{e}^{-\lambda s}\, dL(s). \tag{2.2.22}$$

By (2.2.2) the generating triplet of V and L are related to each other via

$$m_\mu = \frac{1}{\lambda}m + \frac{1}{\lambda}\int_{|y|>1}\frac{y}{|y|}\nu(dy), \quad \sigma_\mu^2 = \frac{1}{2\lambda}\sigma^2, \quad \nu_\mu\,[x,\infty) = \frac{1}{\lambda}\int_x^\infty \frac{\nu\,[y,\infty)}{y}\,dy.$$

In particular, ν_μ has a Lebesgue density given by

$$\nu_\mu(dx) = \frac{\nu\,[x,\infty)}{\lambda x}\,dx \quad \text{for } x > 0, \quad \nu_\mu(dx) = \frac{\nu\,(-\infty,x]}{\lambda|x|}\,dx \quad \text{for } x < 0.$$

Because of (2.2.19) we know, that there exists a supOU process with the same marginal distribution as V.

If the marginal distribution μ with generating triplet $(m_\mu, \sigma_\mu^2, \nu_\mu)$ and λ are known, the generating triplet of L is unique. We denote by ν_μ' the Lévy density of ν_μ and assume that it is differentiable. Then

$$m = \lambda m_\mu - \int_{|y|>1}\frac{y}{|y|}[-\lambda\nu_\mu'(y) - \lambda y\nu_\mu''(y)]\,dy, \sigma^2 = 2\lambda\sigma_\mu^2, \nu\,[x,\infty) = \lambda x\nu_\mu'(x). \tag{2.2.23}$$

Using instead of the Lévy process L the Lévy process $\overline{L} = \{L(t/\lambda)\}_{t\in\mathbb{R}}$, which has generating triplet $(m_{\overline{L}}, \sigma_{\overline{L}}^2, \nu_{\overline{L}}) = (m/\lambda, \sigma^2/\lambda, \nu/\lambda)$, (2.2.23) leads to

$$m_{\overline{L}} = m_\mu - \int_{|y|>1}\frac{y}{|y|}[-\nu_\mu'(y) - y\nu_\mu''(y)]\,dy,\ \sigma_{\overline{L}}^2 = 2\sigma_\mu^2,\ \nu_{\overline{L}}\,[x,\infty) = x\nu_\mu'(x) \tag{2.2.24}$$

and $V \overset{d}{=} \int_0^\infty \mathrm{e}^{-s}\, d\overline{L}(s)$, which is independent of the parameter λ. Changing L to $\overline{L}$ is equivalent to changing Λ to $\overline{\Lambda}$.

Estimating the correlation function of data yields, by (2.2.21), to an estimate for π. Thus, we can model data by a supOU process considering the selfdecomposable marginal distribution and the correlation function. More about supOU models and applications to financial data can be found in Barndorff-Nielsen and Shephard [14, 15]. □

2.3 Regularly varying shot noise processes

Let N be a point process with jump times $\Gamma = \{\Gamma_k\}_{k\in\mathbb{Z}}$ labelled such that

$$-\infty < \ldots < \Gamma_{-1} < \Gamma_0 \le 0 < \Gamma_1 < \ldots < \infty.$$

If the inter-arrival times $\{\Gamma_{k+1} - \Gamma_k\}_{k\in\mathbb{Z}\setminus\{0\}}$ are i. i. d. the counting process N is said to be a *renewal process*. We are concerned with a *stationary renewal process* with intensity μ. This means that $\{\Gamma_{k+1} - \Gamma_k\}_{k\in\mathbb{Z}\setminus\{0\}}$ is i. i. d. positive with d. f. H, $H(0+) = 0$, independent of $\Gamma_1 - \Gamma_0$, $\mathbb{P}(\Gamma_0 < -x, \Gamma_1 > y) = \mu \int_{x+y}^{\infty}(1 - H(s))\,ds$ for $x, y \ge 0$ and $1/\mu = \mathbb{E}(\Gamma_2 - \Gamma_1)$. The name stationary arises since for all $h \in \mathbb{R}$, $n \in \mathbb{N}$, $-\infty < t_1 < \ldots < t_n < \infty$, $h_i > 0$, $s_i = t_i + h_i$, $i = 1, \ldots, n$,

$$(N\,(t_1 + h, s_1 + h]\,, \ldots, N\,(t_n + h, s_n + h]) \stackrel{d}{=} (N\,(t_1, s_1]\,, \ldots, N\,(t_n, s_n]).$$

Furthermore, the *renewal function* $\mathbb{E}N\,(t_1, t_2]$ has value $\mu(t_2-t_1)$ for $t_2-t_1 \ge 0$. More about stationary renewal processes on $\mathbb{R}_+$ can be found in Resnick [126], Section 3.9, and on $\mathbb{R}$ in Karlin and Taylor [82], Chapter 9, Theorem 9.1.

Suppose $\{Z_k\}_{k\in\mathbb{Z}}$ is an i. i. d. sequence of r. v. s and $f : \mathbb{R} \to \mathbb{R}$ is a measurable function. Then

$$Y(t) = \sum_{j=-\infty}^{\infty} f(t - \Gamma_j) Z_j \quad \text{for } t \in \mathbb{R} \tag{2.3.1}$$

is called a *renewal shot noise process*. Often f and Z have support only on the positive real line. We will show in Proposition 2.3.2 that, if Γ is a stationary renewal process, then Y is stationary, too. In this case, we call Y a *stationary renewal shot noise process*. An important special case of a stationary renewal shot noise process is a *Poisson shot noise process*, i. e. a MA process driven by a compound Poisson process. More about Poisson shot noise processes can be found in Example 1.1.9. Necessary and sufficient conditions for the existence of a Poisson shot noise process are given by Proposition 2.2.1. They are satisfied if and only if

$$\begin{aligned}
\widetilde{I}_1 &= \int_{-\infty}^{\infty} \left|\mathbb{E}(f(s)Z\,\mathbf{1}_{\{|f(s)Z|\le 1\}})\right|\,ds < \infty,\\
\widetilde{I}_2 &= \int_{-\infty}^{\infty} \mathbb{E}(f(s)^2 Z^2\,\mathbf{1}_{\{|f(s)Z|\le 1\}})\,ds < \infty, \qquad (2.3.2)\\
\widetilde{I}_3 &= \int_{-\infty}^{\infty} \mathbb{P}(|f(s)Z| > 1)\,ds < \infty.
\end{aligned}$$

These are analog conditions to the three-series theorem (cf. Billingsley [28], Theorem 22.8). Similar conditions are also sufficient for a stationary renewal shot noise process. The assumptions of the following Proposition are analogous to the discrete-time MA processes (Mikosch and Samorodnitsky [112], Lemma A.3). However, Resnick and Willekens [123] have similar conditions for the multivariate case.

Proposition 2.3.1
Let $Z = \{Z_k\}_{k\in\mathbb{Z}}$ be an i. i. d. sequence of regularly varying r. v. s of index $-\alpha$, $\alpha > 0$, in the sense of (2.1.2) with measure σ given by (2.0.9) and $\xi = \{\xi_k\}_{k\in\mathbb{Z}}$ be a sequence of uniformly bounded r. v. s independent of Z satisfying

$$\sum_{k=-\infty}^{\infty} \mathbb{E}|\xi_k|^\delta < \infty \qquad \text{for some } \delta < \alpha. \tag{2.3.3}$$

Assume that one of the following conditions hold:

(a) $\alpha \in (0,1]$ *and* $\delta < \alpha$.

(b) $\alpha \in (1,2)$ *and* $\delta \leq 1$.

(c) $\alpha \in (1,2)$, $\delta < \alpha$, $\delta \in (1,2)$ *and* Z_1 *is symmetric.*

Then $\sum_{k=-\infty}^{\infty} \xi_k Z_k$ *exists a. s. and*

$$\lim_{n\to\infty} n\mathbb{P}\left(\sum_{k=-\infty}^{\infty} \xi_k Z_k > a_n x\right) = x^{-\alpha} \sum_{k=-\infty}^{\infty} \left[p\mathbb{E}\xi_k^{+\,\alpha} + q\mathbb{E}\xi_k^{-\,\alpha}\right]. \tag{2.3.4}$$

Proof.
Step 1. We show $\sum_{k=-\infty}^{\infty} \xi_k Z_k < \infty$ a. s..

Let $\delta \leq 1$. Then

$$\mathbb{E}\left[\sum_{k=-\infty}^{\infty} |\xi_k Z_k|^\delta\right] = \mathbb{E}|Z_1|^\delta \sum_{k=-\infty}^{\infty} \mathbb{E}|\xi_k|^\delta < \infty,$$

where we used monotone convergence, (2.0.11) and the independence of ξ and Z. Thus, $\sum_{k=-\infty}^{\infty} |\xi_k Z_k|^\delta < \infty$ a. s.. By Jensen's-inequality for sequences we obtain $\sum_{k=-\infty}^{\infty} |\xi_k Z_k| \leq (\sum_{k=-\infty}^{\infty} |\xi_k Z_k|^\delta)^{1/\delta} < \infty$ a. s., which is the existence in *(a)* and *(b)*.

In the case *(c)*, where $\delta \in (1,2)$, we use the same techniques as Chatterji [42]. Let $c = \sup_{x\in\mathbb{R}}\{(|1+x| - 1 - \delta x)/|x|^\delta\} \leq 2^{2-\delta}$. Then

$$\begin{aligned} |x+y|^\delta &= |y|^\delta \frac{|x|^\delta}{|y|^\delta}\left[\left|1+\frac{y}{x}\right|^\delta - 1 - \delta\frac{y}{x}\right] + |x|^\delta + \delta |x|^\delta \frac{y}{x} \\ &\leq 2^{2-\delta}|y|^\delta + |x|^\delta + \delta|x|^{\delta-1}\mathrm{sign}(x) y. \end{aligned}$$

By induction and $\sum_{k=n+1}^n := 0$ we obtain for $n > m \geq 1$

$$\left|\sum_{k=m}^n \xi_k Z_k\right|^\delta \leq \sum_{k=m}^n \left[2^{2-\delta}|\xi_k Z_k|^\delta + \delta\left|\sum_{j=k+1}^n \xi_j Z_j\right|^{\delta-1} \mathrm{sign}\left(\sum_{j=k+1}^n \xi_j Z_j\right)\xi_k Z_k\right]. \quad (2.3.5)$$

Taking the independence of Z_k and $|\sum_{j=k+1}^n \xi_j Z_j|^{\delta-1}\mathrm{sign}\left(\sum_{j=k+1}^n \xi_j Z_j\right)\xi_k$ as well as $\mathbb{E}Z_k = 0$ into account and applying (2.3.5) we get

$$\mathbb{E}\left|\sum_{k=m}^n \xi_k Z_k\right|^\delta \leq 2^{2-\delta}\mathbb{E}|Z_1|^\delta \sum_{k=m}^n \mathbb{E}|\xi_k|^\delta.$$

Hence, from (2.3.3) follows that $\{\sum_{k=1}^n \xi_k Z_k\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{L}^\delta$. The space $\mathbb{L}^\delta$ is complete for $\delta > 1$, which implies $\mathbb{E}|\sum_{k=1}^\infty \xi_k Z_k|^\delta < \infty$. Consequently in *(c)* we have $\sum_{k=-\infty}^\infty \xi_k Z_k < \infty$ a. s., too.

Step 2. We compute the tail behavior in (2.3.4).

We show that there exist a $K, y_0 > 0$ such that for $y \geq y_0$

$$\mathbb{P}\left(\left|\sum_{k=-\infty}^\infty \xi_k Z_k\right| > y\right) \leq K\mathbb{P}\left(|Z_1| > y/f^+\right)\sum_{k=-\infty}^\infty \mathbb{E}|\xi_k|^\delta. \quad (2.3.6)$$

Then we have on the one hand

$$\begin{aligned} &\lim_{m\to\infty}\lim_{n\to\infty} n\mathbb{P}\left(\left|\sum_{|k|>m}^\infty \xi_k Z_k\right| > a_n x\right) \quad (2.3.7) \\ &\quad \leq K \lim_{n\to\infty} n\mathbb{P}\left(|Z_1| > a_n x/f^+\right)\lim_{m\to\infty}\sum_{|k|>m}\mathbb{E}|\xi_k|^\delta = 0. \end{aligned}$$

On the other hand, applying (2.1.7) yields, for $n \to \infty$, to

$$\lim_{n\to\infty} n\mathbb{P}\left(\sum_{k=-m}^m \xi_k Z_k > a_n x\right) = x^{-\alpha}\sum_{k=-m}^m \left[p\mathbb{E}{\xi_k^+}^\alpha + q\mathbb{E}{\xi_k^-}^\alpha\right]. \quad (2.3.8)$$

Combining (2.3.7) and (2.3.8) results in (2.3.4); for details see Kokoszka and Taqqu [91], proof of Theorem 2.2.

First, note that

$$\begin{aligned}\mathbb{P}\left(\left|\sum_{k=-\infty}^{\infty}\xi_k Z_k\right|>y\right) &= \mathbb{P}\left(\left|\sum_{k=-\infty}^{\infty}\xi_k Z_k\right|>y,\ \bigvee_{j=-\infty}^{\infty}|\xi_j Z_j|>y\right)\\ &\quad+\mathbb{P}\left(\left|\sum_{k=-\infty}^{\infty}\xi_k Z_k\right|>y,\ \bigvee_{j=-\infty}^{\infty}|\xi_j Z_j|\le y\right) \qquad (2.3.9)\\ &\le \sum_{k=-\infty}^{\infty}\mathbb{P}\left(|\xi_k Z_k|>y\right)+\mathbb{P}\left(\left|\sum_{k=-\infty}^{\infty}\xi_k Z_k\,\mathbf{1}_{\{|\xi_k Z_k|\le y\}}\right|>y\right).\end{aligned}$$

Let G_k be the d. f. of ξ_k, which has bounded support as $|\xi_k|\le f^+$ for $k\in\mathbb{Z}$ and recall that ξ_k is independent of Z_k. By Potter's Theorem A.1.10 *(iv)* there exists a $K_1>1$ such that the first summand of the r. h. s. of (2.3.9) has the upper bound

$$\begin{aligned}\sum_{k=-\infty}^{\infty}\mathbb{P}\left(|\xi_k Z_k|>y\right) &= \sum_{k=-\infty}^{\infty}\int_{-f^+}^{f^+}\mathbb{P}\left(|uZ_k|>y\right)dG_k(u) \qquad (2.3.10)\\ &\le K_1\mathbb{P}\left(|Z_k|>|y|/f^+\right)\sum_{k=-\infty}^{\infty}\int_{-f^+}^{f^+}|u|^{\delta}\,dG_k(u)\\ &= K_1\mathbb{P}\left(|Z_k|>|y|/f^+\right)\sum_{k=-\infty}^{\infty}\mathbb{E}|\xi_k|^{\delta}.\end{aligned}$$

To obtain an upper bound of the second summand of (2.3.9), we will first consider the case where Z_k is symmetric and extend the results afterwards.

Step 2.1. Suppose Z_k is symmetric.

By Markov's inequality and Lemma of Fatou,

$$\begin{aligned}\mathbb{P}\left(\left|\sum_{k=-\infty}^{\infty}\xi_k Z_k\,\mathbf{1}_{\{|\xi_k Z_k|\le y\}}\right|>y\right) &\le \frac{1}{y^2}\mathbb{E}\left|\sum_{k=-\infty}^{\infty}\xi_k Z_k\,\mathbf{1}_{\{|\xi_k Z_k|\le y\}}\right|^2 \qquad (2.3.11)\\ &\le \frac{1}{y^2}\sum_{k=-\infty}^{\infty}\mathbb{E}[\xi_k^2 Z_k^2\,\mathbf{1}_{\{|\xi_k Z_k|\le y\}}]+\frac{1}{y^2}\sum_{\substack{k,j=-\infty\\k\neq 0}}^{\infty}\mathbb{E}[\xi_k Z_k\,\mathbf{1}_{\{|\xi_k Z_k|\le y\}}\,\xi_j Z_j\,\mathbf{1}_{\{|\xi_j Z_j|\le y\}}].\end{aligned}$$

By symmetry and independence, the second summand is 0. By the independence of ξ_k and Z_k we have

$$\mathbb{E}[\xi_k^2 Z_k^2\,\mathbf{1}_{\{|\xi_k Z_k|\le y\}}]=\int_{-f^+}^{f^+}\mathbb{E}[u^2 Z_1^2\,\mathbf{1}_{\{|uZ_1|\le y\}}]\,dG_k(u). \qquad (2.3.12)$$

Taking $\alpha \in (0,2)$ into account and applying Theorem 1, (9.5) of Feller [65], VIII.9, reveals that for fixed u the map $(y \mapsto \mathbb{E}[u^2 Z_1^2 \mathbf{1}_{\{|uZ_1|\le y\}}])$ is a regularly varying function of index $2-\alpha$ such that

$$\mathbb{E}[u^2 Z_1^2 \mathbf{1}_{\{|uZ_1|\le y\}}] \sim \frac{2}{2-\alpha} y^2 \mathbb{P}(|uZ_1| > |y|) \quad \text{for } y \to \infty.$$

Applying again Potter's Theorem A.1.10, there exists a $K_2 > 1$ such that

$$\mathbb{E}[u^2 Z_1^2 \mathbf{1}_{\{|uZ_1|\le y\}}] \le K_2 y^2 |u|^\delta \mathbb{P}\left(|Z_1| > y/f^+\right) \quad \text{for every } u \in [-f^+, f^+],\ y > 0.$$

Together with (2.3.12) we obtain

$$\sum_{k=-\infty}^{\infty} \mathbb{E}[\xi_k^2 Z_k^2 \mathbf{1}_{\{|\xi_k Z_k|\le y\}}] \le K_2 y^2 \mathbb{P}\left(|Z_1| > y/f^+\right) \sum_{k=-\infty}^{\infty} \mathbb{E}|\xi_k|^\delta. \tag{2.3.13}$$

Thus, from (2.3.11) and (2.3.13) the inequality (2.3.6) follows.

Step 2.2. Let $\{Z_k\}_{k\in\mathbb{Z}}$ and $\{\xi_k\}_{k\in\mathbb{Z}}$ be positive sequences and $\delta \le 1$, i.e. $p=1$ and $q=0$ in (2.0.9). We apply a technique used by Mikosch and Samorodnitsky [112], Lemma A.3:

Define $X^m = \sum_{|k|>m} \xi_k Z_k$ for $m \in \mathbb{Z}$. Since $X^m \to 0$ as $m \to \infty$ a.s. by Step 1 the family of r.v.s $\{X^m\}_{m\in\mathbb{N}}$ is tight. Therefore there exists an $N > 0$ such that $\mathbb{P}(X^m \le N) \ge 1/2$ for every $m \in \mathbb{N}$. Let $\widetilde{X}^m = \sum_{|k|>m} \widetilde{\xi}_k \widetilde{Z}_k$ be an independent copy of X^m, such that $\{\widetilde{Z}_k\}_{k\in\mathbb{Z}}$, $\{\widetilde{\xi}_k\}_{k\in\mathbb{Z}}$ are independent copies of $\{Z_k\}_{k\in\mathbb{Z}}$, $\{\xi_k\}_{k\in\mathbb{Z}}$. Then for some $\epsilon \in (0,x)$

$$\begin{aligned}
\frac{1}{2}\mathbb{P}(X^m > a_n x) &\le \mathbb{P}\left(\widetilde{X}^m \le N, X^m > a_n x\right) \\
&= \mathbb{P}\left(\widetilde{X}^m \le N, X^m > a_n x, \sum_{|k|>m} \widetilde{\xi}_k Z_k > a_n \epsilon\right) \\
&\quad + \mathbb{P}\left(\widetilde{X}^m \le N, X^m > a_n x, \sum_{|k|>m} \widetilde{\xi}_k Z_k \le a_n \epsilon\right).
\end{aligned} \tag{2.3.14}$$

For the first summand of (2.3.14) we obtain the estimate

$$\begin{aligned}
&\mathbb{P}\left(\widetilde{X}^m \le N, X^m > a_n x, \sum_{|k|>m} \widetilde{\xi}_k Z_k > a_n \epsilon\right) \\
&\quad \le \mathbb{P}\left(\sum_{|k|>m} \widetilde{\xi}_k (Z_k - \widetilde{Z}_k) > a_n \epsilon - N\right).
\end{aligned} \tag{2.3.15}$$

The second summand of (2.3.14) has the upper bound

$$\begin{aligned}
&\mathbb{P}\left(\widetilde{X}^m \le N, X^m > a_n x, \sum_{|k|>m} \widetilde{\xi}_k Z_k \le a_n \epsilon\right) \\
&\quad \le \mathbb{P}\left(\sum_{|k|>m} (\widetilde{\xi}_k - \xi_k)\widetilde{Z}_k \le N, X^m > a_n x, \sum_{|k|>m} \widetilde{\xi}_k Z_k \le a_n \epsilon\right) \\
&\quad \le \mathbb{P}\left(\sum_{|k|>m} (\xi_k - \widetilde{\xi}_k)(Z_k - \widetilde{Z}_k) > a_n(x-\epsilon) - N\right), \qquad (2.3.16)
\end{aligned}$$

where we used in the second inequality the positivity of $\xi_k, \widetilde{Z}_k$. Furthermore, $\{Z_k - \widetilde{Z}_k\}_{k\in\mathbb{Z}}$ is symmetric and regularly varying of index $-\alpha$ such that for $n \to \infty$

$$n\mathbb{P}(a_n^{-1}(Z_k - \widetilde{Z}_k) \in \cdot) \stackrel{v}{\Longrightarrow} \widetilde{\sigma}(\cdot),$$

where $\widetilde{\sigma}(\cdot) = \alpha|x|^{-\alpha-1}\, dx$. Recall that

$$\sum_{|k|>m} \mathbb{E}|\widetilde{\xi}_k - \xi_k|^\delta \le 2^{\delta+1} \sum_{|k|>m} \mathbb{E}|\xi_k|^\delta < \infty.$$

Then $\sum_{|k|>m}(\widetilde{\xi}_k - \xi_k)(Z_k - \widetilde{Z}_k)$ satisfies the assumptions of Step 2.1, such that by Definition 1.1.1 *(i)*

$$\begin{aligned}
&\lim_{n\to\infty} n\mathbb{P}\left(\sum_{|k|>m} (\widetilde{\xi}_k - \xi_k)(Z_k - \widetilde{Z}_k) > a_n(x-\epsilon) - N\right) \\
&\quad \le K(x-\epsilon)^{-\alpha} \sum_{|k|>m} \mathbb{E}|\xi_k - \widetilde{\xi}_k|^\delta \stackrel{m\to\infty}{\longrightarrow} 0. \qquad (2.3.17)
\end{aligned}$$

Similarly we obtain in (2.3.15)

$$\lim_{n\to\infty} n\mathbb{P}\left(\sum_{|k|>m} \widetilde{\xi}_k(Z_k - \widetilde{Z}_k) > a_n\epsilon - N\right) \le K\epsilon^{-\alpha} \sum_{|k|>m} \mathbb{E}|\xi_k|^\delta \stackrel{m\to\infty}{\longrightarrow} 0. \qquad (2.3.18)$$

Hence by (2.3.14)-(2.3.18) the relation (2.3.7)

$$\lim_{m\to\infty}\lim_{n\to\infty} \mathbb{P}(|X^m| > a_n x) = \lim_{m\to\infty}\lim_{n\to\infty} n\mathbb{P}\left(\sum_{|k|>m} \xi_k Z_k > a_n x\right) = 0$$

holds.

Step 2.3. Let $\{Z_k\}_{k\in\mathbb{Z}}$, $\{\xi_k\}_{k\in\mathbb{Z}}$ be arbitrary and $\delta \le 1$. Then $\overline{X}^m = \sum_{|k|>m} |\xi_k||Z_k|$ satisfies the assumption of Step 2.2 leading to

$$\lim_{m\to\infty}\lim_{n\to\infty} n\mathbb{P}\left(\left|\sum_{|k|>m}\xi_k Z_k\right| > a_n x\right) \le \lim_{m\to\infty}\lim_{n\to\infty} n\mathbb{P}(\overline{X}^m > a_n x) = 0. \qquad \square$$

With this result we can pose necessary conditions for the existence of stationary regularly varying renewal shot noise processes, which are also regularly varying. The result is very similarly to the existence and tail behavior of mixed MA processes in Proposition 2.2.3.

Proposition 2.3.2
Let Y be a shot noise process given by (2.3.1), where $\Gamma = \{\Gamma_k\}_{k\in\mathbb{Z}}$ is a stationary renewal process with intensity μ and $Z = \{Z_k\}_{k\in\mathbb{Z}}$ is an i. i. d. sequence, which is regularly varying of index $-\alpha$, $\alpha > 0$, in the sense of (2.1.2). Furthermore we assume one of the following conditions to hold:

(a) $\alpha \in (0,1]$ *and* $f \in \mathbb{L}^\delta$ *for some* $\delta < \alpha$.

(b) $\alpha \in (1,2)$ *and* $f \in \mathbb{L}^\delta$ *for some* $\delta \le 1$.

(c) $\alpha \in (1,2)$, $f \in \mathbb{L}^\delta$ *for some* $\delta < \alpha$ *and* Z_1 *symmetric.*

Then Y is well defined and stationary with

$$\lim_{n\to\infty} n\mathbb{P}\left(Y(0) > a_n x\right) = \mu x^{-\alpha}\int_{-\infty}^{\infty}\left[p(f^+(s))^\alpha + q(f^-(s))^\alpha\right] ds. \qquad (2.3.19)$$

The process Y is called stationary renewal shot noise process.

Proof.
Define $\xi_k := f(\Gamma_k)$ for $k \in \mathbb{Z}$. For the stationary renewal process Γ and the point process $\kappa = \sum_{k=-\infty}^{\infty}\varepsilon_{\Gamma_k}$ in $\mathbb{R}$ holds

$$\sum_{k=-\infty}^{\infty}\mathbb{E}|\xi_k|^\delta = \mathbb{E}\left[\sum_{k=-\infty}^{\infty}|f(\Gamma_k)|^\delta\right] = \mathbb{E}\left[\int_{-\infty}^{\infty}|f(s)|^\delta\, d\kappa(s)\right] = \mu\int_{-\infty}^{\infty}|f(s)|^\delta\, ds.$$

Then the existence and (2.3.19) follow by Proposition 2.3.1.

It remains to show that Y is stationary: Let $h > 0$ be fixed. Define $\widetilde{\Gamma}_k := \Gamma_{k+N(h)} - h$ and $\widetilde{Z}_k = Z_{k+N(h)}$ for $k \in \mathbb{Z}$. Then $\{\Gamma_k\}_{k\in\mathbb{Z}} \stackrel{d}{=} \{\widetilde{\Gamma}_k\}_{k\in\mathbb{Z}}$ and $\{Z_k\}_{k\in\mathbb{Z}} \stackrel{d}{=} \{\widetilde{Z}_k\}_{k\in\mathbb{Z}}$. We obtain

$$\begin{aligned} Y(t) &= \sum_{k=-\infty}^{\infty} f(t-\Gamma_k)Z_k \overset{d}{=} \sum_{k=-\infty}^{\infty} f(t-\widetilde{\Gamma}_k)\widetilde{Z}_k = \sum_{k=-\infty}^{\infty} f(t+h-\Gamma_{k+N(h)})Z_{k+N(h)} \\ &= \sum_{k=-\infty}^{\infty} f(t+h-\Gamma_k)Z_k = Y(t+h). \end{aligned}$$

Analogously we receive for the finite dimensional distribution

$$(Y(t_1),\dots,Y(t_n)) \overset{d}{=} (Y(t_1+h),\dots,Y(t_n+h)). \qquad \square$$

Similarly to condition (M1) and (M2) for mixed MA processes we state the conditions for the stationary renewal shot noise processes.

Condition (R1).
Let Y with $Y(t)=\sum_{k=-\infty}^{\infty} f(t-\Gamma_k)Z_k$, $k\in\mathbb{Z}$, be a stationary renewal shot noise process as given in (2.3.1), where the stationary renewal process N with jumps $\Gamma=\{\Gamma_k\}_{k\in\mathbb{Z}}$ has intensity μ and $Z=\{Z_k\}_{k\in\mathbb{Z}}$ is an i.i.d. sequence, which is regularly varying of index $-\alpha$, $\alpha>0$, such that there exists a sequence $0<a_n\uparrow\infty$ of constants with

$$n\mathbb{P}(a_n^{-1}Z_k\in\cdot) \overset{v}{\Longrightarrow} \sigma(\cdot)/\mu \quad \text{on } \mathcal{B}(\overline{\mathbb{R}}\backslash\{0\}) \quad \text{for } n\to\infty,$$

where $\sigma(dx)=p\alpha x^{-\alpha-1}\mathbf{1}_{(0,\infty)}(x)\,dx+q\alpha(-x)^{-\alpha-1}\mathbf{1}_{(-\infty,0)}(x)\,dx$ for some $p\in[0,1]$, $q=1-p$. Furthermore, $f:\mathbb{R}\to\mathbb{R}$ is bounded with $f^+=\sup_{t\in\mathbb{R}} f^+(t)=f(\eta^{(1)})<\infty$ and $f^-=\sup_{t\in\mathbb{R}} f^-(t)=f(\eta^{(2)})\le f^+$. Assume that one of the following conditions hold:

- *(a)* $\alpha\in(0,1]$ *and* $f\in\mathbb{L}^\delta$ *for some* $\delta<\alpha$.
- *(b)* $\alpha\in(1,2)$ *and* $f\in\mathbb{L}^\delta$ *for some* $\delta\le 1$.
- *(c)* $\alpha\in(1,2)$, $f\in\mathbb{L}^\delta$ *for some* $\delta<\alpha$ *and* Z_1 *symmetric.*

Condition (R2).
Let Y be a stationary renewal shot noise process given by (2.3.1) satisfying (R1) with a.s. sample paths in $\mathbb{D}(\mathbb{R})$. Define

$$g(t) := \sup_{s\in(-\infty,t]} |f(s)|\,\mathbf{1}_{\left(-\infty,\eta^{(1)}\right]}(t) + \sup_{s\in(t,\infty)} |f(s)|\,\mathbf{1}_{\left(\eta^{(1)},\infty\right)}(t), \qquad (2.3.20)$$

then we suppose $g\in\mathbb{L}^\delta$ for some $\delta<\min\{1,\alpha\}$.

2.4 Point process convergence

We follow Resnick [124, 125] and introduce point processes to describe precisely the extremal behavior of Y. In order to achieve a distributional stability of a sequence of point processes, it is necessary to allow a build up of infinite mass at $[s,t)\times\{\mathbf{0}\}$. This is handled, for our problem, by defining the state space $S=[0,\infty)\times\overline{\mathbb{R}}^d\backslash\{\mathbf{0}\}$. The space S can be metricized as a locally compact, complete and separable Hausdorff space (cf. Lindskøg [99], Theorem 1.5; note that $(0,\infty]$ equipped with the metric $d(x,y)=|1/x-1/y|$ is complete and separable). Compact sets in S are closed sets, which are bounded away from $\mathbf{0}$ and $\pm\infty$. Furthermore $\mathcal{B}(S)$ denotes the Borel σ-field and $M_P(S)$ the class of *point measures* on S, where $M_P(S)$ is equipped with the metric ρ that generates the topology of vague convergence. The space $(M_P(S),\rho)$ is a complete and separable metric space with Borel σ-field $\mathcal{M}_P(S)$. The zero measure is denoted by $\mathbf{0}$. A *point process* in S is a random element in $(M_P(S),\mathcal{M}_P(S))$, i. e. a measurable map from a probability space $(\Omega,\mathcal{A},\mathbb{P})$ into $(M_P(S),\mathcal{M}_P(S))$. A typical example for a point process in extreme value theory is a Poisson random measure, i. e. given a Radon measure ϑ on $\mathcal{B}(S)$, a point process κ is called *Poisson random measure* with mean measure (or intensity measure) ϑ, denoted by PRM(ϑ), if

(a) $\kappa(A)$ is Poisson distributed with mean $\vartheta(A)$ for every $A\in\mathcal{B}(S)$,

(b) for mutually disjoint sets $A_1,\ldots,A_n\in\mathcal{B}(S)$, $n\in\mathbb{N}$, the r. v. s $\kappa(A_1),\ldots,\kappa(A_n)$ are independent.

More about point processes can be found in Daley and Vere-Jones [50] and Kallenberg [81].

For a thorough understanding of the structural behavior of extremes of the continuous-time processes, the knowledge of point processes is desirable. Therefore the results of Davis and Hsing [52] about the point process behavior of a stationary sequence of regularly varying r. v. s under weak dependence are of vitally importance for our studies. Those results were generalized by Davis and Mikosch [53] to multidimensional regularly varying stationary processes. Weak dependence is described by the following mixing condition.

Condition $\mathcal{A}(a_n)$.
Let $\mathbf{Y}=\{\mathbf{Y}_k\}_{k\in\mathbb{Z}}$ be a strictly stationary sequence of regularly varying random

vectors and $0 < a_n \uparrow \infty$ *a sequence of constants satisfying*

$$\lim_{n\to\infty} n\mathbb{P}(|\mathbf{Y}_1| > a_n) = 1. \tag{2.4.1}$$

There exists a set of positive integers $\{r_n\}_{n\in\mathbb{N}}$ *such that* $r_n \to \infty$, $r_n/n \to 0$ *as* $n \to \infty$ *and*

$$\mathbb{E}\exp\left(-\sum_{j=1}^{n} f(\mathbf{Y}_j/a_n)\right) - \left[\mathbb{E}\exp\left(-\sum_{j=1}^{r_n} f(\mathbf{Y}_j/a_n)\right)\right]^{\lfloor n/r_n\rfloor} \overset{n\to\infty}{\longrightarrow} 0 \text{ for all } f \in \mathcal{F}_s,$$

where $\mathcal{F}_s$ *is the collection of bounded non negative step functions on* $\overline{\mathbb{R}}^d\backslash\{\mathbf{0}\}$ *with bounded support.*

The meaning of this condition is that, if $\{\overline{N}_k^{(n)}\}_{k\in\mathbb{N}}$ is an i. i. d. sequence of point processes with the same distribution as $\sum_{k=1}^{r_n} \varepsilon_{\mathbf{Y}_k/a_n}$ and $N_n = \sum_{k=1}^{n} \varepsilon_{\mathbf{Y}_k/a_n}$, then the Laplace functional (Embrechts et al. [60], Theorem 5.2.3) of N_n and $\sum_{k=1}^{\lfloor n/r_n\rfloor} \overline{N}_k^{(n)}$ have the same limit behavior, i. e.

$$\mathbb{E}\exp\left[-\int f(s)\,dN_n(s)\right] - \left[\mathbb{E}\exp\left(-\int f(s)\,d\overline{N}_1^{(n)}(s)\right)\right]^{\lfloor n/r_n\rfloor} \overset{n\to\infty}{\longrightarrow} 0.$$

This is equivalent to

$$\sum_{k=1}^{\lfloor n/r_n\rfloor} \overline{N}_k^{(n)} \overset{w}{\Longrightarrow} \kappa \quad \Longleftrightarrow \quad N_n \overset{w}{\Longrightarrow} \kappa$$

for $n \to \infty$ and some point process κ in $M_P(\overline{\mathbb{R}}^d\backslash\{\mathbf{0}\})$. Condition $\mathcal{A}(a_n)$ is a weak mixing condition and holds, e. g. for a sequence of regularly varying random vectors satisfying the strong mixing condition (Basrak [18], Lemma 2.3.9). It is independent of the particular choice of a_n. The condition is very similar to the $\Delta(a_n)$ condition often used in extreme value theory (cf. Leadbetter and Rootzén [97], Lemma 2.4.2).

For stationary sequences of multivariate regularly varying random vectors satisfying condition $\mathcal{A}(a_n)$ Davis and Mikosch [53] formulate the following point process result.

Theorem 2.4.1 (Davis and Mikosch [53], Theorem 2.8 and Corollary 2.4) *Let* $\mathbf{Y} = \{\mathbf{Y}_k\}_{k\in\mathbb{Z}}$ *be a strictly stationary sequence of random vectors on* $\mathbb{R}^d$ *satisfying condition* $\mathcal{A}(a_n)$*. Assume that all finite-dimensional distributions of* $\mathbf{Y}$ *are jointly regularly varying with index* $-\alpha$, $\alpha > 0$*, such that for every* $l > 0$ *with*

$\widetilde{\mathbf{Y}}^{(l)} = (\mathbf{Y}_{-l}^t, \ldots, \mathbf{Y}_l^t)$ *there exists a random vector* $\Theta^{(l)}$ *with values on the unit sphere* $\mathbb{S}^{(2l+1)d-1} = \{\mathbf{x} \in \mathbb{R}^{(2l+1)d} : |\mathbf{x}| = 1\}$ *such that for every* $x > 0$ *and* $u \to \infty$,

$$\frac{\mathbb{P}(|\widetilde{\mathbf{Y}}^{(l)}| > ux, \widetilde{\mathbf{Y}}^{(l)}/|\widetilde{\mathbf{Y}}^{(l)}| \in \cdot)}{\mathbb{P}(|\widetilde{\mathbf{Y}}^{(l)}| > u)} \stackrel{v}{\Longrightarrow} x^{-\alpha}\mathbb{P}(\Theta^{(l)} \in \cdot) \quad \text{on } \mathcal{B}(\mathbb{S}^{(2l+1)d-1}). \tag{2.4.2}$$

Assume that

$$\lim_{l\to\infty}\lim_{n\to\infty}\mathbb{P}\left(\bigvee_{l\le|k|\le r_n}|\mathbf{Y}_k| > a_n x \,\middle|\, |\mathbf{Y}_0| > a_n x\right) = 0 \quad \text{for } x > 0. \tag{2.4.3}$$

Define the point processes $N_n = \sum_{k=1}^n \varepsilon_{\mathbf{Y}_k/a_n}$ *on* $M_P(\overline{\mathbb{R}}^d\backslash\{\mathbf{0}\})$. *Then the limit*

$$\theta = \lim_{l\to\infty}\mathbb{E}\left(|\Theta_0^{(l)}|^\alpha - \bigvee_{j=1}^{l}|\Theta_j^{(l)}|^\alpha\right)^+ \Big/ \mathbb{E}|\Theta_0^{(l)}|^\alpha \tag{2.4.4}$$

exists. If $\theta = 0$ *then* $N_n \stackrel{w}{\Longrightarrow} \mathbf{0}$ *for* $n \to \infty$. *Else, if* $\theta > 0$, *then for* $l \to \infty$

$$\frac{\mathbb{E}\left(\left[|\Theta_0^{(l)}|^\alpha - \bigvee_{j=1}^{l}|\Theta_j^{(l)}|^\alpha\right]^+ \mathbf{1}\left\{\sum_{|j|\le l}\varepsilon_{\Theta_j^{(l)}} \in \cdot\right\}\right)}{\mathbb{E}\left(|\Theta_0^{(l)}|^\alpha - \bigvee_{j=1}^{l}|\Theta_j^{(l)}|^\alpha\right)^+} \stackrel{w}{\Longrightarrow} \mathcal{Q}(\cdot), \tag{2.4.5}$$

where $\mathcal{Q}$ *is a Radon counting measure on the set*

$$\{\mu \text{ Radon counting measure on } \overline{\mathbb{R}}^d\backslash\{\mathbf{0}\}, \mu(\mathbb{S}^{d-1}) > 0, \mu\{(\mathbf{x} \in \mathbb{R}^d : |\mathbf{x}| > 1)\} = 0\}.$$

Furthermore, for $n \to \infty$,

$$N_n \stackrel{w}{\Longrightarrow} \sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\varepsilon_{\widetilde{P}_k\mathbf{Q}_{kj}},$$

where $\sum_{k=1}^\infty \varepsilon_{\widetilde{P}_k}$ *is a* PRM($\widetilde{\vartheta}$) *on* $\mathbb{R}_+$ *with* $\widetilde{\vartheta}(dx) = \theta\alpha\, x^{-\alpha-1}\,\mathbf{1}_{(0,\infty)}(x)\,dx$ *and* $\sum_{j=1}^\infty \varepsilon_{\mathbf{Q}_{kj}}$ *for* $k \in \mathbb{N}$ *are point processes with the same distribution as* $\mathcal{Q}$. *All point processes are mutually independent.*

Note that θ is the extremal index (Definition A.1.9) of $\{|\mathbf{Y}_k|\}_{k\in\mathbb{Z}}$. Condition (2.4.3) is an anti-clustering condition which, roughly speaking, ensures that high level exceedances have not too much dependence on any short time horizon. In comparison to the $D'(u_n)$ condition, Definition A.3.1, we have the rate r_n instead of n. Long range dependence has already been excluded by the condition $\mathcal{A}(a_n)$. The independence of $\widetilde{P}_k$ and the $\mathbf{Q}_{kj}$ is a consequence of the multivariate regular variation.

There exist further results on the point process behavior of regularly varying random vectors. The following Lemma shows that adding a small random vector to a multivariate regularly varying sequence has no influence on the point process behavior. The meaning of small is that the tail of the absolute value decreases faster than the tail of the absolute value of the multivariate sequence.

Lemma 2.4.2

Let $\mathbf{Z} = \{\mathbf{Z}_k\}_{k\in\mathbb{N}}$, $\mathbf{\Psi} = \{\mathbf{\Psi}_k\}_{k\in\mathbb{N}}$ be sequences of random vectors in $\mathbb{R}^d$. Furthermore, let $\{\Gamma_k\}_{k\in\mathbb{N}}$ be the jump times of a stationary renewal process N with intensity $\mu > 0$, $h \in \mathbb{R}$ be arbitrary and $s_k \in [\Gamma_{k-1} + h, \Gamma_{k+1} + h)$ for $k \in \mathbb{N}$, setting $\Gamma_0 := 0$. Denote by $0 < a_n \uparrow \infty$ a sequence of constants and by

$$\widetilde{\kappa}_n = \sum_{k=1}^{\infty} \varepsilon_{(k/n,\mathbf{Z}_k/a_n)} \qquad \text{and} \qquad \kappa_n = \sum_{k=1}^{\infty} \varepsilon_{(s_k\mu/n,(\mathbf{Z}_k+\mathbf{\Psi}_k)/a_n)}$$

point processes in $M_P(S)$ for $n \in \mathbb{N}$. Suppose there exists a point process κ in $M_P(S)$ with $\kappa\left([s,t) \times \{\mathbf{x}\}\right) = 0$ a.s. for $\mathbf{x} \in \overline{\mathbb{R}}^d \backslash \{\mathbf{0}\}$, $t > s \geq 0$, such that $\widetilde{\kappa}_n \overset{w}{\Longrightarrow} \kappa$ for $n \to \infty$. Furthermore assume that for every $\epsilon, t > 0$,

$$\sum_{k=1}^{\lfloor nt \rfloor} \mathbb{P}(|\mathbf{\Psi}_k| > a_n\epsilon) \overset{n\to\infty}{\longrightarrow} 0. \tag{2.4.6}$$

We suppose that there exists a r. v. W such that

$\mathbb{P}(|\mathbf{Z}_k + \mathbf{\Psi}_k| > x) \leq \mathbb{P}(W > x)$ *for* $x > 0$ *and* $\mathbb{P}(W > a_n x) = O(1/n)$ *for* $n \to \infty$.

Let $I = [s,t) \times \prod_{i=1}^d (c_i, d_i] \subseteq S$. Then

$$\lim_{n\to\infty} \mathbb{P}(\kappa_n(I) \neq \widetilde{\kappa}_n(I)) = 0$$

and $\kappa_n \overset{w}{\Longrightarrow} \kappa$ for $n \to \infty$.

Proof.

Let $\epsilon > 0$ be arbitrary. Denote by $\zeta_n := \sum_{k=1}^{\infty} \varepsilon_{\left(k/n, a_n^{-1}(\mathbf{Z}_k+\mathbf{\Psi}_k)\right)}$ a point process in $M_P(S)$ for $n \in \mathbb{N}$. Define the sets $I_\epsilon^{(1)} = \prod_{i=1}^{d} (c_i - \epsilon, d_i + \epsilon]$, $I_\epsilon^{(2)} = \prod_{i=1}^{d} (c_i + \epsilon, d_i - \epsilon]$ and $I_\epsilon = I_\epsilon^{(1)} \backslash I_\epsilon^{(2)}$. We obtain

$$\begin{aligned} \{\widetilde{\kappa}_n(I) \neq \zeta_n(I)\} \;\subseteq\; & \{\widetilde{\kappa}_n(I_\epsilon) > 0\} \\ & \bigcup_{k\in(ns,nt]} \left\{a_n^{-1}(\mathbf{Z}_k + \mathbf{\Psi}_k) \in I, a_n^{-1}\mathbf{Z}_k \in I_\epsilon^{(1)\,c}\right\} \\ & \bigcup_{k\in(ns,nt]} \left\{a_n^{-1}(\mathbf{Z}_k + \mathbf{\Psi}_k) \in I^c, a_n^{-1}\mathbf{Z}_k \in I_\epsilon^{(2)}\right\}. \end{aligned} \tag{2.4.7}$$

On the one hand,

$$\sum_{k\in(ns,nt]} \mathbb{P}\left(a_n^{-1}(\mathbf{Z}_k+\mathbf{\Psi}_k)\in I, a_n^{-1}\mathbf{Z}_k\in I_\epsilon^{(1)\,c}\right) \leq \sum_{k\in(ns,nt]} \mathbb{P}(|\mathbf{\Psi}_k|>a_n\epsilon) \overset{n\to\infty}{\longrightarrow} 0, \quad (2.4.8)$$

$$\sum_{k\in(ns,nt]} \mathbb{P}\left(a_n^{-1}(\mathbf{Z}_k+\mathbf{\Psi}_k)\in I^c, a_n^{-1}\mathbf{Z}_k\in I_\epsilon^{(2)}\right) \leq \sum_{k\in(ns,nt]} \mathbb{P}(|\mathbf{\Psi}_k|>a_n\epsilon) \overset{n\to\infty}{\longrightarrow} 0, \quad (2.4.9)$$

and on the other hand,

$$\lim_{\epsilon\downarrow 0}\lim_{n\to\infty}\mathbb{P}(\widetilde{\kappa}_n(I_\epsilon)>0) = \lim_{\epsilon\downarrow 0}\mathbb{P}(\kappa(I_\epsilon)>0)=0. \quad (2.4.10)$$

Thus, by (2.4.7)-(2.4.10) we get $\lim_{n\to\infty}\mathbb{P}(\widetilde{\kappa}_n(I)\neq\zeta_n(I))=0$. Applying Rootzén [131], Lemma 3.3, we end with $\zeta_n \overset{w}{\Longrightarrow} \kappa$ as $n\to\infty$.

The rest of this proof is an analogon to the proof of Lemma 1.2.4 and therefore we only sketch it.

Since I is bounded away from $\mathbf{0}$, there exists an $x>0$ such that

$$\mathbb{P}(a_n^{-1}(\mathbf{Z}_k+\mathbf{\Psi}_k)\in I) \leq \mathbb{P}(|\mathbf{Z}_k+\mathbf{\Psi}_k|>a_nx) \leq \mathbb{P}(W>a_nx). \quad (2.4.11)$$

Furthermore,

$$\frac{\Gamma_1-(\Gamma_{N(t)+1}-\Gamma_{N(t)})}{N(t)} + \frac{\sum_{k=1}^{N(t)}(\Gamma_{k+1}-\Gamma_k)}{N(t)} \leq \frac{t}{N(t)} \leq \frac{\Gamma_1}{N(t)} + \frac{\sum_{k=1}^{N(t)}(\Gamma_{k+1}-\Gamma_k)}{N(t)}.$$

Regarding that $\{\Gamma_{k+1}-\Gamma_k\}_{k\in\mathbb{N}}$ is i. i. d. with mean $1/\mu$ by the LLN we obtain

$$N(t)/t\to\mu \quad \text{for } n\to\infty \text{ a. s..} \quad (2.4.12)$$

For a detailed proof of the LLN for renewal processes we refer to Mikosch [111], Theorem 2.2.4. Using (2.4.11) and (2.4.12) the rest of the proof follows along the lines of Lemma 1.2.4. □

Lemma 2.4.2 is an analog result to Theorem 1.2.1 for subexponential distributions in the maximum domain of attraction of the Gumbel distribution. We will give some examples for random vectors $\{\mathbf{\Psi}_k\}_{k\in\mathbb{N}}$ to satisfy (2.4.6):

Example 2.4.3

(a) Suppose there exists a r. v. ψ such that for some $x_0\in\mathbb{R}$ and every $\epsilon>0$, $k\in\mathbb{N}$,

$$\mathbb{P}(|\mathbf{\Psi}_k|>x)\leq\mathbb{P}(\psi>x) \quad \text{for } x\geq x_0 \quad \text{and} \quad \mathbb{P}(\psi>a_n\epsilon)=o(1/n) \quad \text{for } n\to\infty,$$

then the assumption (2.4.6) is satisfied.

(b) Let $\{\widetilde{Z}_k\}_{k\in\mathbb{Z}}$ be a sequence of identically distributed r. v. s, which are regularly varying of index $-\alpha$ in the sense of (2.1.2) and with the same a_n as Z_1. Suppose $\{\widetilde{Z}_k\}_{k\in\mathbb{Z}}$ are independent of the sequence of random vectors $\{\widetilde{\mathbf{\Psi}}_k\}_{k\in\mathbb{Z}}$ in $\mathbb{R}^d$, which have support on $[-f^+, f^+]^d$. Define $\mathbf{\Psi}_k := \widetilde{\mathbf{\Psi}}_k\widetilde{Z}_k$ and assume there exists a $\delta^* > 0$ with $\alpha - \delta^* > 0$ such that

$$\sum_{k=-\infty}^{\infty} \mathbb{E}|\widetilde{\mathbf{\Psi}}_k|^{\alpha-\delta^*} < \infty. \tag{2.4.13}$$

Denote by F_k the d. f. of $\widetilde{\mathbf{\Psi}}_k$. By Potter's Theorem A.1.10 *(iii)* there exists an $n_0 > 0$, $K > 1$ such that

$$\mathbb{P}(f^+|Z_1| > a_n y)/\mathbb{P}(f^+|Z_1| > a_n) \leq K y^{-\alpha+\delta^*} \quad \text{for } y \geq 1,\ n \geq n_0.$$

Then we have for $n \geq n_0$

$$\begin{aligned}
\mathbb{P}(|\widetilde{\mathbf{\Psi}}_k\widetilde{Z}_k| > a_n\epsilon) &= \int_{\overline{\mathbb{R}}^d\setminus\{\mathbf{0}\}} \mathbb{P}(f^+|\widetilde{Z}_k| > a_n\epsilon f^+/|\mathbf{t}|)\, F_k(d\mathbf{t}) \\
&\leq K\mathbb{P}(f^+|\widetilde{Z}_k| > a_n\epsilon) f^{+\,\delta^*-\alpha} \int_{\overline{\mathbb{R}}^d\setminus\{\mathbf{0}\}} |\mathbf{t}|^{\alpha-\delta^*} F_k(d\mathbf{t}) \\
&= K\mathbb{P}(f^+|\widetilde{Z}_1| > a_n\epsilon) f^{+\,\delta^*-\alpha}\mathbb{E}|\widetilde{\mathbf{\Psi}}_k|^{\alpha-\delta^*}.
\end{aligned} \tag{2.4.14}$$

Regarding (2.4.14) and $a_n \to \infty$ as $n \to \infty$, we obtain

$$\lim_{n\to\infty} \sum_{k=1}^{\lfloor nt\rfloor} \mathbb{P}(|\mathbf{\Psi}_k| > a_n\epsilon) \leq K f^{+\,\delta^*-\alpha} \lim_{n\to\infty} \mathbb{P}(f^+|\widetilde{Z}_1| > a_n\epsilon) \sum_{k=1}^{\infty} \mathbb{E}|\widetilde{\mathbf{\Psi}}_k|^{\alpha-\delta^*} = 0.$$

Thus, also $\{\mathbf{\Psi}_k\}_{k\in\mathbb{Z}}$ satisfies the assumption (2.4.6). □

The last Lemma of this section regards on point processes of independent regularly varying sequences. It is analogously to Proposition 1.2.5 for subexponential sequences in the maximum domain of attraction of the Gumbel distribution.

Lemma 2.4.4

For $i = 1, 2$, let $Z^{(i)} = \{Z_k^{(i)}\}_{k\in\mathbb{N}}$ be sequences of identically distributed r. v. s, which are regularly varying of index $-\alpha$, $\alpha > 0$, in the sense of (2.1.2), such that for a sequence $0 < a_n \uparrow \infty$ of constants and $s = 1, 2$,

$$\lim_{n\to\infty} n\mathbb{P}(|Z_1^{(s)}| > a_n x) = c_s x^{-\alpha} \quad \textit{for } x > 0,$$

where $c_s \geq 0$. Assume $Z^{(1)}$ and $Z^{(2)}$ are independent. For $n \in \mathbb{N}$ denote by

$$\kappa_n^{(i)} = \sum_{k=1}^{\infty} \varepsilon_{\left(k/n, Z_k^{(i)}/a_n\right)} \quad \text{and} \quad \kappa_n = \sum_{k=1}^{\infty} \varepsilon_{\left(k/n, (Z_k^{(1)}+Z_k^{(2)})/a_n\right)}$$

point processes in $M_P(S)$. Suppose $\kappa_n^{(i)} \overset{w}{\Longrightarrow} \kappa^{(i)}$ as $n \to \infty$ for some point process $\kappa^{(i)}$ with $\kappa^{(i)}([s,t) \times \{x\}) = 0$ a.s. for $x \in \overline{\mathbb{R}} \backslash \{0\}$, $t > s \geq 0$, and $\kappa^{(1)}$ is independent of $\kappa^{(2)}$. Let $I = [s,t) \times (c,d] \subseteq S$. Then

$$\lim_{n \to \infty} \mathbb{P}(\kappa_n(I) \neq \kappa_n^{(1)}(I) + \kappa_n^{(2)}(I)) = 0$$

and $\kappa_n \overset{w}{\Longrightarrow} \kappa^{(1)} + \kappa^{(2)}$ as $n \to \infty$.

Proof.
Let $\epsilon > 0$ be fixed and $I_\epsilon = [s,t) \times [c-\epsilon, c+\epsilon) \cup [s,t) \times [d-\epsilon, d+\epsilon)$. Then

$$\begin{aligned}
\{\kappa_n(I) \neq \kappa_n^{(1)}(I) + \kappa_n^{(2)}(I)\} &\subseteq \{\kappa_n^{(1)}(I_\epsilon) > 0\} \cup \{\kappa_n^{(2)}(I_\epsilon) > 0\} \\
\bigcup_{k \in [ns,nt)} &\{a_n^{-1}(Z_k^{(1)} + Z_k^{(2)}) \in I, a_n^{-1} Z_k^{(1)} \in I \backslash I_\epsilon, a_n^{-1} Z_k^{(2)} \in I \backslash I_\epsilon\} \\
&\cup \{a_n^{-1}(Z_k^{(1)} + Z_k^{(2)}) \in I, a_n^{-1} Z_k^{(1)} \in I^c \backslash I_\epsilon, a_n^{-1} Z_k^{(2)} \in I^c \backslash I_\epsilon\} \\
&\cup \{a_n^{-1}(Z_k^{(1)} + Z_k^{(2)}) \in I^c, a_n^{-1} Z_k^{(1)} \in I \backslash I_\epsilon\} \\
&\cup \{a_n^{-1}(Z_k^{(1)} + Z_k^{(2)}) \in I^c, a_n^{-1} Z_k^{(2)} \in I \backslash I_\epsilon\}.
\end{aligned}$$

Thus, by the independence of $Z^{(1)}$ and $Z^{(2)}$ we obtain

$$\begin{aligned}
&\mathbb{P}(\kappa_n(I) \neq \kappa_n^{(1)}(I) + \kappa_n^{(2)}(I)) \\
&\leq \mathbb{P}(\kappa_n^{(1)}(I_\epsilon) > 0) + \mathbb{P}(\kappa_n^{(2)}(I_\epsilon) > 0) \\
&\quad + n(t-s)[\mathbb{P}(a_n^{-1} Z_1^{(1)} \in I \backslash I_\epsilon)\mathbb{P}(a_n^{-1} Z_1^{(2)} \in I \backslash I_\epsilon) + \mathbb{P}(|Z_1^{(1)}| > a_n \epsilon)\mathbb{P}(|Z_1^{(2)}| > a_n \epsilon)] \\
&\quad + n(t-s)[\mathbb{P}(|Z_1^{(1)}| > a_n \epsilon)\mathbb{P}(a_n^{-1} Z_1^{(2)} \in I \backslash I_\epsilon) + \mathbb{P}(|Z_1^{(2)}| > a_n \epsilon)\mathbb{P}(a_n^{-1} Z_1^{(1)} \in I \backslash I_\epsilon)] \\
&\overset{n \to \infty}{\longrightarrow} \mathbb{P}(\kappa^{(1)}(I_\epsilon) > 0) + \mathbb{P}(\kappa^{(2)}(I_\epsilon) > 0) \overset{\epsilon \to 0}{\longrightarrow} 0.
\end{aligned} \tag{2.4.15}$$

Taking the independence of $\kappa_n^{(1)}$, $\kappa_n^{(2)}$ and $\kappa^{(1)}$, $\kappa^{(2)}$ into account, we get also $\kappa_n^{(1)} + \kappa_n^{(2)} \overset{w}{\Longrightarrow} \kappa^{(1)} + \kappa^{(2)}$ for $n \to \infty$. Thus by Rootzén [131], Lemma 3.3, and (2.4.15) we have $\kappa_n \overset{w}{\Longrightarrow} \kappa^{(1)} + \kappa^{(2)}$ for $n \to \infty$. □

2.5 Extremal behavior

In this section we study the extremal behavior of a regularly varying mixed MA process Y given as in (2.0.1). Therefore we use, similarly to subexponential MA processes in the maximum domain of attraction of the Gumbel distribution (Section 1.4), a discrete-time skeleton. This means we investigate the extremal behavior of the discrete-time sequence $\{Y(t_n)\}_{n\in\mathbb{N}}$, where the discrete-time sequence $\{t_n\}_{n\in\mathbb{N}}$ is chosen properly by the extremes of the kernel function and the jump times of the Lévy process. The extremes of $\{Y(t_n)\}_{n\in\mathbb{N}}$ coincide with the extremes of Y. We would like to give a short motivation for this:

Consider the mixed Poisson shot noise process $Y(t) = \sum_{k=-\infty}^{\infty} f(R_k, t-\Gamma_k)Z_k$ for $t \in \mathbb{R}$, then

$$Y(\Gamma_k + t) = f(R_k, t)Z_k + \sum_{\substack{j=-\infty \\ j\neq k}}^{\infty} f(R_j, t+\Gamma_k-\Gamma_j)Z_j \quad \text{for } k \in \mathbb{N},\, t \in \mathbb{R}.$$

In the case that the $\{Z_k\}_{k\in\mathbb{Z}}$ are regularly varying it happens, that one Z_k is really large in comparison to $\{Z_j\}_{j\in\mathbb{Z}\setminus\{k\}}$. Then $Y(\Gamma_k + t)$ behaves like $f(R_k, t)Z_k$. The process $\{f(R_k, t)Z_k\}_{t\geq 0}$ achieves its supremum in $\eta^{(1)}$, where $f(R_k, \eta^{(1)}) = f^+$. Similar results hold for large negative jumps and the infimum of the kernel function $\eta^{(2)}$ with $f(R_k, \eta^{(2)}) = -f^-$. This suggests that $\{Y(t_n)\}_{n\in\mathbb{N}}$ with

$$t_n \in \{\Gamma_k + \eta^{(1)} : k \in \mathbb{N}\} \ \cup \{\Gamma_k + \eta^{(2)} : k \in \mathbb{N}\}$$

describes completely the extremes of Y. Clusters of high level exceedances of Y are caused by large jumps of the Lévy process in combination with extremes of the kernel function.

In general we consider a mixed MA process Y as given in (2.0.1). We come back to the decomposition of the i. d. i. s. r. m. $\Lambda = \Lambda_1 + \Lambda_2$ as given in (2.0.4), where Λ_1 is a compound Poisson random measure with generating quadruple $(0, 0, \nu_1, \pi)$ and Λ_2 is an i. d. i. s. r. m. with generating quadruple $(m, \sigma^2, \nu_2, \pi)$. The support of ν_2 is bounded, so that the jump sizes of the underlying Lévy process are bounded, too. The compound Poisson random measure Λ_1 has representation

$$\Lambda_1(A) = \int_A \int_{\mathbb{R}} x\, d\widetilde{N}_1(\omega, x) \quad \text{and} \quad \widetilde{N}_1 = \sum_{k=-\infty}^{\infty} \varepsilon_{(R_k, \Gamma_k, Z_k)},$$

where $-\infty < \ldots < \Gamma_{-1} < \Gamma_0 \leq 0 < \Gamma_1 < \ldots < \infty$ are the jump times of a Poisson process N with intensity $\mu = \nu_1(\mathbb{R})$. Furthermore, Z is an i. i. d. sequence with d. f. $\mathbb{P}(Z_1 \leq x) = \nu_1(-\infty, x]/\mu$ for $x \in \mathbb{R}$ independent of R an i. i. d. sequence with d. f. π. In this section we consider regularly varying mixed MA processes Y as given in (2.0.1), so that we assume (M1) is satisfied. Hence, by Proposition 1.1.2 *(v)* the sequence $0 < a_n \uparrow \infty$ of constants satisfies

$$n\mathbb{P}(a_n^{-1} Z_1 \in \cdot) \overset{v}{\Longrightarrow} \sigma(\cdot)/\mu \quad \text{on } \mathcal{B}(\overline{\mathbb{R}}\backslash\{0\}) \text{ for } n \to \infty, \tag{2.5.1}$$

where $\sigma(dx) = p\alpha x^{-\alpha-1}\,\mathbf{1}_{(0,\infty)}(x)\,dx + q\alpha(-x)^{-\alpha-1}\,\mathbf{1}_{(-\infty,0)}(x)\,dx,\ p,q \in [0,1]$, $p+q=1$. This decomposition of Λ induces also a decomposition of Y in $Y = Y_1 + Y_2$ as given in (2.0.6), where

$$Y_1(t) = \sum_{k=-\infty}^{\infty} f(R_k, t-\Gamma_k) Z_k \quad \text{for } t \in \mathbb{R} \text{ a. s..} \tag{2.5.2}$$

We show that the mixed MA process Y_2 has no influence on the extremal behavior of Y, since the driving Lévy process has only small jumps. Throughout this section the support of f in the second coordinate is of the form $[a,b]$, $(-\infty,a]$, $[a,\infty)$, $a<b$. Furthermore, f is in the second coordinate continuous in the interior of its support and right respectively left continuous at the boundary.

The second model we pay attention to are stationary renewal shot noise processes Y as given in (2.3.1) with

$$Y(t) = \sum_{j=-\infty}^{\infty} f(t-\Gamma_j) Z_j \quad \text{for } t \in \mathbb{R}, \tag{2.5.3}$$

where $\Gamma = \{\Gamma_k\}_{k\in\mathbb{Z}}$ are the jump times of a stationary renewal process N with intensity μ independent of the regularly varying i. i. d. sequence Z. This class of processes is a generalization of Poisson shot noise processes. Here we assume condition (R1) is satisfied. Hence, the sequence $0 < a_n \uparrow \infty$ of constants satisfies

$$n\mathbb{P}(a_n^{-1} Z_1 \in \cdot) \overset{v}{\Longrightarrow} \sigma(\cdot)/\mu \quad \text{on } \mathcal{B}(\overline{\mathbb{R}}\backslash\{0\}) \text{ for } n \to \infty.$$

The extremal behavior for the stationary mixed MA process and the stationary renewal shot noise process are the same, by the identical structure of the Poisson shot noise process Y_1 given by (2.5.2) and the stationary renewal shot noise process given by (2.5.3). Similar results hold for the discrete-time MA processes, which have been investigated by Davis and Resnick [54] and Rootzén [130]. The extremal behavior of heavy tailed shot noise processes, where $f : [0,1] \to [0,1]$ is strictly decreasing and the jump sizes Z are positive, has been thoroughly investigated by McCormick [108,107]. We generalize his result such that f is no longer bounded and positive.

In this section we need the definition of the renewal process $\widetilde{N}$ with jump times $T = \{T_k\}_{k\in\mathbb{Z}}$ where

$$T_k := \Gamma_{k+1} - \Gamma_1 \quad \text{and} \quad T_{-k} := \Gamma_{-k} - \Gamma_0 \quad \text{for } k \in \mathbb{N}_0. \tag{2.5.4}$$

2.5.1 The point process of local maxima

Let $\mathbf{Y}$ be given as in (2.0.14), where Y is either a stationary mixed MA process or a stationary renewal shot noise process and define for $k \in \mathbb{Z}$ the disjoint intervals

$$I_k = \left[\eta^{(1)} + (\Gamma_{k-1} + \Gamma_k)/2, \eta^{(1)} + (\Gamma_k + \Gamma_{k+1})/2\right). \tag{2.5.5}$$

The extremal behavior of the mixed MA process Y as given in (2.0.1) is described by the multivariate point process

$$\kappa_n = \sum_{k=1}^{\infty} \varepsilon_{\left(\Gamma_k/n, \sup_{h\in I_k} Y(h)/a_n, \mathbf{Y}(\Gamma_k)/a_n\right)} \quad \text{in } [0,\infty) \times [0,\infty]^{d+3} \backslash \{\mathbf{0}\}. \tag{2.5.6}$$

This point process can be interpreted as a *marked point process* (Daley and Vere-Jones [50], Section 6.4). Let

$$\widetilde{\mathbf{Y}}_k = (\sup_{h\in I_k} Y(h), \mathbf{Y}(\Gamma_k)), \tag{2.5.7}$$

where $\widetilde{Y}_{k,h}$, $h \in \{1,\ldots,d{+}3\}$ is the h^{th} coordinate of $\widetilde{\mathbf{Y}}_k$. Marked point process means that we consider the point process behavior of $\sum_{k=1}^{\infty} \varepsilon_{\left(\Gamma_k/n, \widetilde{Y}_{k,h}/a_n\right)}$ for some fixed h, and the remaining coordinates of $\widetilde{\mathbf{Y}}_k$ in κ_n describe the behavior of the process, when an excess of $Y_{k,h}$ over a high threshold occurs. In our setting $\mathbf{Y}(\Gamma_k)/a_n$ are the marks, which describe the sample path behavior of the continuous-time process Y, if an extreme occur. They characterize clearly the location of extremes. By adding this mark we obtain complete information about the extremal behavior of Y.

Theorem 2.5.1
Let Y be either a stationary mixed MA process as given in (2.0.1) satisfying (M2) or a stationary renewal shot noise process as given in (2.3.1) satisfying (R2) with $pf^+ > 0$. If $qf^- > 0$ assume furthermore Z_k has the decomposition $Z_k = Z_k^{(1)} - Z_k^{(2)}$, where $\{Z_k^{(1)}\}_{k\in\mathbb{Z}}$, $\{Z_k^{(2)}\}_{k\in\mathbb{Z}}$ are independent sequences and for $s = 1,2$, $\{Z_k^{(s)}\}_{k\in\mathbb{Z}}$ is a sequence of i. i. d. positive r. v. s with

$$\mathbb{P}(Z_1^{(1)} > x) \sim \mathbb{P}(Z_1 > x), \quad \mathbb{P}(Z_1^{(2)} > x) \sim \mathbb{P}(Z_1 < -x) \quad \textit{for } x \to \infty. \tag{2.5.8}$$

Let $\sum_{k=1}^\infty \varepsilon_{(s_k,P_k)}$ be a PRM(ϑ) *with $\vartheta(dt\times dx) = dt\times \alpha x^{-\alpha-1}\mathbf{1}_{(0,\infty)}(x)\,dx$. The sequences $T^{(k)} = \{T_{k,j}\}_{j\in\mathbb{Z}}$ are i. i. d. with $T^{(k)} \stackrel{d}{=} T$, where T is given by (2.5.4), independent of $R = \{R_k\}_{k\in\mathbb{N}}$ an i. i. d. sequence with d. f. π. Let $\chi = \{\chi_k\}_{k\in\mathbb{N}}$ be an i. i. d. sequence with d. f. $\mathbb{P}(\chi_k = 1) = p$ and $\mathbb{P}(\chi_k = -1) = q$. The random elements $\sum_{k=1}^\infty \varepsilon_{(s_k,P_k)}, \{T^{(k)}\}_{k\in\mathbb{N}}, \chi, R$ are independent. Furthermore, define for $k\in\mathbb{N}$, $j\in\mathbb{Z}$,*

$$I_j^{(k)} = \left[\eta^{(1)} + (T_{k,j-1}+T_{k,j})/2, \eta^{(1)} + (T_{k,j}+T_{k,j+1})/2\right).$$

Then for κ_n as given in (2.5.6) and $n\to\infty$,

$$\kappa_n \stackrel{w}{\Longrightarrow} \sum_{k=1}^\infty \sum_{j=-\infty}^\infty \varepsilon_{\left(s_k, \sup_{h\in I_j^{(k)}}\{f(R_k,h)\chi_k P_k\}, \mathbf{f}(R_k,T_{k,j})\chi_k P_k\right)} =: \kappa$$

in $M_P([0,\infty)\times[0,\infty]^{d+3}\backslash\{\mathbf{0}\})$.

Proof.
We define for $k\in\mathbb{Z}$

$$\widetilde{I}_k = \left[\eta^{(1)} + (T_{k-1}+T_k)/2, \eta^{(1)} + (T_k+T_{k+1})/2\right)$$

with T_k given by (2.5.4) and the random functions

$$\begin{aligned} \widetilde{\mathbf{f}}_k^{(1)}(r,t) &:= (\sup\nolimits_{h\in\widetilde{I}_k} f(r,h+t), \mathbf{f}(r,T_k+t)) && \text{for } k\in\mathbb{Z}, r\in\mathbb{R}_+, t\in\mathbb{R},\\ \widetilde{\mathbf{f}}_k^{(2)}(r,t) &:= (\sup\nolimits_{h\in\widetilde{I}_k} -f(r,h+t), -\mathbf{f}(r,T_k+t)) && \text{for } k\in\mathbb{Z}, r\in\mathbb{R}_+, t\in\mathbb{R}, \end{aligned}$$

with values in $\mathbb{R}^{d'}$, where $d' = d+3$ and $\mathbf{f}$ is given by (2.0.13). Note that

$$\widetilde{\mathbf{f}}_k^{(s)}(R_j,-T_j) \stackrel{d}{=} \widetilde{\mathbf{f}}_{k-j}^{(s)}(R_1,0) \tag{2.5.9}$$

for $s=1,2$. We use the decomposition (2.0.6) of $Y = Y_1+Y_2$ and restrict, at first, our attention to the mixed Poisson shot noise process $Y_1(t) = \sum_{k=-\infty}^\infty f(R_k,t-\Gamma_k)Z_k$ for $t\in\mathbb{R}$ a. s. given as in (2.0.7), which is well-defined by (M2), (R2), respectively.

In the case $qf^- = 0$ define $Z_k^{(1)} = Z_k$ and $Z_k^{(2)} = 0$ for $k\in\mathbb{Z}$.

Step 1. For fixed $m>0$ we study the extremal behavior of

$$\widetilde{\mathbf{Y}}_k^{(m)} = \sum_{j=k-m}^{k+m} \widetilde{\mathbf{f}}_k^{(1)}(R_j,-T_j)Z_j^{(1)} + \sum_{j=k-m}^{k+m} \widetilde{\mathbf{f}}_k^{(2)}(R_j,-T_j)Z_j^{(2)} \quad \text{for } k\in\mathbb{Z}.$$

We show that $\{\widetilde{\mathbf{Y}}_k^{(m)}\}_{k\in\mathbb{Z}}$ satisfies the assumptions of Theorem 2.4.1.

Regarding (2.5.1) and (2.5.8) we apply Lemma 2.1.3 and (2.5.9) to end with

$$\begin{aligned}
&\lim_{n\to\infty} n\mathbb{P}(|\widetilde{\mathbf{Y}}_k^m| > a_n x) \\
&\quad = \frac{p}{\mu}\sum_{j=k-m}^{k+m} \mathbb{E}|\widetilde{\mathbf{f}}_k^{(1)}(R_j, -T_j)|^\alpha + \frac{q}{\mu}\sum_{j=k-m}^{k+m} \mathbb{E}|\widetilde{\mathbf{f}}_k^{(2)}(R_j, -T_j)|^\alpha \\
&\quad = \frac{p}{\mu}\sum_{j=-m}^{m} \mathbb{E}|\widetilde{\mathbf{f}}_j^{(1)}(R_1, 0)|^\alpha + \frac{q}{\mu}\sum_{j=-m}^{m} \mathbb{E}|\widetilde{\mathbf{f}}_j^{(2)}(R_1, 0)|^\alpha =: \rho_m. \qquad (2.5.10)
\end{aligned}$$

Observing that $\{\widetilde{\mathbf{Y}}_k^{(m)}\}_{k\in\mathbb{Z}}$ is $(2m+1)$-dependent and taking Lemma 2.4.2 in Leadbetter and Rootzén [97] into account, condition $\mathcal{A}(a_n\rho_m^{1/\alpha})$ holds for $\{\widetilde{\mathbf{Y}}_k^{(m)}\}_{k\in\mathbb{Z}}$. Also by the $(2m+1)$-dependence of $\{\widetilde{\mathbf{Y}}_k^{(m)}\}_{k\in\mathbb{Z}}$, (2.5.10) and $r_n = o(n)$ for $n \to \infty$ of condition $\mathcal{A}(a_n\rho_m^{1/\alpha})$, we obtain for $l > 2m+1$

$$\mathbb{P}\left(\bigvee_{l\le|k|\le r_n} |\widetilde{\mathbf{Y}}_k^{(m)}| > a_n x \,\middle|\, |\widetilde{\mathbf{Y}}_0^{(m)}| > a_n x\right) \le r_n \mathbb{P}(|\widetilde{\mathbf{Y}}_k^{(m)}| > a_n x) \xrightarrow{n\to\infty} 0.$$

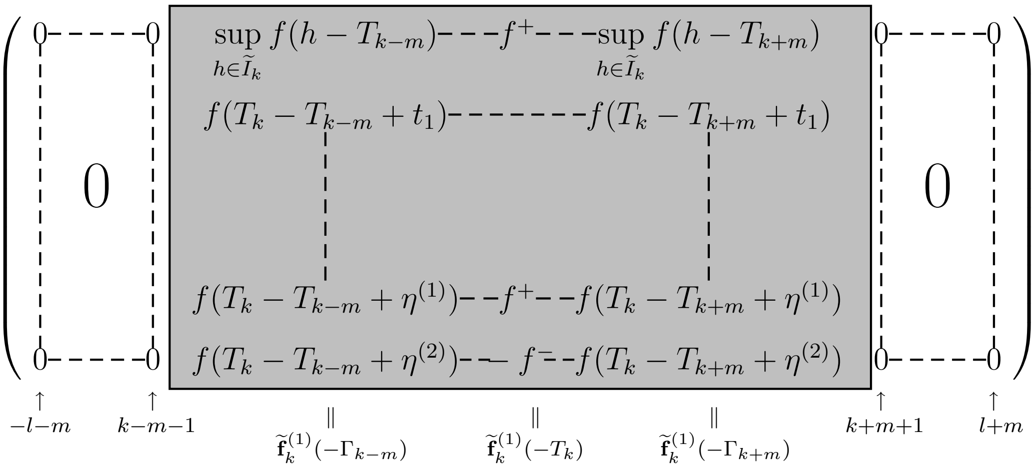
Define for $s = 1, 2$ the random vectors $\mathbf{Z}^{(l,s)} := (Z_{-l-m}^{(s)}, \ldots, Z_{l+m}^{(s)}) \in \mathbb{R}^{2(l+m)+1}$, $l \in \mathbb{N}$, and the random matrices

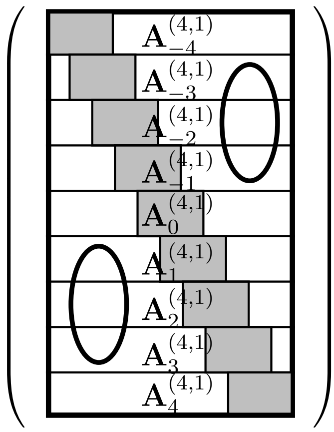
$$\mathbf{A}^{(l,s)} := \begin{pmatrix} \mathbf{A}_{-l}^{(l,s)} \\ \vdots \\ \mathbf{A}_l^{(l,s)} \end{pmatrix} \in \mathbb{R}^{(2l+1)d' \times (2(l+m)+1)},$$

where $\mathbf{A}_k^{(l,s)} \in \mathbb{R}^{d' \times (2(l+m)+1)}$ for $k = -l, \ldots, l$, has entries $(\mathbf{A}_k^{(l,s)})_{i,j}$ in the i^{th} row and j^{th} column with values

$$\begin{aligned}
(\mathbf{A}_k^{(l,1)})_{1,j} &= \sup_{h\in\widetilde{I}_k} f(R_j, h - T_j), \\
(\mathbf{A}_k^{(l,1)})_{i+1,j} &= f(R_j, T_k - T_j + t_i)
\end{aligned}$$

for $j = k-m, \ldots, k+m$, $i = 1, \ldots, d'-1$, $k = -l, \ldots, l$, where $t_{d+1} := \eta^{(1)}$, $t_{d+2} := \eta^{(2)}$. Furthermore, $(\mathbf{A}_k^{(l,1)})_{i,j} = 0$ for $|k-j| > m$, $j = -l-m, \ldots, l+m$, $i = 1, \ldots, d'$ and $k = -l, \ldots, l$. This means $\mathbf{A}_k^{(l,1)}$ has the representation

$\mathbf{A}_k^{(l,1)} =$ 

such that, e. g. for $l = 4$, $m = 1$, $d' = 2$: $\mathbf{A}^{(l,1)} =$ 

,

where the white area marks entries with values 0. The sequence of random matrices $(\mathbf{A}_k^{(l,1)})_{k\in\mathbb{Z}}$ is $(2m+1)$-dependent. Similarly to $\mathbf{A}^{(l,1)}$, we define $\mathbf{A}^{(l,2)}$ by just changing the kernel function from f to $-f$. Then define $\mathbf{A}^{(l)} := (\mathbf{A}^{(l,1)}, \mathbf{A}^{(l,2)}) \in \mathbb{R}^{(2l+1)d' \times 2(2(l+m)+1)}$ and $\mathbf{Z}^{(l)} := (\mathbf{Z}^{(l,1)}, \mathbf{Z}^{(l,2)})^t \in \mathbb{R}^{2(2(l+m)+1)}$. Thus we have

$$\widetilde{\mathbf{Y}}^{(l)} = \mathbf{A}^{(l)}\mathbf{Z}^{(l)} \in \mathbb{R}^{(2l+1)d'}.$$

The matrix $\mathbf{A}^{(l)}$ has at most $2(2m+1)$ entries in a row and $d'(2m+1)$ in a column.

Since $f^+ \le \|\mathbf{A}^{(l)}\| \le 2(2m+1)f^+$ we can apply Lemma 2.1.3 and conclude that $\widetilde{\mathbf{Y}}^{(l)}$ is multivariate regularly varying of index $-\alpha$ with spectral measure

$$\begin{aligned} &\mathbb{P}(\Theta^{(l)} \in \cdot) \qquad (2.5.11)\\ &= \sum_{j=-l-m}^{l+m} \frac{p}{\mu\widetilde{\rho}_m} \mathbb{E}\left(|\mathbf{a}_j^{(l,1)}|^\alpha \mathbf{1}_{\left\{\mathbf{a}_j^{(l,1)}/|\mathbf{a}_j^{(l,1)}| \in \cdot\right\}}\right) + \frac{q}{\mu\widetilde{\rho}_m} \mathbb{E}\left(|\mathbf{a}_j^{(l,2)}|^\alpha \mathbf{1}_{\left\{\mathbf{a}_j^{(l,2)}/|\mathbf{a}_j^{(l,2)}| \in \cdot\right\}}\right), \end{aligned}$$

where $\mathbf{a}_j^{(l,s)} = \mathbf{A}^{(l,s)}\mathbf{e}_j$ with the j^{th} unit vector $\mathbf{e}_j = (0,\dots,0,1,0,\dots,0) \in \mathbb{R}^{2(l+m)+1}$ and

$$\widetilde{\rho}_m = \sum_{j=-l-m}^{l+m} \frac{p}{\mu}\mathbb{E}(|\mathbf{a}_j^{(l,1)}|^\alpha) + \frac{q}{\mu}\mathbb{E}(|\mathbf{a}_j^{(l,2)}|^\alpha).$$

We consider the case $l > m$. Without loss of generality we assume $Z_k = Z_k^{(1)}$, i. e. $p = 1$ and $q = 0$, since the results (2.5.12)-(2.5.15) remain the same by the structure of $\mathbf{A}^{(l)}$. We write $\mathbf{A}^{(l)} = \mathbf{A}^{(l,1)}$, $\widetilde{\mathbf{f}}_k = \widetilde{\mathbf{f}}_k^{(1)}$, $\mathbf{a}_j^{(l)} = \mathbf{a}_j^{(l,1)}$ for $k \in \mathbb{Z}$, $j = -l-m,\dots,l+m$. We obtain for $j = -m,\dots,m$,

$$\begin{aligned}
\bigvee_{k=0}^{l}\bigvee_{i=1}^{d'} |(\mathbf{A}_k^{(l)})_{i,j}|^\alpha - \bigvee_{k=1}^{l}\bigvee_{i=1}^{d'} |(\mathbf{A}_k^{(l)})_{i,j}|^\alpha &= \bigvee_{k=0}^{j+m} |\widetilde{\mathbf{f}}_k(-T_j)|^\alpha - \bigvee_{k=1}^{j+m} |\widetilde{\mathbf{f}}_k(-T_j)|^\alpha,\\
\bigvee_{k=0}^{l}\bigvee_{i=1}^{d'} |(\mathbf{A}_k^{(l)})_{i,j}|^\alpha - \bigvee_{k=1}^{l}\bigvee_{i=1}^{d'} |(\mathbf{A}_k^{(l)})_{i,j}|^\alpha &= 0 \quad \text{for } m < |j| \le l+m,
\end{aligned} \tag{2.5.12}$$

where we used $\bigvee_{k=0}^{l}\bigvee_{i=1}^{d'} |(\mathbf{A}_k^{(l)})_{i,j}|^\alpha = \bigvee_{k=1}^{l}\bigvee_{i=1}^{d'} |(\mathbf{A}_k^{(l)})_{i,j}|^\alpha$ for $j > m$ and $\bigvee_{k=0}^{l}\bigvee_{i=1}^{d'} |(\mathbf{A}_k^{(l)})_{i,j}|^\alpha = 0$ for $j < -m$. Furthermore for $j = -m,\dots,m$,

$$\bigvee_{k=-l}^{l}\bigvee_{i=1}^{d'} |(\mathbf{A}_k^{(l)})_{i,j}| = |\mathbf{a}_j^{(l)}| = \bigvee_{k=j-m}^{j+m} |\widetilde{\mathbf{f}}_k(-T_j)| = f^+. \tag{2.5.13}$$

By taking the conditional probability under Γ_k, R_k, $k = -l-m,\dots,l+m$ and Remark 2.1.4, we can calculate with deterministic quantities and invoke. We apply (2.5.11) to compute

$$\begin{aligned}
&\mathbb{E}\left(\bigvee_{k=0}^{l} |\Theta_k^{(l)}|^\alpha - \bigvee_{k=1}^{l} |\Theta_k^{(l)}|^\alpha\right)\\
&\quad = \frac{1}{\mu\widetilde{\rho}_m} \sum_{j=-l-m}^{l+m} \mathbb{E}\left(|\mathbf{a}_j^{(l)}|^\alpha \left[\bigvee_{k=0}^{l}\bigvee_{i=1}^{d'} \frac{|(\mathbf{A}_k^{(l)})_{i,j}|^\alpha}{|\mathbf{a}_j^{(l)}|^\alpha} - \bigvee_{k=1}^{l}\bigvee_{i=1}^{d'} \frac{|(\mathbf{A}_k^{(l)})_{i,j}|^\alpha}{|\mathbf{a}_j^{(l)}|^\alpha}\right]\right).
\end{aligned}$$

Taking (2.5.9) and (2.5.12) into account we receive that the r. h. s. is equal to

$$\begin{aligned}
&\frac{1}{\mu\widetilde{\rho}_m}\left[\sum_{j=-m}^{m} \mathbb{E}\left(\bigvee_{k=0}^{j+m} |\widetilde{\mathbf{f}}_k(R_j,-T_j)|^\alpha\right) - \sum_{j=-m+1}^{m} \mathbb{E}\left(\bigvee_{k=1}^{j+m} |\widetilde{\mathbf{f}}_k(R_j,-T_j)|^\alpha\right)\right]\\
&= \frac{1}{\mu\widetilde{\rho}_m}\left[\sum_{j=-m}^{m} \mathbb{E}\left(\bigvee_{k=-j}^{m} |\widetilde{\mathbf{f}}_k(R_1,0)|^\alpha\right) - \sum_{j=-m+1}^{m} \mathbb{E}\left(\bigvee_{k=-j+1}^{m} |\widetilde{\mathbf{f}}_k(R_1,0)|^\alpha\right)\right]\\
&= \frac{1}{\mu\widetilde{\rho}_m}\mathbb{E}\left(\bigvee_{k=-m}^{m} |\widetilde{\mathbf{f}}_k(R_1,0)|^\alpha\right) = \frac{f^{+\,\alpha}}{\mu\widetilde{\rho}_m}.
\end{aligned} \tag{2.5.14}$$

Similarly, as $\bigvee_{i=1}^{d'} |(\mathbf{A}_0^{(l)})_{i,j}|^\alpha = 0$ for $|j| > m$, we have by (2.5.10)

$$\begin{aligned}\mathbb{E}|\Theta_0^{(l)}|^\alpha &= \frac{1}{\mu\widetilde{\rho}_m}\mathbb{E}\left(\sum_{j=-l-m}^{l+m}|\mathbf{a}_j^{(l)}|^\alpha \bigvee_{i=1}^{d'}\frac{|(\mathbf{A}_0^{(l)})_{i,j}|^\alpha}{|\mathbf{a}_j^{(l)}|^\alpha}\right) = \frac{1}{\mu\widetilde{\rho}_m}\mathbb{E}\left(\sum_{j=-m}^{m}|\widetilde{\mathbf{f}}_j(R_1,0)|^\alpha\right)\\ &= \frac{1}{\mu\widetilde{\rho}_m}\mu\rho_m = \frac{\rho_m}{\widetilde{\rho}_m}. \qquad (2.5.15)\end{aligned}$$

By (2.5.14) and (2.5.15) we obtain for the extremal index in (2.4.4)

$$\theta_m := f^{+\alpha}/(\mu\rho_m). \qquad (2.5.16)$$

In the last step we have to distinguish between the $Z^{(1)}$ and $Z^{(2)}$. Following the proof of (2.5.14) and taking $l > m$ and (2.5.13) into account we get for $s = 1, 2$, $j = -m, \ldots, m$,

$$\begin{aligned}&\mathbb{E}\left(|\mathbf{a}_j^{(l,s)}|^\alpha\left[\bigvee_{k=0}^{l}\bigvee_{i=1}^{d'}\frac{|(\mathbf{A}_k^{(l,s)})_{i,j}|^\alpha}{|\mathbf{a}_j^{(l,s)}|^\alpha}\mathbf{1}\left\{\sum_{|k|\le l}\varepsilon_{(\mathbf{A}_k^{(l,s)})_{\cdot,j}/|\mathbf{a}_j^{(l,s)}|}\in\cdot\right\}\right]\right)\\ &= \mathbb{E}\left(\bigvee_{k=-j}^{m}|\widetilde{\mathbf{f}}_k^{(s)}(R_1,0)|^\alpha\,\mathbf{1}\left\{\sum_{|k|\le m}\varepsilon_{\widetilde{\mathbf{f}}_k^{(s)}(R_1,0)/f^+}\in\cdot\right\}\right).\end{aligned}$$

Then analog the lines of (2.5.14) we obtain

$$\begin{aligned}&\mathbb{E}\left(\left[|\Theta_0^{(l)}|^\alpha - \bigvee_{j=1}^{l}|\Theta_j^{(l)}|^\alpha\right]^+\mathbf{1}\left\{\sum_{|j|\le l}\varepsilon_{\Theta_j^{(l)}}\in\cdot\right\}\right)\\ &= \frac{f^{+\alpha}}{\widetilde{\rho}_m}\left[\frac{p}{\mu}\mathbb{E}\left(\mathbf{1}\left\{\sum_{|j|\le m}\varepsilon_{\widetilde{\mathbf{f}}_j^{(1)}(R_1,0)/f^+}\in\cdot\right\}\right) + \frac{q}{\mu}\mathbb{E}\left(\mathbf{1}\left\{\sum_{|j|\le m}\varepsilon_{\widetilde{\mathbf{f}}_j^{(2)}(R_1,0)/f^+}\in\cdot\right\}\right)\right]\\ &= \frac{f^{+\alpha}}{\mu\widetilde{\rho}_m}\mathbb{E}\left(\mathbf{1}\left\{\sum_{|j|<m}\varepsilon_{(\sup_{h\in\widetilde{I}_j}\{f(R_1,h)\chi_1\},\mathbf{f}(R_1,T_j)\chi_1)/f^+}\in\cdot\right\}\right). \qquad (2.5.17)\end{aligned}$$

Hence by (2.5.14) and (2.5.17) the measure $\mathcal{Q}$ of (2.4.5) is

$$\mathcal{Q}(\cdot) = \mathbb{P}\left(\sum_{j=-m}^{m}\varepsilon_{(\sup_{h\in\widetilde{I}_j}\{f(R_1,h)\chi_1\}/f^+,\mathbf{f}(R_1,T_j)\chi_1/f^+)}\in\cdot\right).$$

Regarding (2.5.16) we choose in Theorem 2.4.1,

$$\widetilde{\vartheta}(dx) := \theta_m\alpha x^{-\alpha-1}\,\mathbf{1}_{(0,\infty)}(x)\,dx = \alpha f^{+\alpha}/(\mu\rho_m)x^{-\alpha-1}\,\mathbf{1}_{(0,\infty)}(x)\,dx.$$

Taking (2.5.10) into account, we apply Theorem 2.4.1 and obtain for $n \to \infty$

$$\sum_{k=1}^{n} \varepsilon_{\left(\rho_m^{1/\alpha}\widetilde{\mathbf{Y}}_k^{(m)}/a_n\right)} \overset{w}{\Longrightarrow} \sum_{k=1}^{\infty}\sum_{j=-m}^{m} \varepsilon_{((\sup_{h\in \widetilde{I}_j^{(k)}}\{f(R_k,h)\chi_k\},\mathbf{f}(R_k,T_{k,j})\chi_k)\widetilde{P}_k/f^+)},$$

where $\sum_{k=1}^{\infty}\varepsilon_{\widetilde{P}_k}$ is PRM$(\widetilde{\vartheta})$ in $M_P(\overline{\mathbb{R}}^{d'}\backslash\{\mathbf{0}\})$. By Hsing [77], Lemma 4.1.2, the convergence of the sequence of point processes $\kappa_n((0,1]\times\cdot)$ is equivalent to the convergence of κ_n, if the so called $\Delta(a_n)$ condition is satisfied, which is similar to condition $\mathcal{A}(a_n)$. Note, that by the $(2m+1)$-dependence of $\{\widetilde{\mathbf{Y}}_k^{(m)}\}_{k\in\mathbb{Z}}$ the $\Delta(a_n)$ condition holds. This implies, changing $\widetilde{\vartheta}$ by ϑ, and $\{\widetilde{P}_k\}_{k\in\mathbb{N}}$ by $\{P_k\}_{k\in\mathbb{N}}$, respectively, that

$$\sum_{k=1}^{\infty} \varepsilon_{\left(k/(n\mu),\widetilde{\mathbf{Y}}_k^{(m)}/a_n\right)} \overset{w}{\Longrightarrow} \sum_{k=1}^{\infty}\sum_{j=-m}^{m} \varepsilon_{(s_k,(\sup_{h\in \widetilde{I}_j^{(k)}}\{f(R_k,h)\chi_k\},\mathbf{f}(R_k,T_{k,j})\chi_k)P_k)}. \tag{2.5.18}$$

Step 2. For fixed $m > 0$ we study the extremal behavior of

$$\mathbf{Y}_k^{(m)} = \sum_{j=k-m}^{k+m} \mathbf{f}_k^{(1)}(R_j,-\Gamma_j)Z_j^{(1)} + \sum_{j=k-m}^{k+m} \mathbf{f}_k^{(2)}(R_j,-\Gamma_j)Z_j^{(2)} \quad \text{for } k\in\mathbb{Z},$$

where

$$\begin{aligned}
\mathbf{f}_k^{(1)}(r,t) &:= (\sup\nolimits_{h\in I_k} f(r,h+t), \mathbf{f}(r,\Gamma_k+t)) && \text{for } k\in\mathbb{Z},\ r\in\mathbb{R}_+,\ t\in\mathbb{R},\\
\mathbf{f}_k^{(2)}(r,t) &:= (\sup\nolimits_{h\in I_k} -f(r,h+t), -\mathbf{f}(r,\Gamma_k+t)) && \text{for } k\in\mathbb{Z},\ r\in\mathbb{R}_+,\ t\in\mathbb{R}.
\end{aligned}$$

Note that $\widetilde{\mathbf{f}}_k^{(s)}(R_j,-T_j) = \mathbf{f}_{k+1}^{(s)}(R_j,-\Gamma_{j+1})$ for $k,j\in\mathbb{N}_0$ by (2.5.4). Then also $\{\widetilde{\mathbf{Y}}_k^{(m)}\}_{k\geq m} \overset{d}{=} \{\mathbf{Y}_{k+1}^{(m)}\}_{k\geq m}$, although $\{\mathbf{Y}_{k+1}^{(m)}\}_{k\in\mathbb{Z}}$ is not stationary. Thus, the asymptotic point process behavior of $\{\widetilde{\mathbf{Y}}_k^{(m)}\}_{k\in\mathbb{N}}$ and $\{\mathbf{Y}_k^{(m)}\}_{k\in\mathbb{N}}$ are the same. Regarding (2.5.18) we obtain

$$\sum_{k=1}^{\infty} \varepsilon_{\left(k/(n\mu),\mathbf{Y}_k^{(m)}/a_n\right)} \overset{w}{\Longrightarrow} \sum_{k=1}^{\infty}\sum_{j=-m}^{m} \varepsilon_{(s_k,\sup_{h\in I_j^{(k)}}\{f(R_k,h)\chi_k\}P_k,\mathbf{f}(R_k,T_{k,j})\chi_k P_k)}. \tag{2.5.19}$$

Step 3. We study the extremal behavior of $\{\overline{\mathbf{Y}}_k\}_{k\in\mathbb{Z}}$ given by

$$\overline{\mathbf{Y}}_k = \sum_{j=-\infty}^{\infty} \mathbf{f}_k^{(1)}(R_j,-\Gamma_j)Z_j^{(1)} + \sum_{j=-\infty}^{\infty} \mathbf{f}_k^{(2)}(R_j,-\Gamma_j)Z_j^{(2)},$$

which are well-defined by (M2), (R2), respectively. We have to attend to the non-stationarity of the sequences $\{\overline{\mathbf{Y}}_k\}_{k\in\mathbb{Z}}$, $\{\mathbf{Y}_k^{(m)}\}_{k\in\mathbb{Z}}$ and $\{\overline{\mathbf{Y}}_k - \mathbf{Y}_k^{(m)}\}_{k\in\mathbb{Z}}$ (if we replace Γ by T then we obtain stationary sequences, cf. Lemma 3.2.11).

With g as in (2.0.12), (2.3.20), respectively, and (2.5.4) we have

$$\begin{aligned}\mathbb{P}\left(\bigvee_{k=m+1}^{n}|\overline{\mathbf{Y}}_k-\mathbf{Y}_k^{(m)}|>a_nx\right)&\leq\sum_{i=1}^{d'}\sum_{k=1}^{n}\mathbb{P}\left(\sum_{|k-j|>m}g(R_j,\Gamma_k-\Gamma_j+t_i)|Z_j|>a_nx\right)\\&=\sum_{i=1}^{d'}\sum_{k=1}^{n}\mathbb{P}\left(\sum_{j=-\infty}^{-m-1}g(R_j,T_j+t_i)|Z_j|+\sum_{j=m+1}^{k-1}g(R_j,T_j+t_i)|Z_j|\right.\\&\left.\quad+\sum_{j=k-1}^{\infty}g(R_j,T_j+\Gamma_1-\Gamma_0+t_i)|Z_{j+1}|>a_nx\right)\end{aligned}\tag{2.5.20}$$

with $t_{d'}=\eta^{(1)}$. Define

$$c_{k,i}:=\sum_{j=-\infty}^{-m-1}\mathbb{E}|g(R_j,T_j+t_i)|^\delta+\sum_{j=m+1}^{k-1}\mathbb{E}|g(R_j,T_j+t_i)|^\delta+\sum_{j=k-1}^{\infty}\mathbb{E}|g(R_j,T_j+\Gamma_1-\Gamma_0+t_i)|^\delta.$$

Then $c_{k,i}\to\sum_{|j|>m}\mathbb{E}|g(R_j,T_j+t_i)|^\delta$ for $k\to\infty$. If Γ is the stationary renewal process we apply in the next inequality (2.3.6) and, else, if Γ are the jump times of the Poisson process we apply Lemma A.4.1. Then there exists an $n_0\in\mathbb{N},K>1$ such that for $n\geq n_0$

$$\begin{aligned}&\mathbb{P}\left(\bigvee_{k=1}^{n}|\overline{\mathbf{Y}}_k-\mathbf{Y}_k^{(m)}|>a_nx\right)\\&\leq Kn\mathbb{P}(|Z_1|>a_nx)\sum_{i=1}^{d'}\left[\frac{1}{n}\sum_{k=1}^{n}c_{k,i}\right]+\sum_{k=1}^{m}\mathbb{P}\left(|\overline{\mathbf{Y}}_k-\mathbf{Y}_k^{(m)}|>a_nx\right)\\&\overset{n\to\infty}{\longrightarrow}Kx^{-\alpha}\sum_{i=1}^{d'}\sum_{|j|>m}\mathbb{E}|g(R_j,T_j+t_i)|^\delta\overset{m\to\infty}{\longrightarrow}0.\end{aligned}$$

Then following the proof of Resnick [125], Proposition 4.2.7, along the lines gives for $n\to\infty$,

$$\sum_{k=1}^{\infty}\varepsilon_{\left(k/(n\mu),\overline{\mathbf{Y}}_k/a_n\right)}\overset{w}{\Longrightarrow}\kappa.\tag{2.5.21}$$

Step 4. The point process behavior of κ_n.

Let $\mathbf{Y}_k=\sum_{j=-\infty}^{\infty}\mathbf{f}(R_j,\Gamma_k-\Gamma_j)Z_j$ for $k\in\mathbb{Z}$ and $I=[s,t)\times\prod_{i=1}^{d'}(c_i,d_i]\subseteq S$. By Lemma A.4.2 we have, similarly to the proof of Lemma 2.4.4,

$$\lim_{n\to\infty}\mathbb{P}\left(\sum_{k=1}^{\infty}\varepsilon_{(k/(n\mu),\overline{\mathbf{Y}}_k/a_n)}(I)\neq\sum_{k=1}^{\infty}\varepsilon_{(k/(n\mu),(\sup_{h\in I_k}Y(h),\mathbf{Y}_k)/a_n)}(I)\right)=0.$$

Hence, by (2.5.21) and Rootzén [131], Lemma 3.3, for $n \to \infty$,

$$\sum_{k=1}^{\infty} \varepsilon_{(k/(n\mu),(\sup_{h\in I_k} Y(h),\mathbf{Y}_k)/a_n)} \stackrel{w}{\Longrightarrow} \kappa. \tag{2.5.22}$$

In the case of a Lévy driven mixed MA process we need to invoke the decomposition (2.0.6) in $\mathbf{Y}(t) = \mathbf{Y}_1(t) + \mathbf{Y}_2(t)$ for $t \in \mathbb{R}$. Then

$$\mathbf{Y}(\Gamma_k) = \mathbf{Y}_k + \mathbf{Y}_2(\Gamma_k) \quad \text{for } k \in \mathbb{Z}. \tag{2.5.23}$$

Recall that $\widetilde{\mathbf{Y}}_k = (\sup_{h\in I_k} Y(h), \mathbf{Y}(\Gamma_k))$. Similarly as (2.5.20) we have for $0 < \epsilon < x$,

$$\begin{aligned}\mathbb{P}(|\widetilde{\mathbf{Y}}_k| > a_n x) \;\leq\; & \sum_{i=1}^{d'} \left[\mathbb{P}\left(\sum_{j=-\infty}^{\infty} g(R_j, \Gamma_k - \Gamma_j + t_i)|Z_j| > a_n(x-\epsilon)\right) \right. \\ & \left. + \mathbb{P}(\sup_{h\in I_k} |\mathbf{Y}_2(h)| > a_n\epsilon)\right].\end{aligned} \tag{2.5.24}$$

The Lévy measure of $Y_2(t)$ has bounded support by Proposition 2.2.1. Using (1.3.15), (1.3.17) and the independence of I_k and Y_2 yields to

$$\begin{aligned}\mathbb{P}(\sup_{h\in I_k} |\mathbf{Y}_2(h)| > a_n\epsilon) \;&\leq\; d'(1/\mu+1)\mathbb{P}(\sup_{0\leq h\leq 1} |Y_2(t)| > a_n\epsilon) \\ &\leq\; d'(1/\mu+1)\mathrm{e}^{-a_n x}\mathbb{E}\exp(\sup_{0\leq h\leq 1} |Y_2(t)|) = o(1/n)\end{aligned} \tag{2.5.25}$$

for $n \to \infty$. Regarding Lemma 2.2.8, (2.3.6), respectively, (2.5.24) and (2.5.25) we obtain that there exists a r.v. W such that

$$\mathbb{P}(|\widetilde{\mathbf{Y}}_k| > a_n x) \leq \mathbb{P}(W > a_n x) = O(1/n) \quad \text{for } n \to \infty.$$

Thus, by Lemma 2.4.2 and Example 2.4.3 *(a)* the point process behavior of the sequence $\{\widetilde{\mathbf{Y}}_k\}_{k\in\mathbb{Z}}$ is the same as that of $\{(\sup_{h\in I_k} Y(h), \mathbf{Y}_k)\}_{k\in\mathbb{Z}}$. Furthermore we can shift the time scale. This together with (2.5.22) completes the proof. □

Notice that condition (M2) excludes MA processes with the long memory property.

This result shows by considering $\sup_{h\in I_k} Y(h)$ with marks $Y(\Gamma_k + \eta^{(1)})$, $Y(\Gamma_k + \eta^{(2)})$ that clusters of long high level exceedances of the continuous-time process only occur, if an exceedance of the Lévy process meets an extreme value of the kernel function. The properly chosen discrete-time points, where exceedances of the Lévy process occur in combination with extremes of the kernel function, result in exceedances of the mixed MA process. These exceedances are carried on in time by the kernel function and result in the limiting process in clusters of exceedances. Extremes on high levels only appear at this properly chosen discrete-time skeleton. See also Section 3.2.1 for the comparison of subexponential Lévy driven OU processes in the domain of attraction of the Gumbel distribution and Fréchet distribution.

Remark 2.5.2
We conjecture that we can transform our results of Theorem 2.5.1 to an infinite-dimensional setting, where we use as marks the stochastic processes $\{Y(\Gamma_k+t)\}_{t\in[0,1]}$ in $\mathbb{D}[0,1]$ instead of multi-dimensional random vectors $\mathbf{Y}(\Gamma_k)\in\mathbb{R}^{d+2}$ for $k\in\mathbb{N}$. By choosing this marks we would obtain complete information about the sample path behavior of the continuous-time process near extremes. We will give the key aspects to which attention should be paid by this consideration.

(i) Investigating the stochastic processes $\{Y(\Gamma_k+t)\}_{t\in[0,1]}$ with a. s. sample paths in $\mathbb{D}(\mathbb{R})$ for $k\in\mathbb{N}$ as marks, requires a definition of regular variation for stochastic processes in $\mathbb{D}_d[0,1]$, the space of functions $h:[0,1]\to\mathbb{R}^d$, which are right continuous with left hand limits. This definition was introduced by Hult and Lindskøg [79] and is an analog to the Definition 2.1.1 of multivariate regular variation. Define the space $\overline{\mathbb{D}}_d[0,1]=(0,\infty]\times\mathbb{D}_d[0,1]$. A nonzero function $h\in\mathbb{D}_d[0,1]$ is associated with the element $(\sup_{t\in[0,1]}|h(t)|, h/\sup_{t\in[0,1]}|h(t)|)\in\overline{\mathbb{D}}_d[0,1]$. A stochastic process $\mathbf{X}=\{\mathbf{X}_t\}_{t\in[0,1]}$ with sample path in $\mathbb{D}_d[0,1]$ is said to be regularly varying, if there exist a sequence $0<a_n\uparrow\infty$ of constants, and a nonzero boundedly finite measure σ on $\mathcal{B}(\overline{\mathbb{D}}_d[0,1])$ with $\sigma(\overline{\mathbb{D}}_d[0,1]\backslash\mathbb{D}_d[0,1])=0$ such that as $n\to\infty$

$$n\mathbb{P}(a_n^{-1}\mathbf{X}\in\cdot)\overset{\hat{\omega}}{\Longrightarrow}\sigma(\cdot)\quad\text{ on }\mathcal{B}(\overline{\mathbb{D}}_d[0,1]),$$

where $\hat{\omega}$-convergence is the convergence on bounded Borel sets; see Daley and Vere-Jones [50], Section A.2.6 for the definition of $\hat{\omega}$-convergence.

(ii) Note, that the space $\mathbb{D}_d[0,1]$ provided with the Skorohod topology is separable and complete (see Billingsley [28], Section 16). But $\mathbb{D}_d[0,1]$ is not locally compact such that the classical results of vague convergence of point processes, which is in $M_P(\overline{\mathbb{R}}^d\backslash\{\mathbf{0}\})$ equivalent to weak convergence, in the space $M_P(\overline{\mathbb{R}}\backslash\{0\}\times\overline{\mathbb{D}}_d[0,1])$ do not apply. Instead of vague convergence of point processes we will use $\hat{\omega}$-convergence in $M_P(\overline{\mathbb{R}}\backslash\{0\}\times\overline{\mathbb{D}}_d[0,1])$.

We sketch the main steps of the proof in the following:

(a) As mentioned in Remark 2.9 of Davis and Mikosch [53], Theorem 2.4.1 holds also for stationary sequences in a separable locally compact Banach space. We have

to proof that Theorem 2.4.1 also holds for point processes in $M_P(\overline{\mathbb{R}}\backslash\{0\} \times \overline{\mathbb{D}}_d[0,1])$ with the $\hat{\omega}$-convergence.

(b) We need a similar result as Lemma 2.1.2 for regularly varying functions in $\mathbb{D}_d[0,1]$. Therefore define $Y_k(t) := \sum_{j=1}^r f_{k,j}(t)Z_j$ for $t \in [0,1]$, with $f_{k,j} \in \mathbb{D}[0,1]$ bounded for $j = 1,\dots,r$, $k = 1,\dots,d$, and $Z = (Z_1,\dots,Z_r)$ given as in Lemma 2.1.2. Let $\mathbf{a}_j = (f_{1,j},\dots,f_{r,j}) \in \mathbb{D}_r[0,1]$. Then the process $Y(t) = (Y_1(t),\dots,Y_d(t))$ for $t \in [0,1]$ in $\mathbb{D}_d[0,1]$ is regularly varying with the corresponding spectral measure as given in (2.1.5).

Regarding *(a)* and *(b)* the proof of Theorem 2.5.1 goes line by line. □

2.5.2 The marked point process at a discrete-time skeleton

In this section we investigate the point process behavior of the continuous-time process $\mathbf{Y}$ as given in (2.0.14) under less restrictive assumptions than in Theorem 2.5.1. Therefore we do not look at local maxima in this section.

Theorem 2.5.3
Let Y be either a stationary mixed MA process as given in (2.0.1) satisfying (M1) or a stationary renewal shot noise process as given in (2.3.1) satisfying (R1). Further assume that the kernel function f satisfies $f(r,\eta^{(1)}) = f^+ \geq f^-$ for all $r \in \operatorname{supp}(\pi)$. Then with the notation of Theorem 2.5.1,

$$\sum_{k=1}^{\infty} \varepsilon_{(\Gamma_k/n,\mathbf{Y}(\Gamma_k)/a_n)} \stackrel{w}{\Longrightarrow} \sum_{k=1}^{\infty}\sum_{j=-\infty}^{\infty} \varepsilon_{(s_k,\mathbf{f}(T_{k,j})\chi_k P_k)} \text{ in } M_P([0,\infty) \times \overline{\mathbb{R}}^{d+2}\backslash\{\mathbf{0}\}).$$

In particular, for every $t \in \mathbb{R}$,

$$\sum_{k=1}^{\infty} \varepsilon_{(\Gamma_k/n,Y(\Gamma_k+t)/a_n)} \stackrel{w}{\Longrightarrow} \sum_{k=1}^{\infty}\sum_{j=-\infty}^{\infty} \varepsilon_{(s_k,f(T_{k,j}+t)\chi_k P_k)} \quad \text{in } M_P([0,\infty) \times \overline{\mathbb{R}}\backslash\{0\}).$$

The proof is analogous to the proof of Theorem 2.5.1. The additional assumption (M2) in Theorem 2.5.1 was only necessary for the investigation of the local maxima and to apply Lemma A.4.2. Furthermore we need not the decomposition of Y_1 in a mixed MA process driven by a positive and a negative Lévy process such that $f \in \mathbb{L}^1$ is not necessary. In contrast to Theorem 2.5.1 we obtain in Theorem 2.5.3

information also about large minima of the continuous-time process Y, since the point process convergence is in $M_P([0,\infty)\times\overline{\mathbb{R}}^{d+2}\backslash\{\mathbf{0}\})$. The interpretation of large minima is analog to large maxima. They occur in clusters and are caused by large jumps of the Lévy process.

2.5.3 Normalizing constants of running maxima

With the results of the previous section we calculate the normalizing constants of running maxima.

Theorem 2.5.4
Let Y either satisfies the assumptions of Theorem 2.5.1 with $f(r,\eta^{(1)}) = f^+$ and $f(r,\eta^{(2)}) = -f^-$ for all $r \in \operatorname{supp}(\pi)$ or Y is a stationary MA process satisfying (M1) with $\int_{-\infty}^{\infty} \sup_{0\le s\le 1} |f(s+t)|^\delta\, dt < \infty$ for some $\delta < \min\{1,\alpha\}$ with $f(\eta^{(1)}) = f^+$ and $f(\eta^{(2)}) = -f^-$. Suppose $pf^+ > 0$. Define $a_T := a_{\lfloor T\rfloor}$ and $M(T) := \sup_{t\in[0,T]} Y(t)$ for $T > 0$. Then for $x > 0$

$$\lim_{T\to\infty} \mathbb{P}\left(a_T^{-1}M(T)\le x\right) = \exp\left(-x^{-\alpha}\left[p{f^+}^\alpha + q{f^-}^\alpha\right]\right).$$

Proof.
In the case of Y a stationary MA process the proof follows with the same methods as those of Theorem 1.4.8, where we calculated the normalizing constants for subexponential Lévy driven MA processes in MDA(Λ). The only difference is that the point process results for regularly varying processes are here applied, i. e. Theorem 2.5.3 and the results for discrete-time MA processes (see Example 2.5.7). We include the proof only for the case, where the assumption of Theorem 2.5.1 are satisfied. Replacing the discrete-time index n by the continuous-time index T in the definition of κ_n in Theorem 2.5.1, then $\kappa_T \stackrel{w}{\Longrightarrow} \kappa$ for $T\to\infty$. Applying Theorem 2.5.1 we obtain for $x > 0$ and $I = [0,1)\times(x,\infty)^{d+3}$

$$\begin{aligned}\lim_{T\to\infty}\mathbb{P}(a_T^{-1}M(T)\le x) &= \lim_{T\to\infty}\mathbb{P}(\kappa_T(I)=0)\\ &= \mathbb{P}(\kappa(I)=0)\\ &= \mathbb{P}\left(\sum_{k=1}^{\infty}\varepsilon_{(s_k,([f^+\mathbf{1}\{\chi_k=1\}+f^-\mathbf{1}\{\chi_k=-1\}]P_k))}([0,1)\times(x,\infty))=0\right)\\ &= \exp\left(-x^{\alpha}\left[p{f^+}^\alpha+q{f^-}^\alpha\right]\right).\end{aligned}$$

□

Notice that this result rules out MA processes, which exhibits long range dependence, only the case of mixed MA processes with long range dependence is included. In this case the long range dependence is caused by the distribution of π instead of the kernel function.

Theorem 2.5.4 holds only for $pf^+ + qf^- > 0$. More about the extremal behavior of totally skewed α-stable MA processes, which satisfy $pf^+ + qf^- = 0$, can be found in Albin [7, 5]. Already Lebedev [98] calculated the limit distribution of running maxima of subexponential positive shot noise processes restricting his attention to non-decreasing kernel functions with unbounded support. In our result the assumption of a positive process with non-increasing kernel function is not necessary.

2.5.4 Extremal index function

In Section 1.4.4 we defined the extremal index function (Definition 1.4.10) as a measure for dependence in extremes. In the following we calculate the extremal index function for regularly varying mixed MA processes.

Corollary 2.5.5 (Extremal index function)
Let Y be a stationary mixed MA process satisfying the assumptions of Theorem 2.5.4. Then Y has extremal index function

$$\theta(h) = \frac{h[pf^{+\,\alpha} + qf^{-\,\alpha}]}{p\int_{\mathbb{R}_+\times\mathbb{R}} \sup\limits_{t\in[0,h]} (f^+(r,t+s))^\alpha\, ds\,\pi(dr) + q\int_{\mathbb{R}_+\times\mathbb{R}} \sup\limits_{t\in[0,h]} (f^-(r,t+s))^\alpha\, ds\,\pi(dr)}.$$

If additionally (M2) holds, then $\lim_{h\to\infty}\theta(h) = 1$.

Proof.
For $h > 0$ define the sequence $M_k^{(h)} := \sup_{(k-1)h\le t\le kh} Y(t)$ for $k \in \mathbb{N}$. On the one hand we have by (2.2.15) for $n \to \infty$

$$n\mathbb{P}(M(h) > a_n x) \sim x^{-\alpha} h[pf^{+\,\alpha} + qf^{-\,\alpha}]/\theta(h). \tag{2.5.26}$$

On the other hand, by Theorem 2.5.4 and the extremal types Theorem

$$\begin{aligned}\lim_{n\to\infty}\mathbb{P}(\max_{k=1,\dots,n} M_k^{(h)} \le a_n x) &= \lim_{n\to\infty}\mathbb{P}(a_n^{-1}M(nh)\le x)\\ &= \lim_{n\to\infty}\mathbb{P}(a_{nh}^{-1}M(nh)\le xh^{-1/\alpha})\\ &= \exp(-x^{-\alpha}h[pf^{+\,\alpha}+qf^{-\,\alpha}]).\end{aligned} \tag{2.5.27}$$

Statements (2.5.26), (2.5.27) and Definition 1.4.10 give the form of $\theta(h)$. Now assume that (M2) holds. Write

$$\varphi_h = p\int_{\mathbb{R}_+\times\mathbb{R}} \sup_{t\in[0,h]} (f^+(r,t+s))^\alpha \, ds\,\pi(dr) + q\int_{\mathbb{R}_+\times\mathbb{R}} \sup_{t\in[0,h]} (f^-(r,t+s))^\alpha \, ds\,\pi(dr).$$

Then we obtain

$$\begin{aligned}\varphi_h &\geq p\int_0^\infty \int_{\eta^{(1)}-h}^{\eta^{(1)}} \sup_{t\in[0,h]} (f^+(r,t+s))^\alpha \, ds\,\pi(dr)\\ &\quad +q\int_0^\infty \int_{\eta^{(2)}-h}^{\eta^{(2)}} \sup_{t\in[0,h]} (f^-(r,t+s))^\alpha \, ds\,\pi(dr)\\ &= h[pf^{+\,\alpha} + qf^{-\,\alpha}] \qquad (2.5.28)\end{aligned}$$

and with g as in (2.0.12), (2.3.20), respectively,

$$\varphi_h \leq h[pf^{+\,\alpha} + qf^{-\,\alpha}] + \int_{\mathbb{R}_+\times\mathbb{R}} g(r,s)^\alpha \, ds\,\pi(dr). \qquad (2.5.29)$$

Thus, we conclude by (2.5.28), (2.5.29),

$$\lim_{h\to\infty} \varphi_h/h = pf^{+\,\alpha} + qf^{-\,\alpha}.$$

□

The extremal index function $\theta(h) < 1$ for every $h > 0$. This result can be interpreted that in short time intervals, exceedances of $\{M_k^{(h)}\}_{k\in\mathbb{N}}$ occur in clusters, where the mean cluster size tends to 1, as h tends to ∞. This is obvious as cluster sizes will be smaller, because more data are condensed into one block.

Another possibility of building blocks is to divide the positive real line into intervals I_k given by (2.5.5). Then we can also measure cluster sizes.

Corollary 2.5.6 (Point process of exceedances)
Let Y be given as in Theorem 2.5.1 with $f^+ \leq 1$. Suppose $\{\widetilde{s}_k\}_{k\in\mathbb{N}}$ are the jump times of a Poisson process with intensity $x^{-\alpha}$, $x > 0$ be fixed, independent of the i. i. d. sequence $\{\zeta_k\}_{k\in\mathbb{Z}}$ with d. f.

$$\pi_k = \mathbb{P}(\zeta_1 = k) = p\left[\mathbb{E}f_k^{(1)\,\alpha} - \mathbb{E}f_{k+1}^{(1)\,\alpha}\right] - (1-p)\left[\mathbb{E}f_k^{(2)\,\alpha} - \mathbb{E}f_{k+1}^{(2)\,\alpha}\right] \quad \text{for } k\in\mathbb{N},$$

where $f^+ = f_1^{(1)} > f_2^{(1)} > \ldots$ are the order statistics of $\{\sup_{h\in I_j^{(1)}}\{f^+(R_1,h)\}\}_{j\in\mathbb{Z}}$ and $f_1^{(2)} > f_2^{(2)} > \ldots$ are the order statistics of $\{\sup_{h\in I_j^{(1)}}\{-f^-(R_1,h)\}\}_{j\in\mathbb{Z}}$. Then

$$\sum_{k=1}^\infty \varepsilon_{(\Gamma_k/n,\sup_{h\in I_k} Y(h)/a_n)}(\cdot\times(x,\infty)) \overset{w}{\Longrightarrow} \sum_{k=1}^\infty \zeta_k \varepsilon_{\widetilde{s}_k}.$$

Proof.
Note, that

$$\sum_{k=1}^{\infty} \varepsilon_{(s_k, f+P_k)}(\cdot \cap [0,\infty) \times (x,\infty)) \stackrel{d}{=} \sum_{k=1}^{\infty} \varepsilon_{(\widetilde{s}_k, f+J_k)}, \tag{2.5.30}$$

where $\{J_k\}_{k\in\mathbb{N}}$ are i. i. d. with d. f. $F(y) = 1 - y^{-\alpha}x^{\alpha}$ for $y > x$, independent of $\{\widetilde{s}_k\}_{k\in\mathbb{N}}$. By Theorem 2.5.1 and (2.5.30)

$$\begin{aligned}&\sum_{k=1}^{\infty} \varepsilon_{(\Gamma_k/n, \sup_{h\in I_k} Y(h)/a_n)}(\cdot \times (x,\infty))\\ &\qquad \stackrel{w}{\Longrightarrow} \sum_{k=1}^{\infty} \sum_{j=-\infty}^{\infty} \varepsilon_{(\widetilde{s}_k, \sup_{h\in I_j^{(k)}}\{f(R_k,h)\chi_k J_k\})}(\cdot \times (x,\infty)) = \sum_{k=1}^{\infty} \zeta_k \varepsilon_{\widetilde{s}_k},\end{aligned}$$

with $\zeta_k = \operatorname{card}\{j : \sup_{h\in I_j^{(k)}}\{f(R_k,h)\chi_k J_k\} > x\}$. It remains to show that ζ_k has d. f. π. Then for $k \in \mathbb{N}$

$$\begin{aligned}\pi_k &= \mathbb{P}(\zeta_1 = k)\\ &= p\left[\mathbb{P}(f_k^{(1)} J_1 > x) - \mathbb{P}(f_{k+1}^{(1)} J_1 > x)\right] + (1-p)\left[\mathbb{P}(f_k^{(2)} J_1 > x) - \mathbb{P}(f_{k+1}^{(2)} J_1 > x)\right]\\ &= p\left[\mathbb{E} f_k^{(1)\,\alpha} - \mathbb{E} f_{k+1}^{(1)\,\alpha}\right] - (1-p)\left[\mathbb{E} f_k^{(2)\,\alpha} - \mathbb{E} f_{k+1}^{(2)\,\alpha}\right]\end{aligned}$$

and $\pi_0 = \mathbb{P}(\zeta_1 = 0) = 1 - p f^{+\,\alpha} - (1-p)\mathbb{E} f_1^{(2)\,\alpha}$. □

In the case of a positive shot noise process with non-increasing kernel function, the last result represents the cluster intensities among local extremes of the process.

2.5.5 Examples

Example 2.5.7 (Discrete-time MA process)
Let $\xi = \{\xi_k\}_{k\in\mathbb{Z}}$ be an i. i. d. sequence of r. v. s, which are regularly varying in the sense of (2.1.2) with measure σ given by (2.0.9), and let $\{c_k\}_{k\in\mathbb{Z}}$ be a sequence of real constants. Define the discrete-time MA process

$$Y_n = \sum_{k=-\infty}^{\infty} c_{n-k}\xi_k \quad \text{for } n \in \mathbb{Z}.$$

Suppose $\sum_{k=-\infty}^{\infty} |c_k|^{\delta} < \infty$ for $\delta < \alpha$ with either $\delta < 1$ or $\alpha > 1$ and $\mathbb{E}\xi_k = 0$. Then Y is a stationary, regularly varying process (similar to Proposition 2.3.1). As in the proof of Theorem 2.5.1 we have for $n \to \infty$

$$\sum_{k=1}^{\infty} \varepsilon_{(k/n, a_n^{-1} Y_k)} \overset{w}{\Longrightarrow} \sum_{k=1}^{\infty} \sum_{j=-\infty}^{\infty} \varepsilon_{(s_k, c_j P_k)}. \tag{2.5.31}$$

This is the well known result of Davis and Resnick [54], Theorem 2.4, which by our results also hold for long memory processes. If ξ_1 is i. d., this model can be considered as a special case of a Lévy driven MA process Y as in (2.0.3): let $f(t) = \sum_{k=-\infty}^{\infty} c_k \mathbf{1}_{[k-1,k)}(t)$ for $t \in \mathbb{R}$ and $L(1) \overset{d}{=} \xi_1$. Then the continuous-time MA process Y at discrete-time points in $\mathbb{Z}$ is given by

$$Y(n) = \sum_{k=-\infty}^{\infty} c_{n-k}[L(k+1) - L(k)] \text{ for } n \in \mathbb{Z}.$$

It coincides in distribution with the discrete-time MA process $\{Y_n\}_{n\in\mathbb{N}}$. Hence we can reformulate the conditions of the existence of a stationary version, tail behavior and extremal behavior of $\{Y_n\}_{n\in\mathbb{N}}$ on this of the continuous-time MA process $\{Y(t)\}_{t\in\mathbb{R}}$ and obtain with long memory processes a larger class of discrete-time MA processes than Davis and Resnick [54]. $\square$

Example 2.5.8 (supOU process, Continuation of Example 2.2.10)
We investigate the extremal behavior of the supOU process Y given by (2.2.18) with kernel function $f(r,s) = \mathbf{1}_{[0,\infty)}(s)\mathrm{e}^{-rs}$ driven by an i. d. i. s. r. m. Λ with generating quadruple $(m, 0, \nu, \pi)$, where $\int r^{-1}\,\pi(dr) < \infty$ and $(m, 0, \nu)$ is the generating triplet of the Lévy process L. Suppose $L(1)$ is regularly varying of index $-\alpha$, $\alpha > 0$ such that for a sequence $0 < a_n \uparrow \infty$ of constants

$$n\mathbb{P}(a_n^{-1} L(1) \in \cdot) \overset{v}{\Longrightarrow} \sigma(\cdot) \quad \text{ on } \mathcal{B}(\overline{\mathbb{R}}\backslash\{0\}) \text{ for } n \to \infty,$$

with $\sigma(dx) = p\alpha x^{-\alpha-1}\,\mathbf{1}_{(0,\infty)}(x)\,dx + q\alpha(-x)^{-\alpha-1}\,\mathbf{1}_{(-\infty,0)}(x)\,dx$ for some $p \in (0,1]$ and $q = 1-p$. Let $0 = t_0 < t_1 < \cdots < t_d$. Then holds by Theorem 2.5.3

$$\sum_{k=1}^{\infty} \varepsilon_{(\Gamma_k/n, \{Y(\Gamma_k+t_i)/a_n\}_{i=0,\dots,d})} \overset{w}{\Longrightarrow} \sum_{k=1}^{\infty} \sum_{j=0}^{\infty} \varepsilon_{(s_k, \{\exp(-R_k(T_{k,j}+t_i))\chi_k P_k\}_{i=0,\dots,d})}.$$

Thus, if an exceedance over a high level at the discrete-time skeleton $\{\Gamma_k + t_i : k \in \mathbb{N}, i = 0, \dots, d\}$ occurs, then we have an extreme at $Y(\Gamma_k)$ for some $k \in \mathbb{N}$. This exceedance is carried on by the exponential decreasing function $\mathrm{e}^{(-R_k \cdot)}$. Furthermore, if Y has a. s. sample path in $\mathbb{D}(\mathbb{R})$, e. g. $\int_{-\infty}^{\infty} 1 \wedge |x|\,\nu(dx) < \infty$, then the running maxima are in the maximum domain of attraction of the Fréchet distribution with

$$\lim_{T\to\infty} \mathbb{P}(a_T^{-1} M(T) \le x) = \exp(-x^{-\alpha}) \quad \text{for } x > 0,$$

by Theorem 2.5.4. We obtain by Corollary 2.5.5 the extremal index function $\theta(h) = h\lambda/(h+\lambda) < 1$ for $h > 0$. Thus extremes occur in clusters. □

Remark 2.5.9
The results of this chapter can be extended to mixed MA processes driven by an i. d. i. s. r. m. Λ in $\mathbb{R}_+^d \times \mathbb{R}$, whose stationary distribution has the cumulant generating function $\psi_A(u) = \lambda(A)\psi(u)$, where ψ is the cumulant generating function of a Lévy process and $\lambda(d\omega) = \pi_1(dr_1) \times \cdots \times \pi_d(dr_d) \times dt$ for $\omega = (r_1, \ldots, r_d, t) \in \mathbb{R}_+^d \times \mathbb{R}$ and π_i, $i = 1, \ldots, d$, are probability measures on $\mathbb{R}_+$. □

Chapter 3

Extremal behavior of stochastic volatility models

The classical pricing model is the Black-Scholes model given by the SDE

$$dS_t = rS_t\,dt + \sigma S_t\,dB_t\,, \quad S_0 = x \in \mathbb{R}\,, \tag{3.0.1}$$

where $r \in \mathbb{R}$ is the stock-appreciation rate, $\sigma > 0$ is the volatility and B is the standard Brownian motion. The Black-Scholes model is based on the assumption that the relative price changes of the asset form a Gaussian process with stationary and independent increments. The crucial parameter is the volatility σ, which is in this model assumed to be constant. However, empirical analysis of stock volatility has already shown in the 1970ies that volatility is not constant, quite the contrary, it is itself stochastic and varies in time.

This observation has led to a vast number of volatility models in discrete-time as well as in continuous-time. In this chapter we concentrate on continuous-time volatility models. Moreover, we are concerned with the so-called stylized facts of volatility as e. g.

- volatility changes in time,
- volatility is random,
- volatility has heavy tails,
- volatility clusters on high levels.

Introducing a stochastic volatility extends the Black-Scholes model to

$$dS_t = rS_t dt + \sqrt{V_t} S_t \, dB_t \,,$$

where V can in principle be any positive stationary stochastic process.

Within the framework of SDEs quite natural models are easily defined. Common examples are the *Ornstein-Uhlenbeck (OU) process*

$$dV_t = -\lambda V_t dt + \sigma dZ_t \,, \tag{3.0.2}$$

where $\lambda, \sigma > 0$ and Z is a driving process, often a second Brownian motion, independent of B. As this is a Gaussian model, it is not a positive process. Alternatively, a *Cox-Ingersoll-Ross (CIR) model* has been suggested as a volatility model, defined by

$$dV_t = \lambda(a - V_t)dt + \sigma \sqrt{V_t} dZ_t \,, \tag{3.0.3}$$

where $\lambda, a, \sigma > 0$ and $\lambda a \geq \sigma^2/2$. The parameter a is the long-term mean of the process and λ the rate of mean reversion. Again in the classical model Z is a standard Brownian motion, independent of B.

Apart from the fact that Gaussian OU processes are not positive, another stylized fact is also violated: empirical volatility exhibits heavy tails, consequently, again the OU model as a Gaussian model seems not very appropriate. Changing the constant σ to a time dependent diffusion coefficient σV_t^γ for $\gamma \in [1/2, \infty)$ and including a linear drift yields to positive stationary models with arbitrarily heavy tails. This has been shown in Borkovec and Klüppelberg [32]. Such models are called *generalized Cox-Ingersoll-Ross models*, the parameter $\gamma = 1/2$ corresponds to the classical CIR model of (3.0.3).

On the other hand, a constant σ is attractive and an alternative way to generate heavy tails in the volatility is to replace the driving Gaussian process in (3.0.2) by a Lévy process with heavier tailed increments. Furthermore, the upward jumps often observed in empirical volatility cannot be modelled by a continuous process. So Lévy processes with jumps as driving processes seem to be quite natural. Such an OU process is positive, provided the driving Lévy process has only positive increments and no Gaussian component; i. e. it is a *subordinator*. This is exactly what Barndorff-Nielsen and Shephard [15, 16] have suggested, modelling the (right-continuous) volatility process as a *Lévy-OU process*. Their stochastic volatility pricing model is given by

$$\begin{aligned} dS_t &= (a + bV_t)dt + \sqrt{V_t}\, dB_t\,, \\ dV_t &= -\lambda V_t dt + \sigma dL_{\lambda t}\,, \end{aligned} \tag{3.0.4}$$

where $a, b \in \mathbb{R}$, $\lambda > 0$ and L is a subordinator, called the *background driving Lévy process* (BDLP), independent of the Brownian motion B.

A completely different approach to obtain continuous-time volatility models starts with a GARCH model and derives from this discrete-time model a continuous-time model. A natural idea is a diffusion approximation; see e. g. Drost and Werker [59] and the references therein. This approach leads to stochastic volatility models of the type

$$\begin{aligned} dS_t &= \sqrt{V_t}\, dB_t^{(1)}\,, \\ dV_t &= \lambda(a - V_t)dt + \sigma V_t dB_t^{(2)}\,, \end{aligned} \tag{3.0.5}$$

i. e. V is a generalized CIR model with parameter $\gamma = 1$. The two processes $B^{(1)}, B^{(2)}$ are independent Brownian motions.

A different approach has been considered by Klüppelberg et al. [87], who started with a discrete-time GARCH(1,1) model and replaced the noise variables by a Lévy process L with jumps $\Delta L_t = L_t - L_{t-}$, $t \geq 0$. This yields to a stochastic volatility model of the type

$$\begin{aligned} dS_t &= \sqrt{V_t}\, dL_t\,, \qquad (3.0.6) \\ dV_{t+} &= \beta dt + V_t \mathrm{e}^{X_{t-}} d(\mathrm{e}^{-X_t})\,, \end{aligned}$$

where $\beta > 0$ and V is left-continuous. The auxiliary càdlàg process X is defined by

$$X_t = t \log \eta - \sum_{0<s\leq t} \log(1 + \lambda\eta(\Delta L_s)^2)\,, \tag{3.0.7}$$

for $\eta > 1$ and $\lambda \geq 0$. This continuous-time GARCH(1,1) model is called a COGARCH(1,1) model.

Our chapter focuses on the extremal behavior of stationary continuous-time stochastic volatility models. This can be described by the tail behavior of the stationary distribution and by the behavior of the process above high thresholds.

The tail behavior models the size of the fluctuations of V and determines the maximum domain of attraction (MDA) of the model. The notation of MDA is defined in Fisher-Tippett's theorem; see Theorem A.1.5. We distinguish MDA(Φ_α), MDA(Λ) and MDA(Ψ_α), for $\alpha > 0$, respectively. Distribution functions in MDA(Φ_α) have regularly varying tails: they are heavy-tailed in the sense that not all moments are finite; see Definition A.1.1. Distribution functions in MDA(Λ) have tails ranging from semi-heavy tails to very light tails. Distribution functions in MDA(Ψ_α) have support bounded to the right. Financial risk is usually considered as having unbounded support above, hence MDA(Ψ_α) is inappropriate in our context and will play no further role in this chapter.

The description of a continuous-time process above a high threshold depends on the sample path behavior of the process. When classical volatility models, driven by Brownian motion, have continuous sample paths with infinite variation, some discrete-time skeleton is introduced. A standard concept is based on so-called ϵ-upcrossings, see Definition 3.1.4, which is only valid for processes with continuous sample paths.

For Lévy driven models large jumps (for instance larger than one) constitute a natural discrete-time skeleton, which can be utilized. One denotes by $(\Gamma_k)_{k\in\mathbb{N}}$ the random time points on $[0,\infty)$, where the driving Lévy process jumps and exceeds a given threshold. The bivariate process $(\Gamma_k, V_{\Gamma_k})_{k\in\mathbb{N}}$ is interpreted as the coordinates of a point process in $[0,\infty)\times\mathbb{R}_+$. As usual we define point processes via Dirac measures. Recall that for any Borel sets $A\times B\subseteq[0,\infty)\times\mathbb{R}_+$ the measure $\sum_{k=1}^{\infty}\varepsilon\{\Gamma_k, V_{\Gamma_k}\}(A\times B)$ counts how often $\Gamma_k\in A$ and $V_{\Gamma_k}\in B$.

After appropriate normalization in time and space these point processes may converge and the limit process may allow for an interpretation, thus providing a description of the extreme behavior of the volatility process. Under weak dependence in the data we obtain as limit a Poisson random measure with mean measure ϑ (PRM(ϑ)); see Definition A.3.5. Moreover, the two components of ϑ are independent and consist of the Poisson measure in time and the negative logarithm of an extreme value distribution in space. Under strong dependence the limit is a cluster Poisson random measure. All these considerations concern the discrete-time skeleton only and ignore the fact that we deal with continuous-time processes.

In the case of a driving Lévy process with jumps, in principle also the small jumps can influence the extreme behavior. In a very close neighborhood of a jump time Γ_k

infinitely many small jumps can happen; they may contribute to the extreme behavior around Γ_k. To investigate the influence of these small jumps and the Gaussian component we consider the process V at each point Γ_k in a surrounding interval I_k. Finally, in certain situations we investigate also the process V after it has reached a local supremum. With each point Γ_k an excursion of V over a high threshold starts. Interesting questions concern the length of the excursion, the rate of "decrease" after Γ_k. We answer these questions at least for some models considered in this chapter. This is done by attaching marks to the point process $(\Gamma_k, V_{\Gamma_k})_{k\in\mathbb{N}}$. For our model marks are a vector of values of the process V after Γ_k, hence it describes the finite dimensional distributions of V after Γ_k. The limit process turns out to be different in different regimes.

This chapter is organized as follows. In Section 3.1 we review the extremal behavior of the generalized CIR model, which can belong to different maximum domain of attractions; i. e. such models can have arbitrary tails. Unfortunately, they are not appropriate models in the case of high level volatility clusters in the data.

Section 3.2 deals with Lévy-OU volatility models. Their extremal behavior is characterized by the extremal behavior of the driving Lévy process, so that we have to distinguish between different classes of BDLPes. In Section 3.2.1 this is done for subexponential Lévy processes $L = (L_t)_{t\geq 0}$. According to whether $L_1 \in \mathrm{MDA}(\Phi_\alpha)$ for some $\alpha > 0$ or $L_1 \in \mathrm{MDA}(\Lambda)$, the extremal behavior of the Lévy-OU process is quite different. Then, in Section 3.2.2 we study OU processes with exponential tails. As a prominent example we investigate the Γ-OU process, i. e. the stationary volatility is gamma distributed. As an important larger class we study OU processes, whose BDLP belongs to $\mathcal{S}(\gamma)$ for $\gamma > 0$. This class extend subexponential Lévy processes in a natural way; see Definition A.1.3. It turns out that for all OU processes in Section 3.2, high level volatility clusters are exhibited only in the case of regularly varying BDLPes.

The last class of models reviewed in this chapter concerns the COGARCH process in Section 3.3. In contrast to the Lévy-OU processes considered earlier, the COGARCH volatility has heavy tails under quite general conditions on the driving Lévy process L. Furthermore, the COGARCH exhibits high level volatility clusters.

Finally, a short conclusion is given in Section 3.4. Here we compare the models introduced in the different sections before. It turns out that there is a striking similarity concerning the extremal behavior of models with the same stationary distribution. Here we discuss briefly some further empirical facts of volatility.

As not to disturb the flow of arguments we postpone classical definitions and concepts to an Appendix. Throughout this chapter we shall use the following notation: We abbreviate distribution function by d. f. and random variable by r. v. For any d. f. F we denote its tail $\overline{F} = 1 - F$. For two r. v. s X and Y with d. f. s F and G we write $X \stackrel{d}{=} Y$ if $F = G$, and by $\stackrel{T\to\infty}{\Longrightarrow}$ we denote weak convergence as $T \to \infty$. For two functions f and g we write $f(x) \sim g(x)$ as $x \to \infty$, if $\lim_{x\to\infty} f(x)/g(x) = 1$. We also denote $\mathbb{R}_+ = (0, \infty)$. For $x \in \mathbb{R}$, let $x^+ = \max\{x, 0\}$ and $\log^+(x) = \log(\max\{x, 1\})$. Integrals of the form $\int_a^b$ will be interpreted as the integral taken over the interval $(a, b]$.

3.1 Extremal behavior of generalized Cox-Ingersoll-Ross models

In this section we summarize some well-known results on classical volatility models as defined in (3.0.3) and (3.0.5) driven by a standard Brownian motion. This section is based on Borkovec and Klüppelberg [32]; for a review see [84], Section 3.

As all models above fall into the framework of *generalized Cox-Ingersoll-Ross models* (GCIR) models, we restrict ourselves to stationary solutions of the SDE

$$dV_t = \lambda(a - V_t)dt + \sigma V_t^\gamma dB_t \,, \tag{3.1.1}$$

where $\gamma \in \left[\frac{1}{2}, \infty\right)$. For $\lambda, a, \sigma > 0$ (in the case $\gamma = 1/2$ additionally $\lambda a \geq \sigma^2/2$ is needed) these models are ergodic with state space $\mathbb{R}_+$ and have a stationary density.

Associated with the diffusion (3.1.1) is the scale function s and the speed measure m. The *scale function* is defined as

$$s(x) = \int_z^x \exp\left(-\frac{2\lambda}{\sigma^2}\int_z^y \frac{a-t}{t^{2\gamma}}dt\right) dy \qquad \text{for } x \in \mathbb{R}_+ \,, \tag{3.1.2}$$

where z is any interior point of $\mathbb{R}_+$ whose choice does not affect the extremal behavior. For the *speed measure* m we know that it is finite for the GCIR model. Moreover, m is absolutely continuous with Lebesgue density

$$m'(x) = \frac{2}{\sigma^2 x^{2\gamma} s'(x)} \qquad \text{for } x \in \mathbb{R}_+ \,,$$

where s' is the Lebesgue density of s. Then the stationary density of V is given by

$$f(x) = m'(x)/m(\mathbb{R}_+) \qquad \text{for } x \in \mathbb{R}_+ \,. \tag{3.1.3}$$

Proposition 3.1.1

Let V be a GCIR *model given by equation* (3.1.1) *and define* $M(T) = \sup_{t\in[0,T]} V_t$ *for $T > 0$. Then for any initial value $V_0 = y \in \mathbb{R}_+$ and any $u_T \uparrow \infty$,*

$$\lim_{T\to\infty} |\mathbb{P}_y(M(T) \le u_T) - H^T(u_T)| = 0\,,$$

where H is a d. f., defined for any $z \in \mathbb{R}_+$ by

$$H(x) = \exp\left(-\frac{1}{m(\mathbb{R}_+)s(x)}\right) \qquad \text{for } x > z\,. \tag{3.1.4}$$

The function s and the quantity $m(\mathbb{R}_+)$ depend on the choice of z. □

Corollary 3.1.2 (Running maxima)

Let the assumptions of Proposition 3.1.1 hold. Assume further that $H \in \mathrm{MDA}(G)$ *for $G \in \{\Phi_\alpha, \alpha > 0, \Lambda\}$ with norming constants $a_T > 0$, $b_T \in \mathbb{R}$. Then*

$$\lim_{T\to\infty} \mathbb{P}(a_T^{-1}(M(T) - b_T) \le x) = G(x)\,, \quad x \in \mathbb{R}\,.$$

It is clear that the d. f. H decides about the extremal behavior of V. We present four cases.

Example 3.1.3 (Tail behavior of GCIR models)

Let V be a stationary GCIR model given by equation (3.1.1) with stationary density f, corresponding d. f. F, and d. f. H as given in (3.1.4). Recall that a $\Gamma(\mu,\gamma)$ distributed r. v. has probability density

$$p(x) = \frac{\gamma^\mu}{\Gamma(\mu)} x^{\mu-1} \mathrm{e}^{-\gamma x} \qquad \text{for } x > 0\,, \tag{3.1.5}$$

$\mu > 1$ and $\gamma > 0$.

(1) $\gamma = \frac{1}{2}$: The stationary density of V is $\Gamma\left(\frac{2\lambda a}{\sigma^2}, \frac{2\lambda}{\sigma^2}\right)$. The tail of H behaves like

$$\overline{H}(x) \sim \frac{2\lambda^2 a}{\sigma^2} x \overline{F}(x) \qquad \text{for } x \to \infty,$$

so that the tail of H is that of a $\Gamma\left(\frac{2\lambda a}{\sigma^2} + 1, \frac{2\lambda}{\sigma^2}\right)$ distribution. Hence $H \in \mathrm{MDA}(\Lambda)$ with norming constants

$$a_T = \frac{\sigma^2}{2\lambda} \quad \text{and} \quad b_T = \frac{\sigma^2}{2\lambda}\left[\log T + \frac{2\lambda a}{\sigma^2}\log\log T - \log\left(\frac{\lambda}{\Gamma(2\lambda a/\sigma^2)}\right)\right].$$

(2) $\frac{1}{2} < \gamma < 1$: The stationary density of V is given by

$$f(x) = \frac{2}{A\sigma^2} x^{-2\gamma} \exp\left(-\frac{2}{\sigma^2}\left(\frac{\lambda a}{2\gamma-1}x^{-(2\gamma-1)} + \frac{\lambda}{2-2\gamma}x^{2-2\gamma}\right)\right) \qquad \text{for } x > 0,$$

with some normalizing factor $A > 0$. The d. f. H has tail

$$\overline{H}(x) \sim Bx^{2(1-\gamma)}\overline{F}(x) \quad \text{for } x \to \infty, \tag{3.1.6}$$

where $B > 0$. Hence $H \in \text{MDA}(\Lambda)$ with norming constants

$$a_T = \frac{\sigma^2}{2\lambda}\left(\frac{\sigma^2(1-\gamma)}{\lambda}\log T\right)^{\frac{2\gamma-1}{2-2\gamma}},$$

$$b_T = \left(\frac{\sigma^2(1-\gamma)}{\lambda}\log T\right)^{\frac{1}{(2-2\gamma)}}\left(1 - \frac{2\gamma-1}{(2-2\gamma)^2}\frac{\log\left(\frac{\sigma^2(1-\gamma)}{\lambda}\log T\right)}{\log T}\right) + a_T \log\left(\frac{2\lambda}{A\sigma^2}\right).$$

(3) $\gamma = 1$: The stationary density of V is inverse gamma, i. e.

$$f(x) = \left(\frac{\sigma^2}{2\lambda a}\right)^{-\frac{2\lambda}{\sigma^2}-1}\left(\Gamma\left(\frac{2\lambda}{\sigma^2}+1\right)\right)^{-1} x^{-\frac{2\lambda}{\sigma^2}-2}\exp\left(-\frac{2\lambda a}{\sigma^2}x^{-1}\right) \qquad \text{for } x > 0,$$

so that $V_0 \in \mathcal{R}_{-2\lambda/\sigma^2-1}$. In this case $\overline{H}(x) \sim cx^{-2\lambda/\sigma^2-1}$ for $x \to \infty$ and for some $c > 0$. Hence $H \in \text{MDA}(\Phi_\alpha)$ for $\alpha = 2\lambda/\sigma^2 + 1$ with norming constants $a_T = (cT)^{\sigma^2/(2\lambda+\sigma^2)}$ and $b_T = 0$.

(4) $\gamma > 1$: The stationary density f of V has the same form as in *(2)*, but is regularly varying of index $-2\gamma+1$. Now the tail of H becomes very extreme: $\overline{H}(x) \sim cx^{-1}$. Hence $H \in \text{MDA}(\Phi_1)$ with $a_T = cT$ and $b_T = 0$. □

Since all models (3.1.1) are driven by a Brownian motion, they have continuous sample paths; i. e. there is no natural discrete-time skeleton. We follow the standard approach to create a discrete-time skeleton of the process; see e. g. Leadbetter et al. [95], Chapter 12. This allows for a more profound extreme value analysis of V.

Definition 3.1.4
Let V be a stationary version of the diffusion given by (3.1.1). *V is said to have an ϵ-upcrossing of the level u at a point $\Gamma > 0$ if*

$$V_t < u \quad \textit{for} \quad t \in (\Gamma - \epsilon, \Gamma) \quad \textit{and} \quad V_\Gamma = u\,.$$

With this definition we can formulate a further result describing the extreme behavior of a stationary GCIR model.

Theorem 3.1.5 (Point process of ϵ-upcrossings)
Let V be a stationary version of the diffusion given by (3.1.1) *with d. f. H as in* (3.1.4). *Let $a_T > 0$, $b_T \in \mathbb{R}$ be the norming constants as given in Example 3.1.3. Let $(\Gamma_{T,k})_{k\in\mathbb{N}}$ be the time points on $\mathbb{R}_+$, where the ϵ-upcrossings of V of the level $a_T x + b_T$ occur. Let $(j_k)_{k\in\mathbb{N}}$ be the jump times of a Poisson process with intensity* e^{-x} *if $\gamma \in [1/2, 1)$, and $x^{-\alpha}$ with $\alpha = 2\lambda/\sigma^2 + 1$ if $\gamma = 1$ and $\alpha = 1$ if $\gamma > 1$. Then*

$$\sum_{k=1}^{\infty} \varepsilon\left\{\Gamma_{T,k}/T\right\} \stackrel{T\to\infty}{\Longrightarrow} \sum_{k=1}^{\infty} \varepsilon\{j_k\}\,.$$

As is obvious from this result ϵ-upcrossings of V for high levels behave like of i. i. d. data, i. e. such models do not exhibit volatility clusters. They can, however, model heavy tails as the running maxima depend on the d. f. H.

3.2 Extremal behavior of Lévy-OU volatility models

We start with a precise definition of a positive Lévy-OU process as a solution of (3.0.4). For more information on Lévy processes we refer to the excellent monographs by Sato [138], Bertoin [27] and Cont and Tankov [47]. Let L be a subordinator; i. e. L is a Lévy process with increasing sample paths, hence they are of bounded variation, and we assume that they are càdlàg. The Laplace transform is then the natural transform and has for all $t \geq 0$ the representation

$$\mathbb{E}\exp(-\lambda L_t) = \exp(-t\Psi(\lambda)) \qquad \text{for } \lambda \geq 0\,.$$

The Laplace exponent Ψ has representation

$$\Psi(\lambda) = m\lambda + \int_{(0,\infty)} (1 - \mathrm{e}^{-\lambda x})\,\nu(dx)\,.$$

As there is no Gaussian component the characteristic triplet of arbitrary Lévy processes reduces to a pair (m, ν), where $m > 0$ is the drift and the Lévy measure ν has support on $\mathbb{R}_+$ and satisfies

$$\int_{(0,\infty)} (1 \wedge x)\,\nu(dx) < \infty\,.$$

Let $\lambda > 0$. We denote by

$$V_t = \mathrm{e}^{-\lambda t} V_0 + \int_0^t \mathrm{e}^{-\lambda(t-s)}\, dL_{\lambda s} \quad \text{for } t \in \mathbb{R} \tag{3.2.1}$$

the solution to the SDE (3.0.4). Then V becomes a càdlàg process, which is used in this section.

A stationary solution of (3.0.4) exists if and only if $\int_{x>2} \log x\, \nu(dx) < \infty$. Note that this condition is only violated for Lévy measures with extremely heavy tails; more precisely ν needs to have slowly varying tails. As all models considered in this chapter have tails which are regularly varying of some negative index or lighter, all our models satisfy this stationarity condition. Stationarity is then achieved, if V_0 is taken to be independent of the driving Lévy process L and has distribution

$$V_0 \stackrel{d}{=} \int_0^\infty \mathrm{e}^{-s}\, dL_s \,.$$

A convenient representation for the stationary version is

$$V_t = \mathrm{e}^{-\lambda t} \int_{-\infty}^t \mathrm{e}^{\lambda s}\, dL_{\lambda s} \quad \text{for } t \geq 0. \tag{3.2.2}$$

In this representation, L is extended to a Lévy process on the whole real line, by letting $\widetilde{L} = (\widetilde{L}_t)_{t\geq 0}$ be an independent copy of $(L_t)_{t\geq 0}$, and defining $L_t := -\widetilde{L}_{-t-}$ for $t < 0$. The parameter λ in the process L in (3.2.1) ensures that the stationary marginal distribution of V is independent of λ; indeed it is given by (3.2.2).

The r. v. V_0 is infinitely divisible with the characteristic pair (m_V, ν_V), where $m_V = m$ and

$$\nu_V\,[x, \infty) = \int_x^\infty u^{-1} \nu\,[u, \infty)\, du \quad \text{for } x > 0\,. \tag{3.2.3}$$

We are concerned with Lévy processes L, which are heavy or semi-heavy tailed; i. e. whose tails decrease not faster than exponentially. As indicated in (3.2.7) and (3.2.10) this induces a similar tail behavior on V, which is in accordance with empirical findings.

The structure of a Lévy-OU volatility process can be best understood when considering the following example.

Example 3.2.1 (Positive Poisson shot noise process)
Let L be a positive compound Poisson process with characteristic pair $(0, \mu\mathbb{P}_F)$, where $\mu > 0$ and $\mathbb{P}_F$ is a probability measure on $\mathbb{R}_+$ with corresponding d. f. F. Then L has the representation

$$L_t = \sum_{j=1}^{N_t} \xi_j \quad \text{for } t > 0\,, \tag{3.2.4}$$

where $(N_t)_{t\geq 0}$ is a Poisson process on $\mathbb{R}_+$ with intensity $\mu > 0$ and jump times $(\Gamma_k)_{k\in\mathbb{N}}$. The process N is independent of the i. i. d. sequence of positive r. v. s $(\xi_k)_{k\in\mathbb{N}}$ with d. f. F.

The resulting volatility process is then the positive shot noise process

$$V_t = \mathrm{e}^{-\lambda t} V_0 + \int_0^t \mathrm{e}^{-\lambda(t-s)}\, dL_{\lambda s} = \mathrm{e}^{-\lambda t} V_0 + \sum_{j=1}^{N_{\lambda t}} \mathrm{e}^{-\lambda t + \Gamma_j}\, \xi_j \quad \text{for } t \geq 0\,,$$

and from (3.2.3) we get

$$\nu_V\,[x,\infty) = \mu \int_x^\infty u^{-1}\overline{F}(u)\, du \quad \text{for } x > 0.$$

If $\mathbb{E}\log(1+\xi_1) < \infty$, a stationary solution exists in which case V can be represented as

$$V_t = \mathrm{e}^{-\lambda t} \sum_{\substack{j=-\infty \\ j\neq 0}}^{N_{\lambda t}} \mathrm{e}^{\Gamma_j}\xi_j \quad \text{for } t > 0. \tag{3.2.5}$$

Here, we let $(\xi_k)_{k\in-\mathbb{N}_0}$ and $(\Gamma_k)_{k\in-\mathbb{N}_0}$ be sequences of r. v. s such that $(\xi_k)_{k\in\mathbb{Z}}$ and $(\Gamma_k)_{k\in\mathbb{Z}}$ are independent. Furthermore, $(\xi_k)_{k\in\mathbb{Z}}$ is an i. i. d. sequence and $(-\Gamma_k)_{k\in-\mathbb{N}}$ are the jump times of a Poisson process on $\mathbb{R}_+$ with intensity μ, independent of $(\Gamma_k)_{k\in\mathbb{N}}$; further, we define $\Gamma_0 := 0$.

The qualitative extreme behavior of this volatility process can be seen in Figure 3.1. The volatility jumps upwards, whenever $(N_{\lambda t})_{t\geq 0}$ jumps and decreases exponentially fast between two jumps. This means in particular that V has local suprema exactly at the jump times Γ_k/λ (and $t = 0$), i.e

$$V_t = V_{\Gamma_k/\lambda}\mathrm{e}^{-\lambda t+\Gamma_k} \quad \text{for } t \in [\Gamma_k/\lambda, \Gamma_{k+1}/\lambda)\,.$$

Consequently, it is the discrete-time skeleton of V at points Γ_k/λ that determines the extreme behavior of the volatility process. □

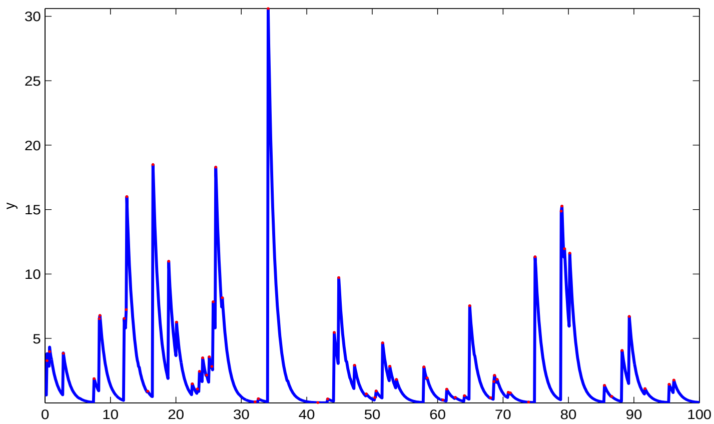


Figure 3.1: Sample path of a Weibull-OU process as given in Example 3.2.3 with $\lambda = 1$ and $p = 1/2$.

For a general subordinator L we decompose

$$L = L^{(1)} + L^{(2)} \tag{3.2.6}$$

into two independent Lévy processes, where $L^{(1)}$ has *characteristic pair* $(0, \nu_1)$ with $\nu_1(x, \infty) = \nu(x, \infty)\, \mathbf{1}_{(1,\infty)}(x)$ and $L^{(2)}$ has characteristic pair (m, ν_2) with $\nu_2(x, \infty) = \nu\,(x, 1]\, \mathbf{1}_{(0,1]}(x)$. Then again $L^{(1)}$ is a compound Poisson process with intensity $\nu(1, \infty)$ and jump sizes with d. f. $\nu_1/\nu(1, \infty)$. All the small jumps and the drift are summarized in $L^{(2)}$.

What is needed, however, are the precise asymptotic links between the tails of V, L and the tail of the Lévy measure $\nu(\cdot, \infty)$. This implies immediately that we have to distinguish different regimes defined by the link between L and $\nu(\cdot, \infty)$.

Any infinitely divisible distribution is asymptotically tail-equivalent to its Lévy measure, whenever it is convolution equivalent; see Definitions A.1.3, A.1.7, and Theorem A.1.4.

The class $\mathcal{S}(0) = \mathcal{S}$ of subexponential d. f. s contains all d. f. s with regularly varying tails, but is much larger. Subexponential distributions can be in two different maximum domain of attractions; see Theorem A.1.5. All d. f. s with regularly varying tail (Definition A.1.1) are subexponential and belong to MDA(Φ_α). Other subexponential d. f. s, as for instance the lognormal and the semi-heavy tailed Weibull d. f. s (see Example 3.2.3), belong to MDA(Λ). On the other hand, d. f. s as the gamma distribution or d. f. s in $\mathcal{S}(\gamma)$ for $\gamma > 0$ belong to MDA(Λ), but are lighter tailed than any subexponential distribution. Consequently, we also consider such exponential models below.

3.2.1 Lévy-OU processes with subexponential tails

In this section we are concerned with the Lévy-OU process given by (3.2.1), whose BDLP is subexponential. This section is based on Chapter 1, Chapter 2.

Proposition 3.2.2 (Tail behavior of subexponential models)
Let V be a stationary version of the Lévy-OU process given by (3.2.1) and define $M(h) = \sup_{t\in[0,h]} V_t$ for $h > 0$.

(a) If $L_1 \in \mathcal{S} \cap \mathrm{MDA}(\Phi_\alpha) = \mathcal{R}_{-\alpha}$, then also $V_0 \in \mathcal{R}_{-\alpha}$ and

$$\mathbb{P}(V_0 > x) \sim \alpha^{-1}\,\mathbb{P}(L_1 > x) \qquad \text{for } x \to \infty\,. \tag{3.2.7}$$

Moreover,

$$\mathbb{P}(M(h) > x) \sim \left[\lambda h + \alpha^{-1}\right]\mathbb{P}(L_1 > x) \qquad \text{for } x \to \infty\,. \tag{3.2.8}$$

(b) If $L_1 \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$, then also $V_0 \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$ and

$$\mathbb{P}(V_0 > x) \sim \mathbb{P}(\exp(-U)L_1 > x) \qquad \text{for } x \to \infty\,, \tag{3.2.9}$$

where U is a uniform r. v. on $(0,1)$, independent of L. In particular, $\mathbb{P}(V_0 > x) = o(\mathbb{P}(L_1 > x))$ as $x \to \infty$. More precisely,

$$\mathbb{P}(V_0 > x) \sim \frac{a(x)}{x}\mathbb{P}(L_1 > x) \qquad \text{for } x \to \infty\,, \tag{3.2.10}$$

where a is the function from the representation (A.1.1):

$$\mathbb{P}(L_1 > x) \sim c\exp\left[-\int_0^x \frac{1}{a(y)}\,dy\right] \qquad \text{for } x \to \infty\,,$$

for some $c > 0$ and $a : \mathbb{R}_+ \to \mathbb{R}_+$ absolutely continuous with $\lim_{x\to\infty} a'(x) = 0$ and $\lim_{x\to\infty} a(x) = \infty$. Finally,

$$\mathbb{P}(M(h) > x) \sim \lambda h\mathbb{P}(L_1 > x) \qquad \text{for } x \to \infty\,. \tag{3.2.11}$$

Proof.
By (3.2.3) we have

$$\frac{\nu_V(x,\infty)}{\nu(x,\infty)} = \frac{\int_x^\infty u^{-1}\nu(u,\infty)du}{\nu(x,\infty)} \qquad \text{for } x>0\,. \tag{3.2.12}$$

Assume that $L_1 \in \mathcal{R}_{-\alpha}$ for some $\alpha > 0$. Then by Theorem A.1.4 *(i)* we have $\nu(\cdot,\infty)/\nu(1,\infty) \in \mathcal{R}_{-\alpha}$. By Karamata's theorem (e. g. Embrechts et al. [60], Theorem A 3.6) we obtain

$$\lim_{x\to\infty} \frac{\nu_V(x,\infty)}{\nu(x,\infty)} = \frac{1}{\alpha}\,.$$

This implies in particular that also $\nu_V(\cdot,\infty)/\nu_V(1,\infty) \in \mathcal{R}_{-\alpha}$ and hence, again by Theorem A.1.4 *(i)*, $V_0 \in \mathcal{R}_{-\alpha}$ and (3.2.7) holds.

If $L_1 \in \mathrm{MDA}(\Lambda)\cap\mathcal{S}$ we can only conclude from (3.2.12) that the right hand side tends to 0. To obtain a precise result we proceed as follows. Denote by ξ_1 the jump distribution of the compound Poisson process $L^{(1)}$ as given in (3.2.6). Taking $\nu(\cdot,\infty)/\nu(1,\infty) \in \mathcal{R}_{-\infty}$ into account and applying l'Hospital's rule yields

$$\begin{aligned}\frac{\nu_V(x,\infty)}{\nu(1,\infty)\mathbb{P}(\exp(-U)\xi_1 > x)} &= \frac{\int_x^\infty u^{-1}\nu(u,\infty)\,du}{\int_0^1 \nu(\mathrm{e}^s x,\infty)\,ds} = \frac{\int_x^\infty u^{-1}\nu(u,\infty)\,du}{\int_x^{x\mathrm{e}} u^{-1}\nu(u,\infty)\,du}\\ &\sim \left[1-\frac{\nu(\mathrm{e}\,x,\infty)}{\nu(x,\infty)}\right]^{-1} \overset{x\to\infty}{\longrightarrow} 1.\end{aligned}$$

The tail-equivalence (3.2.9) follows then from the fact that

$$\nu(1,\infty)\mathbb{P}(\exp(-U)\xi_1 > x) \sim \mathbb{P}(\exp(-U)L_1 > x) \qquad \text{for } x\to\infty$$

and Theorem A.1.4 *(i)*.

For proving (3.2.10), by Theorem A.1.4 *(i)* we may assume without loss of generality that there exists a $x_0 > 0$ such that

$$\nu(x,\infty) = c\,\exp\left[-\int_{x_0}^x \frac{1}{a(y)}\,dy\right] \qquad \text{for } x \ge x_0.$$

Then $\nu(dx) = (a(x))^{-1}\nu(x,\infty)\,dx$ and an application of l'Hospital's rule shows that

$$\begin{aligned}\frac{\nu_V(x,\infty)}{\nu(x,\infty)\,a(x)/x} &\sim \frac{-\nu(x,\infty)/x}{\nu(x,\infty)[a'(x)-a(x)/x]/x - (\nu(x,\infty)/a(x))a(x)/x}\\ &= [-a'(x)+a(x)/x+1]^{-1}\\ &\overset{x\to\infty}{\longrightarrow} 1,\end{aligned}$$

since $a(x)/x \sim a'(x)$ and $a'(x) \to 0$ as $x \to \infty$. Theorem A.1.4 *(i)* then gives (3.2.10).

The results for $M(h)$ are based on Theorem 2.1 of Rosinski and Samorodnitsky [132]. They show that for $L_{\lambda h} + V_0 \in \mathcal{S}$

$$\mathbb{P}(M(h) > x) \sim \nu_{L_{\lambda h}+V_0}(x,\infty) \qquad \text{for } x \to \infty\,,$$

implying the result by Theorem A.1.4 *(i)*. □

Example 3.2.3 (Semi-heavy tailed Weibull distribution)
Let L_1 have distribution tail

$$\mathbb{P}(L_1 > x) \sim K \exp(-x^p) \qquad \text{for } x \to \infty\,,$$

for some $K > 0$ and $p \in (0,1)$. As $a(x) = x^{1-p}/p$, we obtain from (3.2.10) immediately

$$\mathbb{P}(V_0 > x) \sim \frac{K}{p} x^{-p} \exp(-x^p) \qquad \text{for } x \to \infty\,.$$ □

Proposition 3.2.2 shows that in the regularly varying regime the tail of V_0 is equivalent to the tail of L_1. In contrast to that, in the $\mathcal{S} \cap \mathrm{MDA}(\Lambda)$ case, the tail of V_0 becomes lighter, due to the influence of $\exp(-U)$. But in both cases V_0 is subexponential and the tail of $M(h)$ is determined by the tail of L_1, only the constants differ.

The following result gives a complete account of the extreme behavior of the volatility process V for a subexponential BDLP L. There are three components considered in (3.2.14) and (3.2.15). The first one is the scaled jump time process of $(L_{\lambda t})_{t\ge 0}$, where only jumps larger than 1 are included. The second component is the normalized local supremum near that jump, and the third component is a vector of normalized values of V after the jump.

Theorem 3.2.4 (Marked point process behavior of models in $\mathcal{S}$)
Let V be a stationary version of the Lévy-OU process given by (3.2.1). Suppose $\Gamma = (\Gamma_k)_{k\in\mathbb{N}}$ are the jump times of $L^{(1)}$ given by (3.2.6) and $I = (I_k)_{k\in\mathbb{N}}$, where $I_k = \frac{1}{2\lambda}\left[\Gamma_{k-1} + \Gamma_k, \Gamma_k + \Gamma_{k+1}\right)$, $\Gamma_0 := 0$. For $m \in \mathbb{N}$ let $0 = t_0 < t_1 < \cdots < t_m$.

(a) Assume that $L_1 \in \mathcal{S} \cap \mathrm{MDA}(\Phi_\alpha)$ with norming constants $a_T > 0$ such that

$$\lim_{T\to\infty} T\mathbb{P}(L_1 > a_T x) = x^{-\alpha} \qquad \text{for } x > 0. \tag{3.2.13}$$

Take $\Gamma^{(k)} = (\Gamma_{k,i})_{i\in\mathbb{N}}$, $k \in \mathbb{N}$, *as i. i. d. copies of the sequence* Γ *and set* $\Gamma_{k,0} := 0$, $\Gamma_{k,-1} := 0$ *for all* $k \in \mathbb{N}$. *Let* $\sum_{k=1}^{\infty} \varepsilon\{s_k, P_k\}$ *be a* PRM(ϑ) *with mean measure* $\vartheta(dt \times dx) = dt \times \alpha x^{-\alpha}\,\mathbf{1}_{(0,\infty)}(x)\,dx$, *independent of the sequences* $(\Gamma^{(k)})_{k\in\mathbb{N}}$. *Then*

$$\sum_{k=1}^{\infty} \varepsilon\left\{\frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1}\sup_{s\in I_k} V_s, \{a_{\lambda T}^{-1} V_{\Gamma_k/\lambda+t_i}\}_{i=0,\dots,m}\right\}$$
$$\overset{T\to\infty}{\Longrightarrow} \sum_{k=1}^{\infty}\sum_{j=0}^{\infty} \varepsilon\left\{s_k, P_k \mathrm{e}^{-(\Gamma_{k,j-1}+\Gamma_{k,j})/2}, \{P_k \mathrm{e}^{-\lambda t_i - \Gamma_{k,j}}\}_{i=0,\dots,m}\right\}. \tag{3.2.14}$$

(b) Assume that $L_1 \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$ *with norming constants* $a_T > 0$, $b_T \in \mathbb{R}$ *such that*

$$\lim_{T\to\infty} T\mathbb{P}(L_1 > a_T x + b_T) = \exp(-x) \qquad \text{for } x \in \mathbb{R}\,.$$

Let $\sum_{k=1}^{\infty} \varepsilon\{s_k, P_k\}$ *be a* PRM(ϑ) *with mean measure* $\vartheta(dt \times dx) = dt \times \mathrm{e}^{-x}\,dx$. *Then*

$$\sum_{k=1}^{\infty} \varepsilon\left\{\frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1}(\sup_{s\in I_k} V_s - b_{\lambda T}), \{a_{\lambda T}^{-1}(V_{\Gamma_k/\lambda+t_i} - b_{\lambda T})\}_{i=0,\dots,m}\right\}$$
$$\overset{T\to\infty}{\Longrightarrow} \sum_{k=1}^{\infty} \varepsilon\{s_k, P_k, (P_k, 0, \dots, 0)\}\,. \tag{3.2.15}$$

Moreover,

$$\sum_{k=1}^{\infty} \varepsilon\left\{\frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1}(\sup_{s\in I_k} V_s - b_{\lambda T}), \{b_{\lambda T}^{-1} V_{\Gamma_k/\lambda+t_i}\}_{i=0,\dots,m}\right\} \tag{3.2.16}$$
$$\overset{T\to\infty}{\Longrightarrow} \sum_{k=1}^{\infty} \varepsilon\{s_k, P_k, \{\exp(-\lambda t_i)\}_{i=0,\dots,m}\}\,.$$

We first give an interpretation of (3.2.15). The limit relations of the first two components show that the local suprema of V around the Γ_k/λ, normalized with the constants determined via L_λ, converge weakly to the same extreme value distribution as L_λ. Moreover, the third component indicates that for $t_0 = 0$ the second and third component have the same limiting behavior; i. e. the $\sup_{s\in I_k} V_s$ behaves like $V_{\Gamma_k/\lambda}$. For all $t_i > 0$ the last component is negligible, i. e. the process is considerably smaller away from $V_{\Gamma_k/\lambda}$.

In the second and third component of (3.2.14) all points $\Gamma_{k,j}$ and not only $\Gamma_{k,0} = 0$ like in (3.2.15) may influence the limit. This phenomenon has certainly its origin in the very large jumps caused by regular variation. Even though the volatility decreases between the jumps exponentially fast, huge jumps can have a long lasting influence on excursions above high thresholds. This is in contrast to the semi-heavy tailed case, where L is subexponential, but in $\mathrm{MDA}(\Lambda)$.

Both relations (3.2.14) and (3.2.16) exhibit, however, a common effect in the third component: if the Lévy process L has an exceedance over a high threshold, then the OU process decreases after this event exponentially fast.

Corollary 3.2.5 (Point process of exceedances)
Let the assumptions of Theorem 3.2.4 hold.

(a) Assume that $L_1 \in \mathcal{S} \cap \mathrm{MDA}(\Phi_\alpha)$. Let $(j_k)_{k\in\mathbb{N}}$ be the jump times of a Poisson process with intensity $x^{-\alpha}$, $x > 0$ being fixed. Let $(\zeta_k)_{k\in\mathbb{Z}}$ be i. i. d. discrete r. v. s, independent of $(j_k)_{k\in\mathbb{N}}$, with probability distribution

$$\pi_k = \mathbb{P}(\zeta_1 = k) = \mathbb{E}\exp(-\alpha(\Gamma_{k-1} + \Gamma_k)/2) - \mathbb{E}\exp(-\alpha(\Gamma_k + \Gamma_{k+1})/2)\,, \quad k \in \mathbb{N}\,.$$

Then

$$\sum_{k=1}^{\infty} \varepsilon\left\{\frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1} \sup_{s\in I_k} V_s\right\}(\cdot \times (x,\infty)) \overset{T\to\infty}{\Longrightarrow} \sum_{k=1}^{\infty} \zeta_k \varepsilon\{j_k\}\,.$$

(b) Assume that $L_1 \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$. Let $(j_k)_{k\in\mathbb{N}}$ be the jump times of a Poisson process with intensity e^{-x}, $x \in \mathbb{R}$ being fixed. Then

$$\sum_{k=1}^{\infty} \varepsilon\left\{\frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1}(\sup_{s\in I_k} V_s - b_{\lambda T})\right\}(\cdot \times (x,\infty)) \overset{T\to\infty}{\Longrightarrow} \sum_{k=1}^{\infty} \varepsilon\{j_k\}\,.$$

Again the qualitative difference of the two regimes is obvious. In the case of a regularly varying BDLP L the limiting process is a compound Poisson process, where at each Poisson point a cluster appears, whose size is random with distribution $(\pi_k)_{k\in\mathbb{N}}$. In contrast to this, in the $\mathrm{MDA}(\Lambda)$ case, the limit process is simply a homogeneous Poisson process; no clusters appear in the limit.

As the next result shows, the running maxima of the volatility process V have the same behavior as that of an i. i. d. sequence of copies of L_λ.

Corollary 3.2.6 (Running maxima)
Let V be a stationary version of the Lévy-OU process given by (3.2.1), and define $M(T) = \sup_{t\in[0,T]} V_t$ for $T > 0$.

(a) Assume that $L_1 \in \mathcal{S} \cap \mathrm{MDA}(\Phi_\alpha)$ with norming constants $a_T > 0$ given by (3.2.13). Then

$$\lim_{T\to\infty} \mathbb{P}\left(a_{\lambda T}^{-1} M(T) \leq x\right) = \exp(-x^{-\alpha}) \qquad \text{for } x > 0\,.$$

(b) Assume that $L_1 \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$ with norming constants $a_T > 0$, $b_T \in \mathbb{R}$ given by (3.2.15). Then

$$\lim_{T\to\infty} \mathbb{P}\left(a_{\lambda T}^{-1}(M(T) - b_{\lambda T}) \leq x\right) = \exp(-\mathrm{e}^{-x}) \qquad \text{for } x \in \mathbb{R}\,.$$

Finally, we investigate the possibility of volatility clusters in the Lévy-OU process. As the concept of ϵ-upcrossings is only defined for continuous-time processes, which does not fit into our framework, we shall introduce an appropriate method for describing clusters in continuous-time processes with jumps.

As our method will be motivated by the discrete-time skeleton of V, we recall that in a discrete-time process clusters are usually described by the extremal index $\theta \in (0,1]$; see Definition A.1.9. However, continuous-time processes are by nature dependent in small time intervals by the continuity and the structure of the process. Thus it is not adequate to adopt the extremal index concept for stochastic sequences to describe the dependence structure of the continuous-time process on a high level.

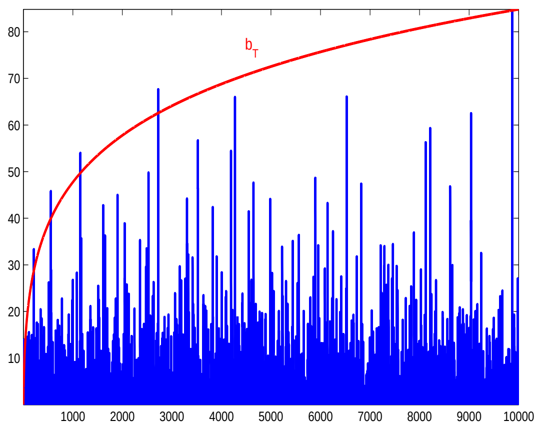


Figure 3.2: Sample path of a Weibull-OU process as given in Example 3.2.3 with $\lambda = 1$ and $p = 1/2$.

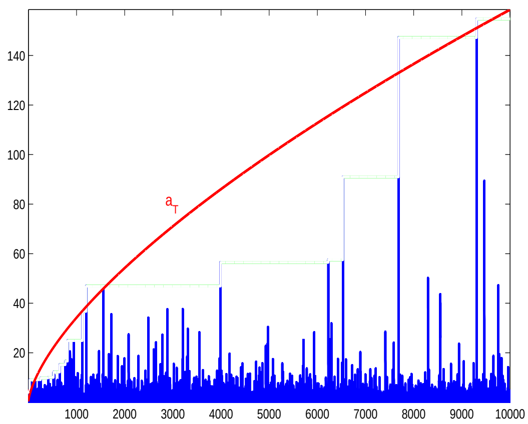


Figure 3.3: Sample path of a 3/2-stable-OU process with $\lambda = 1$

We recall the definition of an extremal index function.

Definition 3.2.7
Let $(V_t)_{t\geq 0}$ be a stationary process. Define the sequence $M_k(h) = \sup_{(k-1)h\leq t\leq kh} V_t$, $k \in \mathbb{N}$, for $h > 0$. Let $\theta(h)$ be the extremal index of the sequence $(M_k(h))_{k\in\mathbb{N}}$. Then we call the function $\theta : (0, \infty) \to [0, 1]$ extremal index function.

The idea is to divide the positive real line into blocks of length h. By taking local suprema of the process over these blocks the natural dependence of the continuous-time process is weakened, in certain cases it even disappears. However, for fixed h the extremal index function is a measure for the expected cluster sizes among these blocks. For an extended discussion on the extremal index in the context of discrete- and continuous-time processes see pp. 83.

Corollary 3.2.8 (Extremal index function)
(a) If $L_1 \in \mathcal{S} \cap \mathrm{MDA}(\Phi_\alpha)$, then $\theta(h) = h\alpha/(h\alpha + 1)$ for $h > 0$.

(b) If $L_1 \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$, then $\theta(h) = 1$ for $h > 0$.

Regularly varying Lévy-OU processes exhibit clusters among blocks, since $\theta(h) < 1$. For increasing h the cluster probabilities tends to 0. So they have the potential to model both volatility features: heavy tails and high level clusters. This is in contrast to Lévy-OU processes in $\mathcal{S} \cap \mathrm{MDA}(\Lambda)$, where no clusters occur.

3.2.2 Lévy-OU processes with exponential tails

In this section we investigate Lévy-OU models having exponential tails, hence are lighter tailed than those considered in the previous section. More precisely, we will concentrate on two classes of models in $\mathcal{L}(\gamma)$, $\gamma > 0$; see Definition A.1.2. The first class concerns the *class of convolution equivalent distributions* $\mathcal{S}(\gamma)$, $\gamma > 0$ (Definition A.1.3). Here Theorem A.1.4 provides the necessary relationship between the tails of the infinitely divisible d. f. and of its Lévy measure, which leads to a comparison between the distribution tail of the stationary r. v. V_0 and the increment L_1 of the BDLP. An important family in $\mathcal{S}(\gamma)$ are d. f. s with tail

$$\overline{F}(x) \sim x^{-\beta} l(x) \mathrm{e}^{-\gamma x - cx^p} \quad \text{for } x \to \infty,$$

where γ, $c \geq 0$, $p < 1$, $l(\cdot)$ is normalized slowly varying, and if $c = 0$, $\beta > 1$ or $\beta = 1$ and $\int_1^\infty l(x)/x\, dx < \infty$, are in $\mathcal{S}(\gamma)$ (Klüppelberg [83], Theorem 2.1, or Pakes [115], Lemma 2.3). The *generalized inverse Gaussian distribution* (GIGD) with probability density

$$p(x) = Kx^{\beta-1}\exp\left(-\left(\delta^2 x^{-1} + \gamma^2 x\right)/2\right) \qquad \text{for } x > 0\,,$$

where K is a normalizing constant, $\beta < 0$ and $\delta^2 > 0$, is a prominent example in $\mathcal{S}\left(\gamma^2/2\right)$. Further examples for distributions in $\mathcal{S}(\gamma)$ can be found e. g. in Cline [45].

The second class of processes with exponential tails, which we investigate in this section, are Γ-*OU processes*. These are defined as stationary Lévy-OU processes, where V_0 is $\Gamma(\mu,\gamma)$ distributed with probability density as defined in (3.1.5) for $\mu > 1$ and $\gamma > 0$. The gamma distribution is infinitely divisible with absolutely continuous Lévy measure given by its density

$$\nu_V(dx) = \mu x^{-1}\mathrm{e}^{-\gamma x}dx \qquad \text{for } x > 0\,.$$

Hence, by (3.2.3) the BDLP L has Lévy density

$$\nu(dx) = \mu\gamma\mathrm{e}^{-\gamma x}dx \qquad \text{for } x > 0\,.$$

Except for the factor μ this is the probability density of an exponential d. f.. Hence L is a positive compound Poisson process with Poisson rate $\mu > 0$ and exponential jumps; for details see Barndorff-Nielsen and Shephard [12]. The exponential and gamma laws with scale parameter $\gamma > 0$ belong to $\mathcal{L}(\gamma)$ but not to $\mathcal{S}(\gamma)$.

In analogy to the Γ-OU process, also for $\mathcal{S}(\gamma)$-OU processes with $\gamma > 0$ we restrict our attention to positive compound Poisson processes as BDLPs; i. e. we work in the framework of positive Poisson shot noise processes as defined in Example 3.2.1. Note that all d. f. s in $\mathcal{L}(\gamma)$ for $\gamma > 0$ belong to MDA(Λ).

Some of the results in this section can be found in Albin [1], who studies the extremal behavior for a larger class of Lévy-OU-processes by purely analytic methods.

For BDLPs in $\mathcal{S}(\gamma)$ for $\gamma > 0$ the relation of the tail of the stationary d. f. and its Lévy measure are stated in the following proposition.

Proposition 3.2.9 (Tail behavior of $\mathcal{S}(\gamma)$-OU models for $\gamma > 0$)
Let V be a stationary version of the Lévy-OU process given by (3.2.1).

(a) Suppose $\nu\left(1,\cdot\right]/\nu(1,\infty) \in \mathcal{L}(\gamma)$, $\gamma > 0$. Then $\nu_V\left(1,\cdot\right]/\nu_V(1,\infty) \in \mathcal{L}(\gamma)$ with

$$\nu_V(x,\infty) \sim \frac{1}{\gamma x}\nu(x,\infty) \qquad \text{for } x \to \infty\,.$$

(b) Suppose $L_1 \in \mathcal{S}(\gamma)$, $\gamma > 0$. Then also $V_0 \in \mathcal{S}(\gamma)$, and

$$\mathbb{P}(V_0 > x) \sim \frac{\mathbb{E}\mathrm{e}^{\gamma V_0}}{\mathbb{E}\mathrm{e}^{\gamma L_1}} \frac{1}{\gamma x} \mathbb{P}(L_1 > x) \qquad \text{for } x \to \infty\,.$$

In particular, $\mathbb{P}(V_0 > x) = o(\mathbb{P}(L_1 > x))$ as $x \to \infty$.

Proof.
(a) By (A.1.1) the Lévy measure ν has representation

$$\nu(x, \infty) = c(x) \exp\left[-\int_1^x \frac{1}{a(y)}\, dy\right] \qquad \text{for } x \geq 1, \tag{3.2.17}$$

for functions $a, c : [1, \infty) \to \mathbb{R}_+$ with $\lim_{x\to\infty} c(x) = c > 0$, $\lim_{x\to\infty} a(x) = 1/\gamma$, and $\lim_{x\to\infty} a'(x) = 0$. Since we are only interested in the tail behavior we may assume without loss of generality that ν is absolutely continuous and $c(\cdot) \equiv c$. Recall from (3.2.3) that $\nu_V(dx) = x^{-1}\nu\,(x, \infty)\, dx$ and let $\nu(dx) = \nu'(x)\, dx$. Part *(a)* follows by an application of l'Hospital's rule, since

$$\frac{\nu_V(x, \infty)}{\nu(x, \infty)/(\gamma x)} \sim \frac{\nu(x, \infty)/x}{[\nu'(x)x + \nu(x, \infty)]/(\gamma x^2)} = \gamma \left[\frac{1}{a(x)} + \frac{1}{x}\right]^{-1} \stackrel{x\to\infty}{\longrightarrow} 1.$$

(b) We first show that $V_0 \in \mathcal{S}(\gamma)$. By Theorem A.1.4 *(i)* it suffices to show that $\nu_V\,(1, \cdot]\,/\nu(1, \infty) \in \mathcal{S}(\gamma)$. Again, we can assume without loss of generality that ν is absolutely continuous and has the representation (3.2.17) with constant $c(\cdot) \equiv c$. For simplicity, we further assume that $c = 1$ and $\nu(1, \infty) = 1$; the general case follows by a simple dilation. As in part *(a)* we use that $\nu_V(dx) = x^{-1}\nu\,(x, \infty)\; dx$. An application of l'Hospital's rule shows that

$$\frac{\nu_V(x - y, \infty)}{\nu_V(x, \infty)} \sim \frac{\nu(x - y, \infty)\; x}{\nu(x, \infty)\; (x - y)} \to \mathrm{e}^{\gamma y} \qquad \text{for } x \to \infty,$$

implying $\nu_V\,(1, \cdot] \in \mathcal{L}(\gamma)$. Denote by ν_V^{2*} the convolution of ν_V restricted to $(1, \infty)$ with itself. Then for $1 < y' < x/2$ we use the usual decomposition of the convolution integral

$$\begin{aligned} \frac{\nu_V^{2*}(dx)}{dx} &= \int_1^x \frac{\nu\,(u, \infty)}{u} \frac{\nu\,(x - u, \infty)}{x - u}\, du \\ &= 2\int_1^{y'} \frac{\nu\,(u, \infty)}{u} \frac{\nu\,(x - u, \infty)}{x - u}\, du + \int_{y'}^{x - y'} \frac{\nu\,(u, \infty)}{u} \frac{\nu\,(x - u, \infty)}{x - u}\, du\,. \end{aligned} \tag{3.2.18}$$

In order to show that $\nu_V\,(1, \cdot] \in \mathcal{S}(\gamma)$, we calculate the limit ratio of the densities of ν_V^{2*} and ν_V. Observe that on every compact set $\nu\,(x - u, \infty)\,/\nu\,(x, \infty)$ converges uniformly in u to $\exp(\gamma u)$ as $x \to \infty$. For the first summand of (3.2.18) we thus obtain

$$\begin{aligned} \lim_{x\to\infty} 2\int_1^{y'} \frac{x}{x-u}\frac{\nu\,(x-u,\infty)}{\nu\,(x,\infty)}\frac{\nu\,(u,\infty)}{u}\,du &= 2\int_1^{y'} \mathrm{e}^{\gamma u}\frac{\nu\,(u,\infty)}{u}\,du \\ &= 2\int_1^{y'} \mathrm{e}^{\gamma u}\,\nu_V(du) \,<\, \infty\,. \end{aligned} \tag{3.2.19}$$

For the second summand in (3.2.18) we estimate

$$\begin{aligned} &\int_{y'}^{x-y'} \frac{x\nu\,(u,\infty)\,\nu\,(x-u,\infty)}{u(x-u)\,\nu\,(x,\infty)}\,du \\ &\qquad \le \frac{x}{y'(x-y')}\int_{y'}^{x-y'} \frac{\nu\,(x-u,\infty)}{\nu\,(x,\infty)}\,\nu\,(u,\infty)\;du. \end{aligned} \tag{3.2.20}$$

Furthermore, since

$$\nu\,(x,\infty)\,/\nu'(x) = a(x) \;\longrightarrow\; 1/\gamma \qquad \text{for } x\to\infty\,,$$

there exist constants $K, x_0 > 0$ such that $\nu\,(x,\infty) \le K\nu'(x)$ for $x \ge x_0$. We obtain for $y' > x_0$

$$\begin{aligned} &\frac{x}{y'(x-y')}\int_{y'}^{x-y'} \frac{\nu\,(x-u,\infty)\,\nu\,(u,\infty)}{\nu\,(x,\infty)}\,du \\ &\qquad \le \frac{Kx}{y'(x-y')}\int_{y'}^{x-y'} \frac{\nu\,(x-u,\infty)}{\nu\,(x,\infty)}\,\nu(du). \end{aligned} \tag{3.2.21}$$

Since $\nu\,(1,\cdot] \in \mathcal{S}(\gamma)$, the same decomposition as in (3.2.18) yields (for details see e. g. Pakes [115], Lemma 5.5)

$$\lim_{y'\to\infty}\lim_{x\to\infty}\int_{y'}^{x-y'} \frac{\nu\,(x-u,\infty)}{\nu\,(x,\infty)}\,\nu(du) = 0. \tag{3.2.22}$$

Furthermore, $\lim_{y'\to\infty}\lim_{x\to\infty} x/[y'(x-y')] = \lim_{y'\to\infty} 1/y' = 0$. By (3.2.18)-(3.2.22) we now obtain $\nu_V^{2*}(dx) \sim (2\int_1^\infty \mathrm{e}^{\gamma u}\,\nu(du))\,\nu_V(dx)$ for $x\to\infty$, showing that $\nu_V\,(1,\cdot]$ and hence V_0 are in $\mathcal{S}(\gamma)$. The assertion on the tail behavior now follows from *(a)* and Theorem A.1.4 *(i)*. □

The following result is an analogon to Theorem 3.2.4 and describes the extremal behavior of V completely.

Theorem 3.2.10 (Point process of exceedances of exponential models)
Let V be a stationary version of the Lévy-OU process given by (3.2.1) with L a positive compound Poisson process as in (3.2.4). *Denote by $(\Gamma_k)_{k\in\mathbb{N}}$ the jump times of the positive compound Poisson process L given by (3.2.4) and define $I_k = \frac{1}{\lambda}[\Gamma_k, \Gamma_{k+1})$ for $k \in \mathbb{N}$. Let $\sum_{k=1}^{\infty} \varepsilon\{s_k, P_k\}$ be* PRM(ϑ) *with mean measure $\vartheta(dt \times dx) = dt \times \mathrm{e}^{-x}\, dx$.*

(a) Assume $L_1 \in \mathcal{S}(\gamma)$, $\gamma > 0$, with norming constants $a_T > 0$, $b_T \in \mathbb{R}$ such that

$$\lim_{T\to\infty} T\mathbb{P}(L_1 > a_T x + b_T) = \frac{\mathbb{E}\mathrm{e}^{\gamma L_1}}{\mathbb{E}\mathrm{e}^{\gamma V_0}} \exp(-x) \quad \text{for } x \in \mathbb{R}\,. \tag{3.2.23}$$

Then

$$\sum_{k=0}^{\infty} \varepsilon\left\{\frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1}(\sup_{s\in I_k} V_s - b_{\lambda T})\right\} \overset{T\to\infty}{\Longrightarrow} \sum_{k=1}^{\infty} \varepsilon\left\{s_k, P_k\right\}. \tag{3.2.24}$$

(b) Assume V is the $\Gamma(\mu,\gamma)$-OU process. Let $a_T > 0, b_T \in \mathbb{R}$ be the norming constants of a $\Gamma(\mu+1,\gamma)$ distributed r. v. W, such that

$$\lim_{T\to\infty} T\mathbb{P}(W > a_T x + b_T) = \mu^{-1} \exp(-x) \quad \text{for } x \in \mathbb{R}. \tag{3.2.25}$$

Then

$$\sum_{k=0}^{\infty} \varepsilon\left\{\frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1}(\sup_{s\in I_k} V_s - b_{\lambda T})\right\} \overset{T\to\infty}{\Longrightarrow} \sum_{k=1}^{\infty} \varepsilon\left\{s_k, P_k\right\}. \tag{3.2.26}$$

The proof is divided into several steps. We shall utilize classical results for extreme value theory of stationary discrete-time processes. As a discrete-time skeleton $(V_{\Gamma_k/\lambda})_{k\in\mathbb{Z}}$ seems to be a good candidate. However, $V_{\Gamma_k/\lambda} = \sum_{j=-\infty, j\neq 0}^{k} \mathrm{e}^{-(\Gamma_k-\Gamma_j)}\xi_j$, $k \in \mathbb{N}$, is not stationary. As we will show in Lemma 3.2.11 the process

$$\widetilde{V}_k = \sum_{j=-\infty}^{k} \mathrm{e}^{-(\Gamma_k-\Gamma_j)}\xi_j = V_{\Gamma_k/\lambda} + \mathrm{e}^{-\Gamma_k}\xi_0 \quad \text{for } k \in \mathbb{N}\,, \tag{3.2.27}$$

is stationary, where $\Gamma_0 := 0$. For increasing k the process $\mathrm{e}^{-\Gamma_k}\xi_0$ tends to 0. Thus it has no influence on the extremal behavior. We shall show that the point process behavior is the same for $(V_{\Gamma_k/\lambda})_{k\in\mathbb{N}}$ and for $(\widetilde{V}_k)_{k\in\mathbb{N}}$. For the proof we also need that the D and D' conditions hold for $(\widetilde{V}_k)_{k\in\mathbb{N}}$. The highly technical Lemma A.3.2, where this is confirmed, is postponed to the Appendix.

Lemma 3.2.11
Let V be a stationary version of the Lévy-OU process given by (3.2.1) with L a positive compound Poisson process as in (3.2.4). Then $(\widetilde{V}_k)_{k\in\mathbb{Z}}$ as defined in (3.2.27) is stationary.

Proof.

Let $h \in \mathbb{R}$ be fixed. Note that $(\Gamma_{h+j} - \Gamma_h)_{j\in\mathbb{Z}} \stackrel{d}{=} (\Gamma_j)_{j\in\mathbb{Z}}$. Then

$$\begin{aligned}\widetilde{V}_{k+h} &= \sum_{j=-\infty}^{k+h} \mathrm{e}^{-(\Gamma_{k+h}-\Gamma_j)}\xi_j = \sum_{j=-\infty}^{k+h} \mathrm{e}^{-(\Gamma_{k+h}-\Gamma_h-(\Gamma_j-\Gamma_h))}\xi_j \stackrel{d}{=} \sum_{j=-\infty}^{k+h} \mathrm{e}^{-(\Gamma_k-\Gamma_{j-h})}\xi_j \\ &= \sum_{j=-\infty}^{k} \mathrm{e}^{-(\Gamma_k-\Gamma_j)}\xi_{j+h} \stackrel{d}{=} \sum_{j=-\infty}^{k} \mathrm{e}^{-(\Gamma_k-\Gamma_j)}\xi_j = \widetilde{V}_k.\end{aligned}$$

Similarly, for $l \in \mathbb{N}$ we obtain $(\widetilde{V}_{k_1+h},\ldots,\widetilde{V}_{k_l+h}) \stackrel{d}{=} (\widetilde{V}_{k_1},\ldots,\widetilde{V}_{k_l})$, $k_1,\ldots,k_l \in \mathbb{N}$. $\square$

Proof of Theorem 3.2.10.

Since V is decreasing between jumps, it follows that $\sup_{s\in I_k} V_s = V_{\Gamma_k/\lambda}$. Recall that $\widetilde{V}_k = V_{\Gamma_k/\lambda} + \mathrm{e}^{-\Gamma_k}\xi_0 \stackrel{d}{=} V_0 + \xi_1$ and that $(\widetilde{V}_k)_{k\in\mathbb{N}}$ is stationary. We show first that the norming constants $a_n > 0, b_n \in \mathbb{R}$ given by (3.2.23) and (3.2.25) satisfy

$$\lim_{n\to\infty} n\mathbb{P}(\widetilde{V}_k > a_n x + b_n) = \mu^{-1}\exp(-x) \qquad \text{for } x \in \mathbb{R}. \tag{3.2.28}$$

To show this in case *(a)*, observe that $\mathbb{P}(V_0 > x) = o(\mathbb{P}(\xi_1 > x))$ for $x \to \infty$ by Proposition 3.2.9 *(b)*, so that Theorem A.1.4 *(i,ii)* yields

$$\mathbb{P}(\widetilde{V}_k > x) \sim \mathbb{E}\mathrm{e}^{\gamma V_0}\mathbb{P}(\xi_1 > x) \sim \mathbb{E}\mathrm{e}^{\gamma V_0}[\mathbb{E}\mathrm{e}^{\gamma L_1}]^{-1}\mu^{-1}\mathbb{P}(L_1 > x) \quad \text{for } x \to \infty. \tag{3.2.29}$$

From this (3.2.28) follows immediately, and further we see that $\widetilde{V}_k \in \mathcal{S}(\gamma)$.

In case *(b)*, $\widetilde{V}_k$ is $\Gamma(\mu+1,\gamma)$ distributed as an independent sum of a $\Gamma(\mu,\gamma)$ and an $\mathrm{Exp}(\gamma)$ r. v., and (3.2.28) is immediate. The norming constants of a Γ distribution can be found in Table 3.4.4 of Embrechts et al. [60].

Note that in both cases *(a)* and *(b)*, we have $\widetilde{V}_k \in \mathcal{L}(\gamma)$. Thus, by Lemma A.3.2 and Leadbetter et al. [95], Theorem 5.5.1,

$$\sum_{k=0}^{\infty} \varepsilon\left\{k/(\mu n), a_n^{-1}(\widetilde{V}_k - b_n)\right\} \stackrel{n\to\infty}{\Longrightarrow} \sum_{k=1}^{\infty} \varepsilon\left\{s_k, P_k\right\}. \tag{3.2.30}$$

Define the point processes

$$\widetilde{\kappa}_n = \sum_{k=0}^{\infty} \varepsilon\left\{k/(\mu n), a_n^{-1}(\widetilde{V}_k - b_n)\right\} \quad \text{and} \quad \kappa_n = \sum_{k=0}^{\infty} \varepsilon\left\{k/(\mu n), a_n^{-1}(V_{\Gamma_k/\lambda} - b_n)\right\}.$$

For $\epsilon > 0$ and $I = [s,t) \times (x,\infty) \subseteq \mathbb{R}_+ \times \mathbb{R}$ define $I_\epsilon = [s,t) \times (x, x+\epsilon]$. Taking into account that $V_{\Gamma_k/\lambda} \leq \widetilde{V}_k$ we have for $\delta \in (0,1)$

$$\begin{aligned}
&\mathbb{P}(\kappa_n(I) \neq \widetilde{\kappa}_n(I)) \\
&\quad \leq \mathbb{P}(\widetilde{\kappa}_n(I_\epsilon) > 0) + \sum_{k \in [sn\mu, tn\mu)} \mathbb{P}(\widetilde{V}_k > u_n + \epsilon\, a_n, V_{\Gamma_k/\lambda} \leq u_n) \\
&\quad \leq \mathbb{P}(\widetilde{\kappa}_n(I_\epsilon) > 0) + \sum_{k \in \left[0, n^\delta t\mu\right)} \mathbb{P}(\widetilde{V}_k > u_n + \epsilon\, a_n) + \sum_{k \in \left[n^\delta t\mu, nt\mu\right)} \mathbb{P}(\mathrm{e}^{-\Gamma_k}\xi_0 > \epsilon\, a_n).
\end{aligned}$$

We shall show below that

$$\lim_{n\to\infty} \mathbb{P}(\kappa_n(I) \neq \widetilde{\kappa}_n(I)) = 0. \tag{3.2.31}$$

Then by Rootzén [131], Lemma 3.3, the limit behavior of $\widetilde{\kappa}_n$ and κ_n is the same. Relation (3.2.26) then follows by transforming the time scale as in Lemma 1.2.4.

To show (3.2.31), observe that by (3.2.30) we have

$$\lim_{n\to\infty} \mathbb{P}(\widetilde{\kappa}_n(I_\epsilon) > 0) = 1 - \exp[(t-s)(\exp(-x) - \exp(-(x+\epsilon)))] \xrightarrow{\epsilon\downarrow 0} 0.$$

Furthermore, since $\delta < 1$, we have by (3.2.28)

$$\lim_{n\to\infty} \sum_{k \in \left[0, n^\delta t\mu\right)} \mathbb{P}(\widetilde{V}_k > u_n + \epsilon\, a_n) \leq \lim_{n\to\infty} n^\delta t\mu \mathbb{P}(\widetilde{V}_k > a_n(x+\epsilon) + b_n) = 0.$$

Applying (A.3.4) we obtain

$$\begin{aligned}
&\sum_{k \in \left[n^\delta t\mu, nt\mu\right)} \mathbb{P}(\mathrm{e}^{-\Gamma_k}\xi_0 > \epsilon\, a_n) \\
&\quad \leq \sum_{k \in \left[n^\delta t\mu, nt\mu\right)} \Big(\mathbb{P}\big(\mathrm{e}^{-\Gamma_k}\xi_0 > \epsilon\, a_n, \Gamma_k \geq k/(2\mu)\big) + \mathbb{P}\big(\mathrm{e}^{-\Gamma_k}\xi_0 > \epsilon\, a_n, \Gamma_k < k/(2\mu)\big)\Big) \\
&\quad \leq \sum_{k \in \left[n^\delta t\mu, nt\mu\right)} \mathbb{P}(\mathrm{e}^{-k/(2\mu)}\xi_0 > \epsilon\, a_n) + \sum_{k \in \left[n^\delta t\mu, nt\mu\right)} K/k^2.
\end{aligned}$$

The last summand tends to 0 as $n \to \infty$, since $\sum_{k=1}^\infty 1/k^2 < \infty$. Moreover, there exists an $n_0 \in \mathbb{N}$ such that $a_n \geq 1/(2\gamma)$ and $k\mathrm{e}^{-k/(2\mu)} \leq 1/2$ for $n, k \geq n_0$. Then the first exponential moment of $\gamma k \mathrm{e}^{-k/(2\mu)}\xi_0$ exists, and for $n^\delta t\mu, n \geq n_0$ we obtain

$$\begin{aligned}
\sum_{k \in \left[n^\delta t\mu, nt\mu\right)} \mathbb{P}(\mathrm{e}^{-k/(2\mu)}\xi_0 > \epsilon\, a_n) &\leq \sum_{k \in \left[n^\delta t\mu, nt\mu\right)} \mathbb{E}[\exp(\gamma k \mathrm{e}^{-k/(2\mu)})\xi_0]\mathrm{e}^{-k\gamma\epsilon\, a_n} \\
&\leq \mathbb{E}[\exp(\gamma\xi_0/2)] \sum_{k=\lfloor n^\delta t\mu\rfloor}^{\infty} \mathrm{e}^{-k\epsilon/2} \xrightarrow{n\to\infty} 0,
\end{aligned}$$

since $\sum_{k=1}^\infty \mathrm{e}^{-k\varepsilon/2} < \infty$. This shows (3.2.31). $\square$

Results (3.2.24) and (3.2.26) show that local extremes of such exponential models have no cluster behavior on high levels. The following two corollaries are immediate from Theorem 3.2.10.

Corollary 3.2.12 (Point process of local maxima)
Let the assumptions of Theorem 3.2.10 hold. Denote by $(j_k)_{k\in\mathbb{N}}$ *the jump times of a Poisson process with intensity* e^{-x} *for fixed* $x \in \mathbb{R}$. *Then*

$$\sum_{k=1}^{\infty} \varepsilon\left\{\frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1}(\sup_{s\in I_k} V_s - b_{\lambda T})\right\}(\cdot \times (x,\infty)) \stackrel{T\to\infty}{\Longrightarrow} \sum_{k=1}^{\infty} \varepsilon\{j_k\}\,.$$

Corollary 3.2.13 (Running maxima)
Let V *be a stationary version of the Lévy-OU process (3.2.1), where* L *is a positive, compound Poisson process as in* (3.2.4). *Define* $M(T) = \sup_{0\le t\le T} V_t$ *for* $T > 0$.

(a) Assume $L_1 \in \mathcal{S}(\gamma)$, $\gamma > 0$, *with norming constants given by (3.2.23). Then*

$$\lim_{T\to\infty} \mathbb{P}(a_{\lambda T}^{-1}(M(T) - b_{\lambda T}) \le x) = \exp(-\mathrm{e}^{-x}) \quad \text{for } x \in \mathbb{R}.$$

(b) Assume V *is the* $\Gamma(\mu,\gamma)$*-OU process with norming constants given by (3.2.25). Then*

$$\lim_{T\to\infty} \mathbb{P}(a_{\lambda T}^{-1}(M(T) - b_{\lambda T}) \le x) = \exp(-\mathrm{e}^{-x}) \quad \text{for } x \in \mathbb{R}. \tag{3.2.32}$$

For a subexponential Lévy-OU process the r. v. $M(h) = \sup_{0\le t\le h} V_t$, $h > 0$ be fixed, is tail-equivalent to the increment of the Lévy process; cf. (3.2.8) and (3.2.11). In the class $\mathcal{S}(\gamma)$, $\gamma > 0$, this is much more involved; see Braverman and Samorodnitsky [33]. Although the large jumps of the Lévy process determine the tail behavior, small jumps also have a non-negligible influence. For any $h > 0$, the tail of $M(h)$ is of the same order of magnitude as the tail of the increment of the BDLP, but in general it is only possible to give upper and lower bounds on the asymptotic ratio of the two tails. Using Corollary 3.2.13 one can calculate this constant for Lévy-OU processes explicitly.

Corollary 3.2.14 (Extremal index function)
Let V *be a stationary version of the Lévy-OU process given by (3.2.1), where* L *is a positive compound Poisson process as in* (3.2.4). *Define* $M(h) = \sup_{0\le t\le h} V_t$ *for* $h > 0$.

(a) Let $L_1 \in \mathcal{S}(\gamma)$, $\gamma > 0$. *Then* $M(h) \in \mathcal{L}(\gamma)$ *if and only if*

$$\mathbb{P}(M(h) > x) \sim \lambda h\, \mathbb{E}\mathrm{e}^{\gamma V_0}[\mathbb{E}\mathrm{e}^{\gamma L_1}]^{-1}\mathbb{P}(L_1 > x) \quad \text{for } x \to \infty\,. \tag{3.2.33}$$

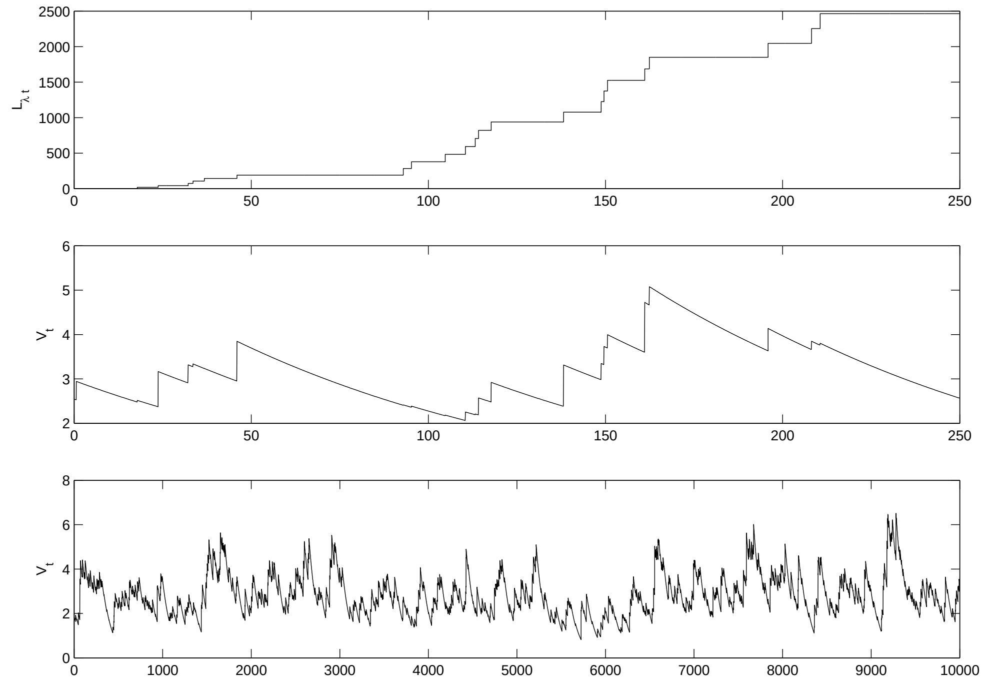


Figure 3.4: Sample path of a Γ-OU process with $\gamma = 3$, $\mu = 8.5$ and $\lambda = 0.01$.

In that case $M(h) \in \mathcal{S}(\gamma)$ *and* $\theta(\cdot) \equiv 1$.

(b) Let V *be the* $\Gamma(\mu,\gamma)$*-OU process with norming constants given by (3.2.25) and let* W *be a* $\Gamma(\mu+1,\gamma)$ *r. v.. Then* $\theta(\cdot) \equiv 1$ *and*

$$\mathbb{P}(M(h) > x) \sim \lambda h \mu \mathbb{P}(W > x) \qquad \text{for } x \to \infty\,. \tag{3.2.34}$$

Proof.

(a) First we assume $M(h) \in \mathcal{L}(\gamma)$. Let $\widetilde{a}_n > 0$, $\widetilde{b}_n \in \mathbb{R}$ and $\widetilde{u}_n = \widetilde{a}_n x + \widetilde{b}_n$ be constants such that

$$\lim_{n\to\infty} n\mathbb{P}(M(h) > \widetilde{u}_n) = \exp(-x).$$

Denote by $\widetilde{M_k}$ an i. i. d. sequence of copies of $M(h)$. Then we obtain by Lemma A.3.2 *(b)* and Leadbetter et al. [95], Theorem 3.5.1, for $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{P}(\widetilde{a}_n^{-1}(M(nh) - \widetilde{b}_n) \le x) = \lim_{n\to\infty} \mathbb{P}(\widetilde{a}_n^{-1}(\max_{k=1,\dots,n} \widetilde{M_k} - \widetilde{b}_n) \le x) = \exp(-\mathrm{e}^{-x}),$$

showing in particular that $\theta(h) = 1$. On the other hand, by Corollary 3.2.13,

$$\lim_{n\to\infty} \mathbb{P}(a_{\lambda nh}^{-1}(M(nh) - b_{\lambda nh}) \le x) = \exp(-\mathrm{e}^{-x}) \qquad \text{for } x \in \mathbb{R}.$$

Then by the convergence to types theorem (see e. g. Theorem A 1.5 of Embrechts et al. [60]), $\widetilde{a}_n/a_{\lambda nh} \stackrel{x\to\infty}{\longrightarrow} 1$ and $\widetilde{b}_n - b_{\lambda nh} \stackrel{x\to\infty}{\longrightarrow} 0$. Applying the convergence to types theorem a second time yields

$$\lim_{n\to\infty} \mathbb{P}(a_{\lambda nh}^{-1}(\max_{k=1,\dots,n} \widetilde{M}_k - b_{\lambda nh}) \le x) = \exp(-\mathrm{e}^{-x}) \qquad \text{for } x \in \mathbb{R}.$$

This implies by Leadbetter et al. [95], Theorem 1.5.1 that

$$\lim_{n\to\infty} n\mathbb{P}(M(h) > u_{\lambda nh}) = \exp(-x),$$

with $u_{\lambda nh} = a_{\lambda nh}x + b_{\lambda nh}$. By (3.2.28) also $\lim_{n\to\infty} n\mathbb{P}(\widetilde{V}_k > u_{\lambda nh}) = \exp(-x)/(\lambda\mu h)$. Hence $\mathbb{P}(M(h) > x) \sim h\lambda\mu\mathbb{P}(\widetilde{V}_k > x)$ for $x \to \infty$, and (3.2.33) follows from (3.2.29).

Conversely, if (3.2.33) holds, then it is clear that $L_1 \in \mathcal{S}(\gamma) \subseteq \mathcal{L}(\gamma)$ implies $M(h) \in \mathcal{L}(\gamma)$ by tail-equivalence. By Lemma A.3.2 *(b)* follows $\theta(h) = 1$.

(b) We refer to Albin [1], Theorem 3, for (3.2.34). That $\theta(h) = 1$ follows then from (3.2.25), (3.2.32) and (3.2.34). □

Remark 3.2.15

For OU processes with BDLP in $\mathcal{S}(\gamma)$, $\gamma > 0$, $\mathcal{S} \cap \mathrm{MDA}(\Lambda)$ and the Γ-OU process

$$\mathbb{P}(M(h) > x) \sim h\lambda\mu\mathbb{P}(\widetilde{V}_k > x) \qquad \text{for } x \to \infty.$$

□

In both cases the extremal index function is equal to one, so that for any $h > 0$ the sequence $M_k = \sup_{(k-1)h\le t\le kh} V_t$ behaves like i. i. d. data. Hence such models cannot explain volatility clusters on high levels.

3.3 Extremal behavior of the COGARCH model

The volatility of the COGARCH process as introduced in (3.0.6) is the (càdlàg) solution to the SDE (3.0.7), which is given by

$$V_t = V_0 + \beta t - \log\eta \int_0^t V_s\,ds + \lambda\eta \sum_{0<s<t} V_s(\Delta L_s)^2 \qquad \text{for } t \ge 0, \tag{3.3.1}$$

given for $\eta > 1$, $\lambda \geq 0$, $\beta > 0$, see Klüppelberg et al. [87,88] for details. An essential feature of the COGARCH model is that the same Lévy process drives the price process S and the volatility process V. Denote by ν the Lévy measure of L. There exists a stationary version of the volatility process V (i. e. V_0 independent of L can be chosen such that V is stationary) if and only if

$$\int_{\mathbb{R}} \log(1 + \lambda\eta y^2)\, \nu(dy) < \log \eta. \tag{3.3.2}$$

With the auxiliary càdlàg process $(X_t)_{t\geq 0}$ defined in (3.0.7) by

$$X_t = t \log \eta - \sum_{0<s\leq t} \log(1 + \lambda\eta(\Delta L_s)^2) \qquad \text{for } t \geq 0\,, \tag{3.3.3}$$

the stationary volatility process has representation

$$V_t = \left(\beta \int_0^t \mathrm{e}^{X_s} ds + V_0\right) \mathrm{e}^{-X_{t-}} \qquad \text{for } t \geq 0\,, \tag{3.3.4}$$

and $V_0 \stackrel{d}{=} \beta \int_0^\infty \mathrm{e}^{-X_t} dt$, independent of L. The auxiliary process $(X_t)_{t\geq 0}$ itself is a spectrally negative Lévy process of bounded variation with drift $\gamma_X = \log \eta$, no Gaussian component, and Lévy measure ν_X given by

$$\nu_X\,[0,\infty) = 0, \quad \nu_X\,(-\infty, -x] = \nu\left(\{y \in \mathbb{R} \,:\, |y| \geq \sqrt{(\mathrm{e}^x - 1)x/(\lambda\eta)}\}\right) \quad \text{for } x > 0.$$

We work with the Laplace transform $\mathbb{E}\mathrm{e}^{-sX_t} = \mathrm{e}^{t\Psi(s)}$, where the Laplace exponent is

$$\Psi(s) = -s \log \eta + \int_{\mathbb{R}} ((1 + \lambda\eta y^2)^s - 1)\, \nu(dy) \qquad \text{for } s \geq 0\,. \tag{3.3.5}$$

For fixed $s \geq 0$, $\mathbb{E}\mathrm{e}^{-sX_t}$ exists (i. e. is finite) for one and hence all $t > 0$, if and only if the integral appearing in (3.3.5) is finite. This is equivalent to $\mathbb{E}|L_1|^{2s} < \infty$. Further, if there exists some $s > 0$ such that $\Psi(s) \leq 0$, then (3.3.2) holds, and hence a stationary version of the volatility process exists.

The qualitative extreme behavior of this volatility process can be seen in Figures 3.5, 3.6, where the driving Lévy process is a compound Poisson process. As in the case of a Lévy OU process the volatility jumps upwards, whenever the driving Lévy process L jumps and decreases exponentially fast between two jumps.

The next Theorem (cf. Klüppelberg et al. [88], Theorem 6) shows that, under weak conditions on the moments of L, the volatility process has Pareto like tails. Since we shall apply a similar argument in the proof of Theorem 3.3.3, we sketch the idea of the proof.

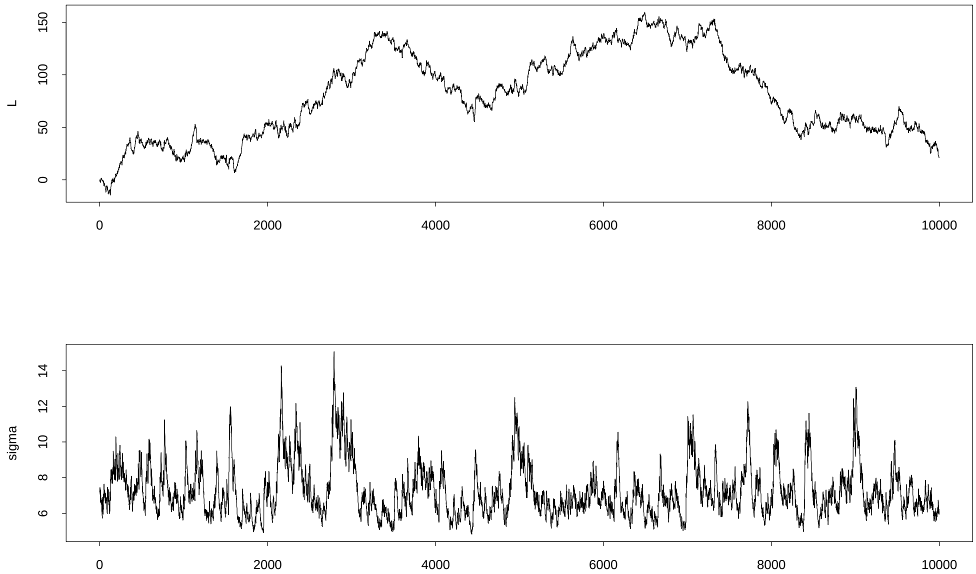


Figure 3.5: Sample path of a compound Poisson driving process with rate 1 and normal jumps with mean 0 and variance 1 (top) and corresponding sample path of the COGARCH process driven by this Lévy process with COGARCH parameters $\beta = 1$, $\lambda = 0.04$ and $\eta = 1.064$ (bottom).

Theorem 3.3.1 (Pareto tail behavior of COGARCH models)
Suppose there exists $\alpha > 0$ such that

$$\mathbb{E}|L_1|^{2\alpha} \log^+ |L_1| < \infty \quad \textit{and} \quad \Psi(\alpha) = 0. \tag{3.3.6}$$

Let V be a stationary version of the volatility process given by (3.3.1)*. Then for some constant $C > 0$ we have*

$$\lim_{x\to\infty} x^\alpha \mathbb{P}(V_0 > x) = C. \tag{3.3.7}$$

Proof.
From (3.3.4) it is seen that the stationary volatility process V satisfies

$$V_t = \mathrm{e}^{-X_{t-}} V_0 + \beta \int_0^t \mathrm{e}^{X_s - X_{t-}}\, ds \qquad \text{for } t > 0,$$

where V_0 is independent of $(\mathrm{e}^{-X_{t-}}, \beta \int_0^t \mathrm{e}^{X_s - X_{t-}}\, ds)_{t\geq 0}$. Thus the stationary solution V_0 satisfies for every $t > 0$ the distributional fix point equation

$$V_0 \stackrel{d}{=} A_t V_0 + B_t,$$

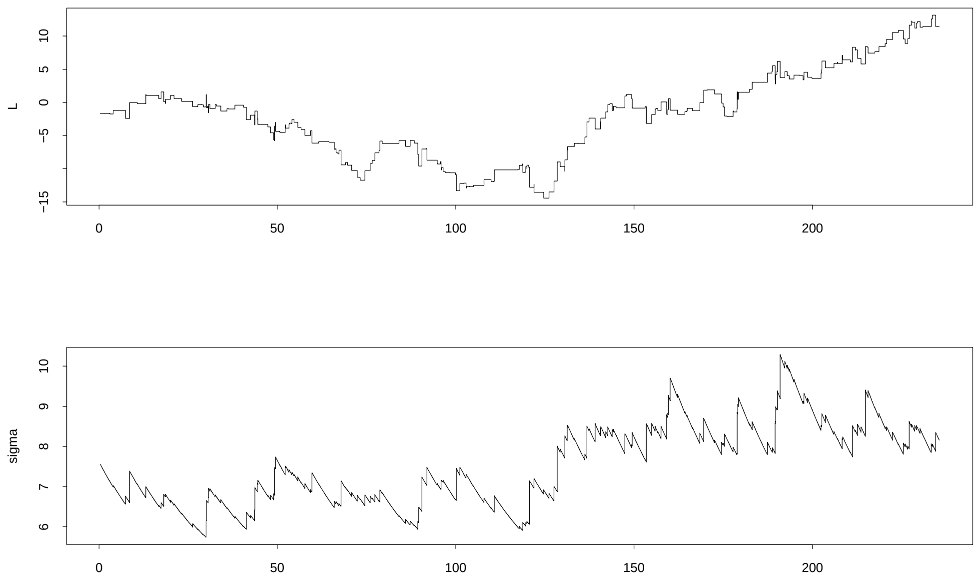


Figure 3.6: First 2500 observations of the sample paths of Figure 3.5 indicating the micro behavior of the COGARCH model.

where V_0 is independent of $(A_t, B_t)_{t\geq 0}$ and

$$A_t \stackrel{d}{=} \mathrm{e}^{-X_{t-}}, \quad B_t \stackrel{d}{=} \beta \int_0^t \mathrm{e}^{X_s - X_{t-}}\, ds.$$

The result now follows from Theorem A.2.1, by choosing t such that (A_t, B_t) satisfies the assumptions. This is possible because of the structure of the process and condition (3.3.6), for details see Klüppelberg et al. [88], Theorem 6. $\square$

The following remark gives a simple sufficient condition for (3.3.6) to hold.

Remark 3.3.2

Let $D := \{d \in [0,\infty) : \mathbb{E}|L_1|^{2d} < \infty\}$ and $d_0 := \sup D \in [0,\infty]$. Suppose $d_0 \notin D$, or that there exists an $s_0 > 0$ such that $0 < \Psi(s_0) < \infty$. Further suppose that $(V_t)_{t\geq 0}$ is strictly stationary. Then (3.3.6) and hence (3.3.7) hold (cf. Klüppelberg et al. [88], Proposition 5). $\square$

We aim at a precise asymptotic description of the COGARCH model above a high threshold like in Section 3.2. It is, however, clear from the definition of V that the influence of the spectrally negative Lévy process X is hard to analyze. In particular, the influence of the small jumps of L to V needs special treatment. We shall restrict ourselves again to the case of a compound Poisson driving process L as given in (3.2.4) by $L_t = \sum_{j=1}^{N_t} \xi_j$ for $t \geq 0$. Here, however, ξ has support on $\mathbb{R}$.

In this case the auxiliary process X simplifies to

$$X_t = t \log \eta - \sum_{k=1}^{N_t} \log(1 + \lambda\eta\xi_k^2) \qquad \text{for } t \geq 0\,, \tag{3.3.8}$$

and the Laplace exponent becomes

$$\Psi(s) = -(s \log \eta + \mu) + \mu \mathbb{E}(1 + \lambda\eta\xi_1^2)^s\,. \tag{3.3.9}$$

In the stationary volatility model we know that $V_t \geq \beta/\log\eta$ a.s. and V jumps if and only if L jumps (cf. Klüppelberg et al. [88], Proposition 2 *(a)*). The jump sizes are positive and depend on the level of the process at that time. As shown in Proposition 2 *(b)* and *(c)* of Klüppelberg et al. [88],

$$V_{\Gamma_k+} - V_{\Gamma_k} = \lambda\eta V_{\Gamma_k}\xi_k^2 \qquad \text{for } k \in \mathbb{N}\,, \tag{3.3.10}$$

and the process decreases exponentially in between jumps:

$$V_t = \frac{\beta}{\log\eta} + \left(V_{\Gamma_k+} - \frac{\beta}{\log\eta}\right) \mathrm{e}^{-(t-\Gamma_k)\log\eta} \qquad \text{for } t \in (\Gamma_k, \Gamma_{k+1}]\,. \tag{3.3.11}$$

As described in Example 3.2.1, the compound Poisson driven volatility process V achieves local suprema only at its right limits at the jump times (and at $t = 0$). This indicates that the discrete-time sequence $(V_{\Gamma_k+})_{k\in\mathbb{N}}$ in combination with the deterministic behavior of V between jumps suffices to describe the extremal behavior of the COGARCH process. Consequently, we investigate the discrete-time skeleton

$$\widetilde{V}_k := V_{\Gamma_k+} \qquad \text{for } k \in \mathbb{N}\,. \tag{3.3.12}$$

Using (3.3.10) and (3.3.11) we obtain

$$\widetilde{V}_{k+1} = \widetilde{V}_k\,(1 + \lambda\eta\xi_{k+1}^2)\mathrm{e}^{-(\Gamma_{k+1}-\Gamma_k)\log\eta} + \frac{\beta}{\log\eta}\left(1 + \lambda\eta\xi_{k+1}^2\right)\left(1 - \mathrm{e}^{-(\Gamma_{k+1}-\Gamma_k)\log\eta}\right)\,,$$

and we see that $(\widetilde{V}_k)_{k\in\mathbb{N}}$ satisfies the stochastic recurrence equation

$$\widetilde{V}_k = \widetilde{A}_k\widetilde{V}_{k-1} + \widetilde{B}_k \qquad \text{for } k \in \mathbb{N}\,, \tag{3.3.13}$$

with $\widetilde{V}_0 = V_0$, where

$$\widetilde{A}_k = (1+\lambda\eta\xi_k^2)\mathrm{e}^{-(\Gamma_k-\Gamma_{k-1})\log\eta}\,, \tag{3.3.14}$$
$$\widetilde{B}_k = \frac{\beta}{\log\eta}(1+\lambda\eta\xi_k^2)\left(1-\mathrm{e}^{-(\Gamma_k-\Gamma_{k-1})\log\eta}\right), \tag{3.3.15}$$

and $((\widetilde{A}_k,\widetilde{B}_k))_{k\in\mathbb{N}}$ is an i. i. d. sequence. It is an interesting observation that by (3.3.14)

$$\log\prod_{j=1}^{k}\widetilde{A}_j = \sum_{j=1}^{k}\log\widetilde{A}_j = -\Gamma_k\log\eta + \sum_{j=1}^{k}\log(1+\lambda\eta\xi_j^2) = -X_{\Gamma_k}\,. \tag{3.3.16}$$

On the other hand, by (3.3.15) and $X_s - X_{\Gamma_k} = \log(1+\lambda\eta\xi_k^2) + (s-\Gamma_k)\log\eta$ for $s\in(\Gamma_k,\Gamma_{k+1})$,

$$\widetilde{B}_k = \beta\int_{\Gamma_{k-1}}^{\Gamma_k}\mathrm{e}^{X_s - X_{\Gamma_k}}\,ds.$$

Denote by $(\widetilde{A},\widetilde{B})$ a copy of $(\widetilde{A}_1,\widetilde{B}_1)$ independent of L. Then it follows that

$$\widetilde{A}\stackrel{d}{=}\widetilde{A}_k\stackrel{d}{=}\mathrm{e}^{-X_{\Gamma_1}}\quad\text{and}\quad\widetilde{B}\stackrel{d}{=}\widetilde{B}_k\stackrel{d}{=}\beta\mathrm{e}^{-X_{\Gamma_1}}\int_0^{\Gamma_1}\mathrm{e}^{X_s}ds\qquad\text{for } k\in\mathbb{N}. \tag{3.3.17}$$

Moreover,

$$\widetilde{V}_k = \widetilde{V}_0\prod_{j=1}^{k}\widetilde{A}_j + \sum_{i=1}^{k}\widetilde{B}_i\prod_{j=i+1}^{k}\widetilde{A}_j = \mathrm{e}^{-X_{\Gamma_k}}\left[\widetilde{V}_0 + \sum_{i=1}^{k}\widetilde{B}_i\mathrm{e}^{X_{\Gamma_i}}\right]\qquad\text{for } k\in\mathbb{N}\,.$$

We are now ready to present the analogue of Theorem 3.3.1 for the sequence $(\widetilde{V}_k)_{k\in\mathbb{N}}$. As can be seen from (3.3.8), the process $(X_{\Gamma_k})_{k\in\mathbb{N}}$ is a random walk with increments

$$X_{\Gamma_k} - X_{\Gamma_{k-1}} = (\Gamma_k-\Gamma_{k-1})\log\eta - \log(1+\lambda\eta\xi_k^2)\qquad\text{for } k\in\mathbb{N}.$$

Theorem 3.3.3 (Pareto tail behavior of $\widetilde{\mathbf{V}}$)
Suppose there exists some $\alpha > 0$ such that

$$\mathbb{E}|L_1|^{2\alpha}\log^+|L_1| < \infty\quad\text{and}\quad\Psi(\alpha) = 0. \tag{3.3.18}$$

Then a stationary solution $(\widetilde{V}_k)_{k\in\mathbb{N}}$ of (3.3.13) exists. Its marginal stationary distribution $\widetilde{V}_\infty\stackrel{d}{=}\widetilde{V}_1$ is the unique solution of the random fix point equation

$$\widetilde{V}_\infty\stackrel{d}{=}\widetilde{A}\widetilde{V}_\infty + \widetilde{B},$$

where $(\widetilde{A}, \widetilde{B})$ *is given by (3.3.17) and is independent of* L*. Furthermore,*

$$\mathbb{P}(\widetilde{V}_\infty > x) \sim \widetilde{C} x^{-\alpha} \qquad \text{for } x \to \infty\,,$$

where

$$\widetilde{C} = \frac{\mathbb{E}\left[(\widetilde{A}\widetilde{V}_\infty + \widetilde{B})^\alpha - (\widetilde{A}\widetilde{V}_\infty)^\alpha\right]}{\alpha \mathbb{E}|\widetilde{A}|^\alpha \log^+ |\widetilde{A}|} > 0\,. \tag{3.3.19}$$

Proof.

We shall show that conditions *(i)*-*(iv)* of Theorem A.2.1 are satisfied: by definition, $\log \widetilde{A} \stackrel{d}{=} -\Gamma_1 \log \eta + \log(1 + \lambda\eta\xi_1^2)$, where Γ_1 is exponentially distributed. Consequently, *(i)* follows.

To show *(ii)* note that by the independence of Γ_1 and ξ_1, for $\alpha > 0$ we have by (3.3.9)

$$\begin{aligned} \mathbb{E}|\widetilde{A}|^\alpha &= \mathbb{E}\mathrm{e}^{-\Gamma_1 \alpha \log \eta} \mathbb{E}(1 + \lambda\eta\xi_1^2)^\alpha \\ &= \frac{\mu}{\mu + \alpha \log \eta} \frac{\mu + \alpha \log \eta + \Psi(\alpha)}{\mu} \\ &= 1 + \frac{1}{\mu + \alpha \log \eta} \Psi(\alpha) \;=\; 1\,, \end{aligned}$$

by the second assumption in (3.3.18).
In order to prove *(iii)* note that

$$\mathbb{E}|\widetilde{A}|^\alpha \log^+ |\widetilde{A}| \;\leq\; \mathbb{E}|1 + \lambda\eta\xi_1^2|^\alpha \log^+ \left(1 + \lambda\eta\xi_1^2\right) < \infty\,,$$

if and only if the first assumption in (3.3.18) holds, see Sato [138], Theorem 25.3.
Finally, *(iv)* follows from

$$\mathbb{E}|\widetilde{B}|^\alpha \leq (\beta/\log \eta)^\alpha \mathbb{E}|1 + \lambda\eta\xi_1^2|^\alpha < \infty\,.$$

That the constant $\widetilde{C}$ is indeed strictly positive follows from the fact that $\widetilde{A}$, $\widetilde{B}$ and $\widetilde{V}_\infty$ are strictly positive, almost surely. □

Remark 3.3.4

(i) X_{Γ_k} tends almost surely to ∞ if and only if $\mathbb{E}X_{\Gamma_1} > 0$ or, equivalently, $\mu\mathbb{E}\log(1+\lambda\eta\xi_1^2) < \log\eta$. Notice that for this model the stationarity condition (3.3.2) is equivalent to $\mathbb{E}X_{\Gamma_1} > 0$.

(ii) In a sense it is remarkable that the tail of the stationary r. v. of the continuous-time model V_∞ and of the discrete-time skeleton $\widetilde{V}_\infty$ are so similar. As the discrete-time skeleton considers only local suprema of the process, one expects $\widetilde{V}_\infty$ to be stochastically larger. As the Pareto index α is the same for both models, any difference can only appear in the constants C and $\widetilde{C}$. Brockwell, Chadraa and Lindner [37] have established a precise relationship between the distributions of V_∞ and $\widetilde{V}_\infty$, showing that

$$\left(V_\infty - \frac{\beta}{\log\eta}\right) \stackrel{d}{=} \mathrm{e}^{-(\log\eta)\Gamma}\left(\widetilde{V}_\infty - \frac{\beta}{\log\eta}\right),$$

where $\Gamma \stackrel{d}{=} \Gamma_1$ is exponentially distributed with parameter μ and independent of $\widetilde{V}_\infty$. Using a result of Breiman [34], it then follows that

$$C = \mathbb{E}\left(\mathrm{e}^{-(\log\eta)\Gamma}\right)^\alpha \widetilde{C} = \frac{\mu}{\mu+\alpha\log\eta}\widetilde{C} = \frac{1}{\mathbb{E}(1+\lambda\eta\xi_1^2)^\alpha}\widetilde{C}, \tag{3.3.20}$$

where the last equation follows from (3.3.9). □

The extremal behavior of solutions of stochastic recurrence equations is studied in de Haan et al. [57]. Their results can be applied to the stationary discrete-time skeleton of the volatility process $(\widetilde{V}_k)_{k\in\mathbb{N}}$ as defined in (3.3.12).

Theorem 3.3.5 (Extremal behavior of the COGARCH model)
Let V be a stationary version of the volatility process given by (3.3.1) *and define $M(T) = \sup_{0\le t\le T} V_t$ for $T>0$. Suppose there exists some $\alpha>0$ such that*

$$\mathbb{E}|L_1|^{2\alpha}\log^+|L_1| < \infty \quad \text{and} \quad \Psi(\alpha) = 0.$$

Let $\widetilde{C}$ be the constant in (3.3.19) *and define $a_T := (\mu T)^{1/\alpha}$ for $T>0$. Then*

$$\lim_{T\to\infty}\mathbb{P}(a_T^{-1}M(T)\le x) = \exp(-\widetilde{C}\theta x^{-\alpha}) \quad \text{for } x>0,$$

where $\theta = 1-\mathbb{E}\left[\sup_{t\ge\Gamma_1}\{\mathrm{e}^{-\alpha X_t^+}\}\right] \in (0,1)$. Denote by $(\Gamma_k)_{k\in\mathbb{N}}$ the jump times of the compound Poisson process L given by (3.2.4) and define $I_k = (\Gamma_k,\Gamma_{k+1}]$ for $k\in\mathbb{N}$. Let $(j_k)_{k\in\mathbb{N}}$ be the jump times of a Poisson process with intensity $\widetilde{C}\theta x^{-\alpha}$. Let $(\zeta_k)_{k\in\mathbb{N}}$ be i. i. d. discrete r. v. s, independent of $(j_k)_{k\in\mathbb{N}}$, with probability distribution

$$\pi_k = \mathbb{P}(\zeta_1 = k) = (\theta_k - \theta_{k+1})/\theta \quad \text{for } k\in\mathbb{N}.$$

Then for $x > 0$,

$$\sum_{k=1}^{\infty} \varepsilon\left\{\frac{\Gamma_k}{T}, a_T^{-1} \sup_{s\in I_k} V_s\right\}(\cdot \times (x,\infty)) \overset{T\to\infty}{\Longrightarrow} \sum_{k=1}^{\infty} \zeta_k \varepsilon\{j_k\}. \tag{3.3.21}$$

Here

$$\begin{aligned}\theta_k &:= \mathbb{E}\left[\exp(\alpha \min\{T_{k-1},0\}) - \exp(\alpha \min\{T_k,0\})\right] \\ &= \int_0^1 \mathbb{P}\left(\mathrm{card}\left\{j\in\mathbb{N} : \mathrm{e}^{-\alpha X_{\Gamma_j}} > y\right\} = k-1\right) dy,\end{aligned}$$

where $\infty = T_0 \geq T_1 \geq \ldots$ *are the ordered values of the sequence* $(-X_{\Gamma_k})_{k\in\mathbb{N}}$ *and* $\theta = \theta_1$.

Proof.
Since $\sup_{s\in I_k} V_s = \widetilde{V}_k$, Theorem 3.3.3 and de Haan et al. [57], Theorem 2.1, show that

$$\sum_{k=1}^{\infty} \varepsilon\left\{\frac{k}{\mu n}, a_n^{-1}\widetilde{V}_k\right\}(\cdot \times (x,\infty)) \overset{T\to\infty}{\Longrightarrow} \sum_{k=1}^{\infty} \zeta_k \varepsilon\{j_k\}$$

and that $(\widetilde{V}_k)_{k\in\mathbb{N}}$ has extremal index $\theta \in (0,1)$, given by

$$\begin{aligned}\theta &= \alpha \int_1^{\infty} \mathbb{P}\left(\bigvee_{j=1}^{\infty}\prod_{k=1}^{j} \widetilde{A}_k \leq y^{-1}\right) y^{-\alpha-1}\, dy \\ &= \alpha \int_1^{\infty} \mathbb{P}\left(\bigvee_{j=1}^{\infty} \exp(-X_{\Gamma_j}) \leq y^{-1}\right) y^{-\alpha-1}\, dy \\ &= \int_0^1 \mathbb{P}\left(\sup_{t\geq\Gamma_1}\{\mathrm{e}^{-\alpha X_t}\} \leq z\right) dz \\ &= 1 - \mathbb{E}\left[\min\left\{1, \sup_{t\geq\Gamma_1}\{\mathrm{e}^{-\alpha X_t}\}\right\}\right].\end{aligned}$$

For the first expression for θ_k, see de Haan et al. [57], and the second expression follows by a similar calculation as above. By an application of Lemma 1.2.4 we transform the time scale, such that (3.3.21) holds. Then we obtain

$$\begin{aligned}\lim_{T\to\infty} \mathbb{P}(a_T^{-1} M(T) \leq x) &= \lim_{T\to\infty} \mathbb{P}\left(\sum_{k=1}^{\infty} \varepsilon\left\{\frac{\Gamma_k}{T}, a_T^{-1}\sup_{s\in I_k} V_s\right\}((0,1)\times(x,\infty)) = 0\right) \\ &= \mathbb{P}\left(\sum_{k=1}^{\infty} \zeta_k \varepsilon\{j_k\}((0,1)) = 0\right) = \exp(-\widetilde{C}\theta x^{-\alpha}).\end{aligned}$$

□

Remark 3.3.6
(i) By the Poisson result (3.3.21) we observe clusters in local extremes of the continuous-time process. So the COGARCH is a suitable model for heavy tailed volatility models with clusters on high levels.

(ii) Note that we have shown in the proof above that θ is the extremal index of $(\widetilde{V}_k)_{k\in\mathbb{N}}$. □

3.4 Conclusion

In this chapter we have investigated the extremal behavior of the most popular continuous-time volatility models. We have concentrated on models with tails ranging from exponential to regularly varying; i. e. tails as they are found in empirical volatilities. The stochastic quantities derived for such models include

- the tail of the stationary volatility V_0 and the relation to the tail of the distribution governing the extreme behavior,
- the asymptotic distribution of the running maxima, i. e. their MDA and the norming constants,
- the cluster behavior of the model on high levels.

We found interesting similarities in the extremal behavior of certain models, which was quite unexpected.

Recall the GCIR model of Example 3.1.3, where the tail of the stationary distribution F of V_0 is compared to the tail of H, the d. f. describing the extreme behavior. Example 3.1.3 *(2)* belongs to $\mathcal{S} \cap \mathrm{MDA}(\Lambda)$, it has stationary distribution with a semi-heavy Weibull like tail. Relation (3.2.10) is mimicked by the fact that (3.1.6) can be rewritten to

$$\overline{F}(x) \sim \frac{a(x)}{x}\overline{H}(x) \qquad \text{for } x \to \infty\,.$$

Moreover, as the norming constants in the GCIR examples are calculated based on the d. f. H, analogously, by the above Corollary 3.2.6, for the Lévy-OU process in $\mathrm{MDA}(\Lambda)$, the norming constants are derived from L_1 and $M(1)$, respectively, and not from the stationary distribution of the process V.

Analogous results hold for Example 3.1.3 *(3)*, which belongs to $\mathcal{S} \cap \mathrm{MDA}(\Phi_\alpha)$ for some $\alpha > 0$. Here the tails of F and H are both regularly varying of the same index; this corresponds to (3.2.8).

Also in the case, where V_0 is gamma distributed, the behavior of the running maxima of the GCIR model in Example 3.1.3 *(1)* and of the Γ-OU process as given in (3.2.25) and (3.2.32), respectively, are identical.

This means also that, if the stationary distribution of a GCIR model coincides with the stationary distribution of a Lévy-OU model, then also the norming constants and the behavior of the running maxima coincide. The role of $M(h)$ for $L_1 \in \mathcal{S}(\gamma)$, $\gamma \geq 0$ corresponds for the GCIR models to the d. f. H; the influence of the driving Brownian motion plays no role whatsoever for the extreme behavior.

Concerning volatility clusters, no OU process in $\mathrm{MDA}(\Lambda)$ presented in this chapter can model such clusters on high levels. Whereas regularly varying Lévy-OU models have the potential to model them. A way to introduce clusters into subexponential models in $\mathrm{MDA}(\Lambda)$ is to replace the the exponentially decreasing kernel function by a kernel function with more than one maxima (Theorem 1.4.1, Corollary 1.4.11).

The COGARCH model resembles the GCIR models only in the sense that heavy tails occur, although the driving process can be very light tailed; the difference being that the COGARCH model always has heavy tails. There is no obvious relationship between the tail behavior of the stationary r. v. V_0 and L_1; the heavy tails occur by the very intrinsic dependence structure of the model.

With respect to volatility clusters, only regularly varying OU processes and COGARCH processes exhibit volatility clusters on high levels, which can be described quite precisely by the distribution of the cluster sizes; see Corollary 3.2.5 and Theorem 3.3.5.

In this chapter we have refrained from discussing another important stylized fact of empirical volatility: it exhibits often long memory in the sense that the autocovariance function decreases very slowly. This phenomenon can have various reasons, as for instance discussed in Mikosch and Stărică [113]. On the other hand, it is an important fact, which should not be completely ignored. All models presented in this chapter have an exponentially decreasing covariance functions, which only exhibit some visual long memory, when the process is close to non-stationarity.

For diffusion models like the GCIR models, a remedy, which introduces long range dependence in such models, is to replace the driving Brownian motion by a fractional Brownian motion. This generates a new class of stationary long memory models. Such models have been suggested and analyzed in [41, 40].

For the OU process the exponentially decreasing covariance function is due to the exponential kernel function; see (3.2.2). The often observed long-range dependence effect in the empirical volatility cannot be modelled this way. There are two ways to introduce long memory into such models. The first one is to replace the exponential kernel function by a hyperbolic kernel function of the form $f(x) \sim |x|^{-\beta}$ for $|x| \to \infty$ and some $\beta \in (0.5, 1)$. This introduces long memory into the model. This can be modelled by the regularly varying Lévy driven MA processes of Chapter 2. The second method has been suggested by Barndorff-Nielsen and Shephard [14]: a superposition of several regularly varying Lévy-OU processes also creates long memory; Example 2.5.8 shows that they also exhibit volatility clusters.

Appendix A

Appendix

A.1 Basic notation and definition

In this Appendix we summarize some definitions and concepts used throughout the thesis.

For details and further references see Embrechts et al. [60].

Definition A.1.1
A positive measurable function $u : \mathbb{R} \to \mathbb{R}_+$ *is regularly varying with index* α*, denoted by* $u \in \mathcal{R}_\alpha$ *for* $\alpha \in \mathbb{R}$*, if*

$$\lim_{t\to\infty} \frac{u(tx)}{u(t)} = x^\alpha \qquad \text{for } x > 0\,.$$

The function u *is said to be slowly varying if* $\alpha = 0$*, and rapidly varying, denoted by* $u \in \mathcal{R}_{-\infty}$*, if the above limit is equal to* 0 *for* $x > 1$ *and to* ∞ *for* $0 < x < 1$.

Definition A.1.2
A d.f. F *belongs to the class* $\mathcal{L}(\gamma)$, $\gamma \geq 0$ *if for every* $y \in \mathbb{R}$,

$$\lim_{x\to\infty} \overline{F}(x-y)/\overline{F}(x) = \mathrm{e}^{\gamma y}\,.$$

The class $\mathcal{L}(\gamma)$ is related to the class $\mathcal{R}_{-\gamma}$ by the fact that

$$F \in \mathcal{L}(\gamma) \quad \text{if and only if} \quad \overline{F} \circ \log \in \mathcal{R}_{-\gamma}.$$

Thus the convergence of $\overline{F}(x-y)/\overline{F}(x)$ in Definition A.1.2 is uniform on compact y-intervals. For an excellent monograph on regular variation we refer to Bingham et al. [29].

Applying Karamata's representation for regularly varying functions to the class $\mathcal{L}(\gamma)$ we obtain for $F \in \mathcal{L}(\gamma)$, $\gamma \geq 0$, the representation

$$\overline{F}(x) = c(x) \exp\left[-\int_0^x \frac{1}{a(y)}\, dy\right] \quad \text{for } x > 0, \tag{A.1.1}$$

where $a, c : \mathbb{R}_+ \to \mathbb{R}_+$ and $\lim_{x\to\infty} c(x) = c > 0$ and a is absolutely continuous with $\lim_{x\to\infty} a(x) = 1/\gamma$ and $\lim_{x\to\infty} a'(x) = 0$.

Definition A.1.3 (Convolution equivalent distributions)
Let $\gamma \geq 0$ and X have d. f. F. We say that F or X belongs to the class $\mathcal{S}(\gamma)$, if the following properties hold.

(i) $F \in \mathcal{L}(\gamma)$,

(ii) $\displaystyle \lim_{x\to\infty} \frac{\overline{F^{*2}}(x)}{\overline{F}(x)} = 2\widehat{f}(\gamma) < \infty\,,$

where $\widehat{f}(\gamma) = \mathbb{E}\mathrm{e}^{\gamma X}$ is the moment generating function of X at γ. The class $\mathcal{S} := \mathcal{S}(0)$ is called the class of subexponential distributions.

Theorem A.1.4

(i) Let F be infinitely divisible with Lévy measure ν and $\gamma \geq 0$. Then

$$F \in \mathcal{S}(\gamma) \quad \Leftrightarrow \quad \nu\,(1,\cdot]\,/\nu(1,\infty) \in \mathcal{S}(\gamma) \quad \Leftrightarrow \quad \lim_{x\to\infty} \overline{F}(x)/\nu(x,\infty) = \widehat{f}(\gamma)\,.$$

(ii) Suppose $F \in \mathcal{S}(\gamma)$, $\lim_{x\to\infty} \overline{F_i}(x)/\overline{F}(x) = q_i \geq 0$ and $\widehat{f_i}(\gamma) < \infty$ for $i = 1, 2$. Then

$$\lim_{x\to\infty} \frac{\overline{F_1 * F_2}(x)}{\overline{F}(x)} = q_1 \widehat{f}_2(\gamma) + q_2 \widehat{f}_1(\gamma).$$

If $q_i > 0$ for some $i \in \{1,2\}$, then also F_i, $F_1 * F_2 \in \mathcal{S}(\gamma)$.

(iii) Let N be a Poisson r. v. with mean μ and $(X_k)_{k\in\mathbb{N}}$ be an i. i. d. sequence with d. f. $F \in \mathcal{S}(\gamma)$. The r. v. $Y = \sum_{k=1}^N X_k$ has d. f. $G = \mathrm{e}^{-\mu} \sum_{n=0}^\infty \frac{\mu^n}{n!} F^{*n}$. Then $G \in \mathcal{S}(\gamma)$ and

$$\overline{G}(x) \sim \mu \widehat{f}(\gamma) \overline{F}(x) \quad \text{for } x \to \infty.$$

The following is the fundamental theorem in extreme value theory.

Theorem A.1.5 (Fisher-Tippett Theorem)
Suppose we can find sequences of real numbers $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$\lim_{n\to\infty} \mathbb{P}(a_n^{-1}(M_n - b_n) \le x) = \lim_{n\to\infty} F^n(a_n x + b_n) = G(x) \quad \text{for } x \in \mathbb{R}\,, \tag{A.1.2}$$

for some non-degenerate d. f. G (we say F is in the maximum domain of attraction of G and write $F \in \mathrm{MDA}(G)$). Then there are $a > 0$, $b \in \mathbb{R}$ such that $x \mapsto G(ax+b)$ is one of the following three extreme value d. f. s:

- *Fréchet* $\Phi_\alpha(x) = \begin{cases} 0, & x \le 0, \\ \exp\left(-x^{-\alpha}\right), & x > 0, \end{cases}$ *for* $\alpha > 0$.
- *Gumbel* $\Lambda(x) = \exp\left(-\mathrm{e}^{-x}\right), \quad x \in \mathbb{R}\,.$
- *Weibull* $\Psi_\alpha(x) = \begin{cases} \exp\left(-\left(-x\right)^{\alpha}\right), & x \le 0, \\ 1, & x > 0, \end{cases}$ *for* $\alpha > 0\,.$

We summarize some well-known facts related to domains of attraction.

Proposition A.1.6

(a) *The following Poisson characterization holds: $F \in \mathrm{MDA}(G)$ if and only if $a_n > 0, b_n \in \mathbb{R}$ exist such that*

$$\lim_{n\to\infty} n\overline{F}(a_n x + b_n) = -\log G(x) \quad \text{for } x \in \mathbb{R}. \tag{A.1.3}$$

(b) *If $F \in \mathcal{L}(\gamma)$ for $\gamma > 0$, then $F \in \mathrm{MDA}(\Lambda)$ with $a_n \to 1/\gamma$ as $n \to \infty$ and $\mathrm{e}^{b_n} \in \mathcal{R}_{1/\gamma}$.*

(c) *If $F \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$, then $b_n \to \infty$, $a_n \to \infty$ and $b_n/a_n \to \infty$ as $n \to \infty$.*

(d) *If $F \in \mathrm{MDA}(\Phi_\alpha) = \mathcal{R}_{-\alpha}$ for $\alpha > 0$, then $b_n = 0$, $a_n \in \mathcal{R}_{1/\alpha}$ and $a_n \to \infty$ as $n \to \infty$.*

The following concept has proved useful in comparing tails.

Definition A.1.7 (Tail-equivalence)
Two d. f. s F and G (or two measures μ and ν) are called tail-equivalent *if both have support unbounded to the right and there exists some $c > 0$ such that*

$$\lim_{x\to\infty} \overline{F}(x)/\overline{G}(x) = c \quad \text{or} \quad \lim_{x\to\infty} \nu(x,\infty)/\mu(x,\infty) = c\,.$$

Important in the context of our chapter is that all the following classes are closed with respect to tail-equivalence: MDA(G) for $G \in \{\Phi_\alpha, \alpha > 0, \Lambda\}$, $\mathcal{R}_{-\alpha}$ for $\alpha \in [0,\infty)$, $\mathcal{L}(\gamma)$ for $\gamma \geq 0$, $\mathcal{S}(\gamma)$ for $\gamma \geq 0$. Moreover, for two tail-equivalent d. f. s in some MDA(G) one can choose the same norming constants.

Definition A.1.8 (Poisson random measure)
Let $(A, \mathcal{A}, \vartheta)$ be a measurable space, where ϑ is σ-finite, and $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A Poisson random measure N *with mean measure ϑ, denoted by* PRM(ϑ), *is a collection of r. v. s $(N(A))_{A\in\mathcal{A}}$ with $N(\emptyset) = 0$, such that:*

(a) *Given any sequence $(A_n)_{n\in\mathbb{N}}$ of mutually disjoint sets in $\mathcal{A}$:*
$$N\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} N(A_n) \quad \text{a.s..}$$
(b) *$N(A)$ is Poisson distributed with mean $\vartheta(A)$ for every $A \in \mathcal{A}$.*
(c) *For mutually disjoint sets $A_1, \dots, A_n \in \mathcal{A}$, $n \in \mathbb{N}$, the r. v. s $N(A_1), \dots, N(A_n)$ are independent.*

Definition A.1.9 (Extremal index)
Let $X = (X_n)_{n\in\mathbb{Z}}$ be a strictly stationary sequence and $\theta \geq 0$. If for every $\tau > 0$ there exists a sequence $u_n(\tau)$ with
$$\lim_{n\to\infty} n\mathbb{P}(X_1 > u_n(\tau)) = \tau \quad \text{and} \quad \lim_{n\to\infty} \mathbb{P}(\max_{k=1,\dots,n} X_k \leq u_n(\tau)) = \mathrm{e}^{-\theta\tau},$$
then θ is called the extremal index of X and has value in $[0,1]$.

Theorem A.1.10 (Potter's Theorem, Bingham et al. [29], Theorem 1.5.6)

(i) *If l is slowly varying then for any chosen constants $A > 1$, $\delta > 0$ there exists a $X(\delta, A)$ such that*
$$l(y)/l(x) \leq A \max\{(y/x)^\delta, (y/x)^{-\delta}\} \quad \text{for } x, y \geq X(\delta, A).$$
(ii) *If, further, l is bounded away from 0 and ∞ on every compact subset of $[0,\infty)$, then for every $\delta > 0$ there exists an $A(\delta) > 1$ such that*
$$l(y)/l(x) \leq A(\delta) \max\{(y/x)^\delta, (y/x)^{-\delta}\} \quad \text{for } x, y > 0.$$
(iii) *If f is regularly varying of index α then for any chosen $A > 1, \delta > 0$ there exists an $X(A, \delta)$ such that*
$$f(y)/f(x) \leq A \max\{(y/x)^{\alpha+\delta}, (y/x)^{\alpha-\delta}\} \quad \text{for } x, y \geq X(\delta, A).$$
(iv) *If, further, f is regularly varying of index α and bounded away from 0 and ∞ on every compact subset of $[0,\infty)$, then for every $\delta > 0$ there exists an $A(\delta) > 1$ such that*
$$f(y)/f(x) \leq A(\delta) \max\{(y/x)^{\alpha+\delta}, (y/x)^{\alpha-\delta}\} \quad \text{for } x, y > 0.$$

A.2 Stationary solution of a random recurrence equation

The following result is the central result for proving stationarity and the tail behavior of a stochastic process defined by a random recurrence equation.

Theorem A.2.1 (Vervaat [142], Goldie [69], Theorem 4.1, Lemma 2.2)
Let $(Y_k)_{k\in\mathbb{N}}$ be a stochastic process defined by $Y_k = A_k Y_{k-1} + B_k$, where $((A_k, B_k))_{k\in\mathbb{N}}$, (A, B) are i. i. d. sequences. Assume that the following conditions are satisfied:

(i) *The law of* $\log|A|$, *given* $|A| \neq 0$, *is not concentrated on a lattice* $-\infty \cap r\mathbb{Z}$ *for any* $r > 0$.

(ii) $\mathbb{E}|A|^\alpha = 1$.

(iii) $\mathbb{E}|A|^\alpha \log^+ |A| < \infty$.

(iv) $\mathbb{E}|B|^\alpha < \infty$.

Then the equation $Y_\infty \stackrel{d}{=} AY_\infty + B$, *where* Y_∞ *is independent of* (A, B), *has the solution unique in distribution*

$$Y_\infty \stackrel{d}{=} \sum_{m=1}^{\infty} B_m \prod_{k=1}^{m} A_k.$$

The process $(Y_k)_{k\in\mathbb{N}}$ *with* $Y_0 \stackrel{d}{=} Y_\infty$ *is stationary and has tails*

$$\mathbb{P}(Y_\infty > x) \sim \frac{\mathbb{E}\left[((AY_\infty + B)^+)^\alpha - ((AY_\infty)^+)^\alpha\right]}{\alpha\mathbb{E}|A|^\alpha \log^+ |A|} x^{-\alpha} \qquad \text{for } x \to \infty\,.$$

A.3 The conditions $D_r(u_n)$ and $D'(u_n)$

Classical results for the extremal behavior of weakly stationary sequences are based on two conditions: the first one is a specific type of asymptotic dependence, and the second is an anti-clustering condition.

Definition A.3.1
Let $X = (X_n)_{n\in\mathbb{N}}$ *be a strictly stationary sequence, such that for* $m = 1, \dots, r$, *the sequences of constants* $(u_n^{(m)})_{n\in\mathbb{N}}$ *and* $(u_n)_{n\in\mathbb{N}}$ *satisfies* $\lim_{n\to\infty} n\overline{F}(u_n^{(m)}) = \tau^{(m)}$ *and* $\lim_{n\to\infty} n\overline{F}(u_n) = \tau$.

(a) *For any integers* p, q *and* n *let*

$$1 \leq i_1 < \cdots < i_p < j_1 < \cdots < j_q \leq n$$

such that $j_1 - i_p \geq l$ *and* $\mathbf{v}_n = (v_n^{(1)}, \dots, v_n^{(p)})$, $\mathbf{w}_n = (w_n^{(1)}, \dots, w_n^{(q)})$ *with* $v_n^{(l)}, w_n^{(s)} \in \{u_n^{(1)}, \dots, u_n^{(r)}\}$. *Write* $I = \{i_1, \dots, i_p\}$, $J = \{j_1, \dots, j_q\}$, $\mathbf{X}_I = (X_{i_1}, \dots, X_{i_p})$ *and* $\mathbf{X}_J = (X_{j_1}, \dots, X_{j_q})$. *If for each choice of indices* I, J

$$|\mathbb{P}(\mathbf{X}_I \leq \mathbf{v}_n, \mathbf{X}_J \leq \mathbf{w}_n) - \mathbb{P}(\mathbf{X}_I \leq \mathbf{v}_n)\, \mathbb{P}(\mathbf{X}_J \leq \mathbf{w}_n)| \leq \alpha_{n,l},$$

where $\alpha_{n,l_n} \to 0$ *as* $n \to \infty$ *for some sequence* $l_n = o(n)$, *then* X *satisfies the condition* $\boldsymbol{D_r(u_n)}$.

(b) X *satisfies the condition* $D'(u_n)$, *if*

$$\lim_{k\to\infty} \limsup_{n\to\infty} n \sum_{j=2}^{\lfloor n/k \rfloor} \mathbb{P}(X_1 > u_n\,,\, X_j > u_n) = 0\,.$$

We show that $(\widetilde{V}_k)_{k\in\mathbb{Z}}$ satisfies the $\boldsymbol{D_r(u_n)}$ and $D'(u_n)$ conditions. The result is an analogon for discrete-time MA processes given in Rootzén [131], Lemma 3.2.

Lemma A.3.2
Let V *be a stationary version of the Lévy-OU process given by* (3.2.5) *with* L *a positive, compound Poisson process as in (3.2.4).*

(a) *Assume* $\widetilde{V}_k = V_{\Gamma_k/\lambda} + \mathrm{e}^{-\Gamma_k}\xi_0 \in \mathcal{L}(\gamma)$, $\gamma > 0$, *such that for* $a_n > 0$, $b_n \in \mathbb{R}$ *and* $u_n = a_n x + b_n$,

$$\lim_{n\to\infty} n\mathbb{P}(\widetilde{V}_k > u_n) = \mathrm{e}^{-x} \qquad \text{for } x \in \mathbb{R}. \tag{A.3.1}$$

For $r \in \mathbb{N}$ *and* $\mathbf{x} = (x_1, \dots, x_r)$ *let* $\mathbf{u}_n = (a_n x_1 + b_n, \dots, a_n x_r + b_n)$. *Then* $(\widetilde{V}_k)_{k\in\mathbb{N}}$ *satisfies the* $\boldsymbol{D_r(u_n)}$ *and* $D'(u_n)$ *conditions.*

(b) *Let* L_1 *be in* $\mathcal{S}(\gamma)$, $\gamma > 0$. *Define* $M_k = \sup_{(k-1)h \leq t \leq kh} V_t$ *for* $h > 0$, $k \in \mathbb{N}$. *Suppose* $M_1 \in \mathcal{L}(\gamma)$ *such that for* $a_n > 0, b_n \in \mathbb{R}$ *and* $u_n = a_n x + b_n$,

$$\lim_{n\to\infty} n\mathbb{P}(M_k > u_n) = \mathrm{e}^{-x} \qquad \text{for } x \in \mathbb{R}. \tag{A.3.2}$$

For $r \in \mathbb{N}$ *and* $\mathbf{x} = (x_1, \dots, x_r)$ *let* $\mathbf{u}_n = (a_n x_1 + b_n, \dots, a_n x_r + b_n)$. *Then* $(M_k)_{k\in\mathbb{N}}$ *satisfies the* $\boldsymbol{D_r(u_n)}$ *and* $D'(u_n)$ *conditions.*

Proof.

(a) To show the $\boldsymbol{D_r(u_n)}$ condition, let $u_n^{(m)} = a_n x_m + b_n$, $x_m \in \mathbb{R}$, $m = 1,\dots,r$. Let $\mathbf{v}_n = (v_n^{(1)},\dots,v_n^{(p)})$, $\mathbf{w}_n = (w_n^{(1)},\dots,w_n^{(q)})$ with $v_n^{(l)}, w_n^{(s)} \in \{u_n^{(1)},\dots,u_n^{(r)}\}$. Let $1 \le i_1 < \cdots < i_p < j_1 < \cdots < j_q \le n$. Define $V_k^n = \sum_{j=k-n}^{k} \mathrm{e}^{-(\Gamma_k-\Gamma_j)}\xi_j$ for $k,n \in \mathbb{N}$, $\mathbf{V}_I = (\widetilde{V}_{i_1},\dots,\widetilde{V}_{i_p})$, $\mathbf{V}_I^n = (V_{i_1}^n,\dots,V_{i_p}^n)$, and similarly $\mathbf{V}_J = (\widetilde{V}_{j_1},\dots,\widetilde{V}_{j_q})$, $\mathbf{V}_J^n = (V_{j_1}^n,\dots,V_{j_q}^n)$. Then $(V_k^n)_{k\in\mathbb{Z}}$ is stationary and $\mathbf{V}_I^{\lfloor n\delta\rfloor}$ is independent of $\mathbf{V}_J^{\lfloor n\delta\rfloor}$ for $j_1 - i_p > \lfloor n\delta\rfloor$, $\delta > 0$. Since $\mathbb{P}(\xi_k < 0) = 0$, we obtain $\mathbf{V}_I^n \le \mathbf{V}_I$. Here, $\mathbf{x} = (x_1,\dots,x_p) \le \mathbf{y} = (y_1,\dots,y_p)$ means that $x_i \le y_i$ for all $i = 1,\dots,p$. It now follows that for any $\epsilon > 0$,

$$
\begin{aligned}
&\mathbb{P}\left(\mathbf{V}_I \le \mathbf{v}_n, \mathbf{V}_J \le \mathbf{w}_n\right) \le \mathbb{P}\left(\mathbf{V}_I^{\lfloor n\delta\rfloor} \le \mathbf{v}_n, \mathbf{V}_J^{\lfloor n\delta\rfloor} \le \mathbf{w}_n\right)\\
&= \mathbb{P}\left(\mathbf{V}_I^{\lfloor n\delta\rfloor} \le \mathbf{v}_n\right)\mathbb{P}\left(\mathbf{V}_J^{\lfloor n\delta\rfloor} \le \mathbf{w}_n\right)\\
&\le \mathbb{P}\left(\mathbf{V}_I \le \mathbf{v}_n + \epsilon(a_n,\dots,a_n)\right)\mathbb{P}\left(\mathbf{V}_J \le \mathbf{w}_n + \epsilon(a_n,\dots,a_n)\right)\\
&\quad + n\mathbb{P}\left(|\widetilde{V}_1 - V_1^{\lfloor n\delta\rfloor}| > \epsilon a_n\right)\\
&\le \mathbb{P}\left(\mathbf{V}_I \le \mathbf{v}_n\right)\mathbb{P}\left(\mathbf{V}_J \le \mathbf{w}_n\right) + n\mathbb{P}\left(|\widetilde{V}_1 - V_1^{\lfloor n\delta\rfloor}| > \epsilon a_n\right)\\
&\quad + \sum_{m=1}^{r} n\mathbb{P}(u_n^{(m)} \le \widetilde{V}_1 \le u_n^{(m)} + \epsilon a_n).
\end{aligned}
$$

Similarly, we can find a lower bound, such that for $j_1 - i_p > \lfloor n\delta\rfloor$,

$$
\begin{aligned}
\alpha_{n,\lfloor n\delta\rfloor} &:= \left|\mathbb{P}\left(\mathbf{V}_I \le \mathbf{v}_n, \mathbf{V}_J \le \mathbf{w}_n\right) - \mathbb{P}\left(\mathbf{V}_I \le \mathbf{v}_n\right)\mathbb{P}\left(\mathbf{V}_J \le \mathbf{w}_n\right)\right|\\
&\le n\mathbb{P}(|\widetilde{V}_1 - V_1^{\lfloor n\delta\rfloor}| > \epsilon a_n) + \sum_{m=1}^{r} n\mathbb{P}(u_n^{(m)} - \epsilon a_n \le \widetilde{V}_1 \le u_n^{(m)} + \epsilon a_n)\\
&=: \widetilde{\alpha}_{n,\lfloor n\delta\rfloor,\epsilon}.
\end{aligned} \tag{A.3.3}
$$

Let $X_i := \Gamma_i - \Gamma_{i-1} - 1/\mu$, $i \in \mathbb{N}$. Then $(X_i)_{i\in\mathbb{N}}$ is a centered i. i. d. sequence such that $\sum_{i=1}^n X_i = \Gamma_n - n/\mu$. It follows that there exists a constant $K > 0$, such that for every $n \in \mathbb{N}$,

$$
\begin{aligned}
\mathbb{E}(\Gamma_n - n/\mu)^6 &= n\mathbb{E}X_1^6 + \binom{n}{3}\binom{6}{2}\binom{4}{2}(\mathbb{E}X_1^2)^3\\
&\quad + \binom{n}{2}\binom{6}{3}(\mathbb{E}X_1^3)^2 + \binom{n}{2}\binom{6}{2}(\mathbb{E}X_1^2)(\mathbb{E}X_1^4)\\
&\le n^3 K.
\end{aligned}
$$

Hence, by Markov's inequality there is a constant $\widetilde{K} > 0$ such that for all $n \in \mathbb{N}$,

$$
\mathbb{P}(\Gamma_n < n/(2\mu)) \le \mathbb{P}(|\Gamma_n - n/\mu| > n/(2\mu)) \le (2\mu/n)^6\mathbb{E}(\Gamma_n - n/\mu)^6 \le \widetilde{K}/n^3. \tag{A.3.4}
$$

Thus we obtain for $n \in \mathbb{N}$,

$$\begin{aligned} n\mathbb{P}(|\widetilde{V}_1 - V_1^{\lfloor n\delta \rfloor}| > \epsilon a_n) &= n\mathbb{P}\left(\mathrm{e}^{-\Gamma_{\lfloor n\delta \rfloor}} \sum_{j=-\infty}^{-\lfloor n\delta \rfloor -1} \mathrm{e}^{(\Gamma_{\lfloor n\delta \rfloor}+\Gamma_j)} \xi_j > \epsilon a_n \right) \\ &\leq n\mathbb{P}\left(\mathrm{e}^{-\lfloor n\delta \rfloor/(2\mu)} \widetilde{V}_1 > \epsilon a_n \right) + \widetilde{K}/n^2. \end{aligned} \quad \text{(A.3.5)}$$

Note, that the first exponential moment of $\gamma\beta \mathrm{e}^{-\lfloor n\delta \rfloor/(2\mu)} \widetilde{V}_1$ exists for $\beta \mathrm{e}^{-\lfloor n\delta \rfloor/(2\mu)} < 1$. Choose $\beta_n = 2/(\epsilon(1-\epsilon)) \log n$. There exists $n_0 = n_0(\delta, \epsilon) \in \mathbb{N}$ such that $\beta_n \mathrm{e}^{-\lfloor n\delta \rfloor/(2\mu)} < (1-\epsilon)$ and $a_n \geq (1-\epsilon)/\gamma$ for $n \geq n_0$. The first term of the right hand side of (A.3.5) is by Markov's inequality for $n \geq n_0$ bounded above by

$$n\mathbb{E} \exp\left[\beta_n \gamma \left(\mathrm{e}^{-\lfloor \delta n \rfloor/(2\mu)} \widetilde{V}_1\right)\right] \mathrm{e}^{-\beta_n \epsilon \gamma a_n} \leq n\mathbb{E} \exp\left[(1-\epsilon)\gamma \widetilde{V}_1\right] \mathrm{e}^{-2\log n}, \quad \text{(A.3.6)}$$

which converges to 0 as $n \to \infty$. Together with (A.3.1), (A.3.3) and (A.3.5) this gives

$$\lim_{n\to\infty} \widetilde{\alpha}_{n,\lfloor n\delta \rfloor,\epsilon} = \sum_{m=1}^{r} [\mathrm{e}^{-(x_m-\epsilon)} - \mathrm{e}^{-(x_m+\epsilon)}],$$

so that

$$\lim_{n\to\infty} \alpha_{n,\lfloor n\delta \rfloor} \leq \lim_{\epsilon \downarrow 0} \lim_{n\to\infty} \widetilde{\alpha}_{n,\lfloor n\delta \rfloor,\epsilon} = 0,$$

which implies the $\boldsymbol{D_r(u_n)}$ condition by Lemma 3.2.1 in Leadbetter et al. [95].

To show the $D'(u_n)$ condition, let $\epsilon > 0$. Then there exists an $x_0 > 0$ such that $\mathbb{P}(\Gamma_2 - \Gamma_1 < x_0) = \epsilon$. Since $(\Gamma_{i+1} - \Gamma_i)_{i\in\mathbb{N}}$ is a positive i. i. d. sequence, it follows that

$$\mathbb{P}(\Gamma_j - \Gamma_1 < x_0) \leq \mathbb{P}(\Gamma_j - \Gamma_{j-1} < x_0, \ldots, \Gamma_2 - \Gamma_1 < x_0) = \epsilon^{j-1}, \quad j \geq 2.$$

Now choose β such that $1/2 < \beta < (1 + \mathrm{e}^{-x_0})^{-1}$ and $\delta > 0$ such that $1 + \delta < 2\beta$. Then, for any $k, n \in \mathbb{N}$,

$$\sum_{j=2}^{\lfloor n/k \rfloor} \mathbb{P}(\widetilde{V}_1 > u_n, \widetilde{V}_j > u_n) = \sum_{j=2}^{\lfloor n^\delta \rfloor} \mathbb{P}(\widetilde{V}_1 > u_n, \widetilde{V}_j > u_n) + \sum_{j=\lfloor n^\delta \rfloor+1}^{\lfloor n/k \rfloor} \mathbb{P}(\widetilde{V}_1 > u_n, \widetilde{V}_j > u_n) \text{(A.3.7)}$$

We first show that the first summand, when multiplied by n, tends to 0 as $n \to \infty$. Note that by the independence of $\widetilde{V}_1$ and $\Gamma_j - \Gamma_1$,

$$\mathbb{P}(\widetilde{V}_1 > u_n, \widetilde{V}_j > u_n) \leq \epsilon^{j-1}\, \mathbb{P}(\widetilde{V}_1 > u_n) + \mathbb{P}(\widetilde{V}_1 + \widetilde{V}_j > 2u_n, \Gamma_j - \Gamma_1 \geq x_0). \quad \text{(A.3.8)}$$

Let $\widetilde{V}_1'$ be an independent copy of $\widetilde{V}_1$. Then

$$\begin{aligned}
&\mathbb{P}\left(\widetilde{V}_1+\widetilde{V}_j>2u_n,\Gamma_j-\Gamma_1\ge x_0\right)\\
&\quad=\mathbb{P}\left((1+\mathrm{e}^{-(\Gamma_j-\Gamma_1)})\widetilde{V}_1+\sum_{k=2}^{j}\mathrm{e}^{-(\Gamma_j-\Gamma_k)}\xi_k>2u_n,\Gamma_j-\Gamma_1\ge x_0\right)\\
&\quad\le\mathbb{P}\left((1+\mathrm{e}^{-x_0})\widetilde{V}_1+\sum_{k=2}^{j}\mathrm{e}^{-(\Gamma_j-\Gamma_k)}\xi_k>2u_n\right)\\
&\quad\le\mathbb{P}\left((1+\mathrm{e}^{-x_0})\widetilde{V}_1+\widetilde{V}_1'>2u_n\right). \qquad\qquad (A.3.9)
\end{aligned}$$

Since the first exponential moment of $\beta\gamma((1+\mathrm{e}^{-x_0})\widetilde{V}_1+\widetilde{V}_1')$ exists if $\beta(1+\mathrm{e}^{-x_0})<1$, the last expression is bounded above by

$$\mathbb{E}\exp\left[\beta\gamma((1+\mathrm{e}^{-x_0})\widetilde{V}_1+\widetilde{V}_1')\right]\mathrm{e}^{-2\beta\gamma u_n} \qquad (A.3.10)$$

by the Markov inequality. Since $\left(n\mapsto\mathrm{e}^{-2\beta\gamma u_n}\right)\in\mathcal{R}_{-2\beta}$ for fixed x, it follows that $\left(n\mapsto n^{1+\delta}\mathrm{e}^{-2\beta\gamma u_n}\right)\in\mathcal{R}_{(1+\delta)-2\beta}$, where $1+\delta-2\beta<0$. We then obtain by (A.3.1), (A.3.8)-(A.3.10) and Bingham et al. [29], Proposition 1.5.1, that

$$\begin{aligned}
&n\sum_{j=2}^{\lfloor n^\delta\rfloor}\mathbb{P}(\widetilde{V}_1>u_n,\widetilde{V}_j>u_n) \qquad\qquad (A.3.11)\\
&\quad\le n\mathbb{P}(\widetilde{V}_1>u_n)\sum_{j=2}^{\lfloor n^\delta\rfloor}\epsilon^{j-1}+\mathbb{E}[\beta\gamma((1+\mathrm{e}^{-x_0})\widetilde{V}_1+\widetilde{V}_1')]n^{1+\delta}\mathrm{e}^{-2\beta\gamma u_n}\\
&\quad\stackrel{n\to\infty}{\longrightarrow}\mathrm{e}^{-x}\,\frac{\epsilon}{1-\epsilon}.
\end{aligned}$$

For the second term in (A.3.7), using the independence of $V_j^{\lfloor n^\delta\rfloor}$ and $V_1^{\lfloor n^\delta\rfloor}$ for $j>\lfloor n^\delta\rfloor$, we obtain

$$\begin{aligned}
&n\sum_{j=\lfloor n^\delta\rfloor+1}^{\lfloor n/k\rfloor}\mathbb{P}(\widetilde{V}_1>u_n,\widetilde{V}_j>u_n)\\
&\le n\sum_{j=\lfloor n^\delta\rfloor+1}^{\lfloor n/k\rfloor}\mathbb{P}(V_1^{\lfloor n^\delta\rfloor}>u_n-\epsilon\,a_n,V_j^{\lfloor n^\delta\rfloor}>u_n-\epsilon\,a_n)+2n^2\mathbb{P}(|\widetilde{V}_1-V_1^{\lfloor n^\delta\rfloor}|>\epsilon\,a_n)\\
&=n\sum_{j=\lfloor n^\delta\rfloor+1}^{\lfloor n/k\rfloor}\mathbb{P}(V_1^{\lfloor n^\delta\rfloor}>u_n-\epsilon\,a_n)\mathbb{P}(V_j^{\lfloor n^\delta\rfloor}>u_n-\epsilon\,a_n)+2n^2\mathbb{P}(|\widetilde{V}_1-V_1^{\lfloor n^\delta\rfloor}|>\epsilon\,a_n)\\
&\le(n^2/k)\,\mathbb{P}(\widetilde{V}_1>u_n-\epsilon\,a_n)^2+2n^2\mathbb{P}(|\widetilde{V}_1-V_1^{\lfloor n^\delta\rfloor}|>\epsilon\,a_n). \qquad\qquad (A.3.12)
\end{aligned}$$

Analogously to (A.3.5) and (A.3.6), with $\beta_n = 3/(\epsilon(1-\epsilon))\log n$, we have $n^2\mathbb{P}(|\widetilde{V}_1 - V_1^{\lfloor n^\delta \rfloor}| > \epsilon\, a_n) \to 0$ as $n \to \infty$. Using (A.3.1), we also have

$$\lim_{n\to\infty} (n^2/k)\,\mathbb{P}(\widetilde{V}_1 > u_n - \epsilon\, a_n)^2 = \exp(-2(x-\epsilon))/k, \tag{A.3.13}$$

which converges to 0 as $k \to \infty$. Then by (A.3.7), (A.3.11)-(A.3.13), and letting $\epsilon \downarrow 0$, the $D'(u_n)$ condition holds.

(b) To prove condition $\boldsymbol{D_r(u_n)}$, we replace V_k^n in *(a)* by

$$M_k^n := \sup_{(k-1)h \le t \le kh} \int_{t-nh}^{t} \mathrm{e}^{-\lambda(t-s)}\, dL_{\lambda s}.$$

We then obtain an analogue result to (A.3.3). Further, since

$$\begin{aligned} |M_k - M_k^n| &\le \sup_{(k-1)h \le t \le kh} \int_{-\infty}^{t-nh} \mathrm{e}^{-\lambda(t-s)}\, dL_{\lambda s} \\ &\overset{d}{=} \mathrm{e}^{-\lambda nh} \sup_{(k-1)h \le t \le kh} \int_{-\infty}^{t} \mathrm{e}^{-\lambda(t-s)}\, dL_{\lambda s} = \mathrm{e}^{-\lambda nh} M_k, \end{aligned}$$

we obtain for any $\delta > 0$ that

$$n\mathbb{P}(|M_k - M_k^{\lfloor n\delta \rfloor}| > \epsilon\, a_n) \le n\mathbb{P}(\mathrm{e}^{-\lambda \lfloor n\delta \rfloor h} M_k > \epsilon\, a_n). \tag{A.3.14}$$

Since the first exponential moment of $\beta\gamma\mathrm{e}^{-\lambda\lfloor n\delta\rfloor h} M_k$ exists for $\beta\mathrm{e}^{-\lambda\lfloor n\delta\rfloor h} < 1$, similar reasoning as in (A.3.5) and (A.3.6) shows that $\lim_{n\to\infty} n\mathbb{P}(|M_k - M_k^{\lfloor n\delta\rfloor}| > \epsilon\, a_n) = 0$. As in the proof of *(a)* we then conclude that the $\boldsymbol{D_r(u_n)}$ condition holds.

For the proof of the $D'(u_n)$ condition we use

$$\begin{aligned} M_k &\le \int_{-\infty}^{\infty} \sup_{(k-1)h \le t \le kh} \mathrm{e}^{-\lambda(t-s)}\, \mathbf{1}_{(-\infty,t]}(s)\, dL_{\lambda s} \\ &= L_{\lambda kh} - L_{\lambda(k-1)h} + \int_{-\infty}^{(k-1)h} \mathrm{e}^{-\lambda((k-1)h-s)}\, dL_{\lambda s} =: \overline{V}_k. \end{aligned}$$

Let $j \geq 3$. Then we have the upper bound

$$\begin{aligned}\overline{V}_j &= L_{\lambda jh} - L_{\lambda(j-1)h} + \mathrm{e}^{-\lambda(j-1)h} \int_{-\infty}^{0} \mathrm{e}^{\lambda s}\, dL_{\lambda s} \\ &\quad + \mathrm{e}^{-\lambda(j-2)h} \int_0^h \mathrm{e}^{-\lambda(h-s)}\, dL_{\lambda s} + \int_h^{(j-1)h} \mathrm{e}^{-\lambda((j-1)h-s)}\, dL_{\lambda s} \\ &\leq L_{\lambda jh} - L_{\lambda(j-1)h} + \mathrm{e}^{-\lambda(j-2)h} \int_{-\infty}^{0} \mathrm{e}^{\lambda s}\, dL_{\lambda s} \\ &\quad + \mathrm{e}^{-\lambda(j-2)h} L_{\lambda h} + \int_h^{(j-1)h} \mathrm{e}^{-\lambda((j-1)h-s)}\, dL_{\lambda s} \\ &\leq L_{\lambda jh} - L_{\lambda(j-1)h} + \mathrm{e}^{-\lambda h}\overline{V}_1 + \int_h^{(j-1)h} \mathrm{e}^{-\lambda((j-1)h-s)}\, dL_{\lambda s}.\end{aligned}$$

Let $\overline{V}_1'$ be an independent copy of $\overline{V}_1$. Then

$$\begin{aligned} n\sum_{j=3}^{\lfloor n^\delta \rfloor} \mathbb{P}(M_1 > u_n, M_j > u_n) &\leq n\sum_{j=3}^{\lfloor n^\delta \rfloor} \mathbb{P}(\overline{V}_1 > u_n, \overline{V}_j > u_n) \qquad \text{(A.3.15)} \\ &\leq n\sum_{j=3}^{\lfloor n^\delta \rfloor} \mathbb{P}(\overline{V}_1 + \overline{V}_j > 2u_n) \\ &\leq n^{1+\delta}\, \mathbb{P}((1+\mathrm{e}^{-\lambda h})\overline{V}_1 + \overline{V}_1' > 2u_n).\end{aligned}$$

The tail of $\overline{V}_1$ behaves by Proposition 3.2.9 *(b)* and Theorem A.1.4 *(ii)* like

$$\mathbb{P}(\overline{V}_1 > x) = \mathbb{P}(L_{\lambda h} + V_0 > x) \sim \mathbb{E}\mathrm{e}^{\gamma V_0}\mathbb{P}(L_{\lambda h} > x) \quad \text{for } x \to \infty,$$

so that $\overline{V}_1 \in \mathcal{S}(\gamma)$. An analogue result to (A.3.10) gives

$$\lim_{n\to\infty} n^{1+\delta}\, \mathbb{P}((1+\mathrm{e}^{-\lambda h})\overline{V}_1 + \overline{V}_1' > 2u_n) = 0,$$

and, arguing similarly as in (A.3.12) and (A.3.13), we obtain

$$\lim_{k\to\infty} \limsup_{n\to\infty} n \sum_{j=3}^{\lfloor n/k \rfloor} \mathbb{P}(M_1 > u_n, M_j > u_n) = 0.$$

It remains to show that

$$\lim_{n\to\infty} n\mathbb{P}(M_1 > u_n, M_2 > u_n) = 0. \qquad \text{(A.3.16)}$$

Note that

$$\begin{aligned}\mathbb{P}(M_1 > u_n) &= \mathbb{P}(M_1 > u_n, N_{\lambda h} > 0) + \mathbb{P}(M_1 > u_n, N_{\lambda h} = 0) \\ &\geq \mathbb{P}(V_{\Gamma_1/\lambda} > u_n, N_{\lambda h} > 0) \\ &\geq \mathbb{P}(\xi_1 > u_n)\mathbb{P}(N_{\lambda h} > 0),\end{aligned}$$

so that

$$n\mathbb{P}(M_1 > u_n, M_2 > u_n) \leq \frac{n\mathbb{P}(M_1 > u_n)}{\mathbb{P}(N_{\lambda h} > 0)} \frac{\mathbb{P}(M_1 > u_n, M_2 > u_n)}{\mathbb{P}(\xi_1 > u_n)}. \tag{A.3.17}$$

Furthermore, we have the upper bound

$$\begin{aligned}\mathbb{P}(M_1 > u_n, M_2 > u_n) &\leq \mathbb{P}(\overline{V}_1 + \overline{V}_2 > 2u_n) \\ &\leq \mathbb{P}\left(\frac{L_{2\lambda h} - L_{\lambda h}}{2} + \frac{1+\mathrm{e}^{-\lambda h}}{2}\int_{-\infty}^{0} \mathrm{e}^{\lambda s}\, dL_{\lambda s} + \frac{1}{2}\int_0^h (1+\mathrm{e}^{-\lambda h+\lambda s})\, dL_{\lambda s} > u_n\right).\end{aligned}$$

The three summands are independent, and we shall show that for each of them the probability to be greater than u_n is of order $o(\mathbb{P}(\xi_1 > u_n))$ for $n \to \infty$, so that by Theorem A.1.4 *(ii)* (analog to Proposition 1.1.2 *(iii)*)

$$\lim_{n\to\infty} \mathbb{P}(M_1 > u_n, M_2 > u_n)/\mathbb{P}(\xi_1 > u_n) = 0. \tag{A.3.18}$$

Equation (A.3.16) and hence condition $D'(u_n)$ then follow from (A.3.2), (A.3.17) and (A.3.18).

The rapidly varying tails and Theorem A.1.4 *(i)* give

$$\lim_{x\to\infty} \frac{\mathbb{P}(L_{2\lambda h} - L_{\lambda h} > 2x)}{\mathbb{P}(\xi_1 > x)} = \lim_{x\to\infty} \mu \frac{\mathbb{P}(L_{\lambda h} > x)}{\nu(x,\infty)} \frac{\mathbb{P}(L_{\lambda h} > 2x)}{\mathbb{P}(L_{\lambda h} > x)} = 0,$$

which is the assertion for the first summand. Further, also by the rapidly varying tails, Proposition 3.2.9 *(b)* and Theorem A.1.4 *(i)*,

$$\frac{\mathbb{P}\left((1+\mathrm{e}^{-\lambda h})\int_{-\infty}^0 \mathrm{e}^{\lambda s}\, dL_{\lambda s} > 2x\right)}{\mathbb{P}(\xi_1 > x)} = \frac{\mathbb{P}\left((1+\mathrm{e}^{-\lambda h})V_0 > 2x\right)}{\mathbb{P}(V_0 > x)} \frac{\mu\mathbb{P}(V_0 > x)}{\nu(x,\infty)} \stackrel{x\to\infty}{\longrightarrow} 0.$$

For the last summand we use that

$$X := \int_0^h (1+\mathrm{e}^{-\lambda h}\,\mathrm{e}^{\lambda s})\, dL_{\lambda s} \stackrel{d}{=} \sum_{i=1}^{N_{\lambda h}} (1+\mathrm{e}^{-\lambda h}\mathrm{e}^{\lambda h U_i})\xi_i \stackrel{d}{=} \sum_{i=1}^{N_{\lambda h}} (1+\mathrm{e}^{-\lambda h U_i})\xi_i,$$

where $(U_i)_{i\in\mathbb{N}}, U$ are i. i. d. uniform on $(0,1)$ and independent of L (cf. Lemma 1.3.2). From Theorem A.1.4 *(iii)* then follows

$$\frac{\mathbb{P}\left(\int_0^h (1+\mathrm{e}^{-\lambda s})\, dL_{\lambda s} > 2x\right)}{\mathbb{P}(\xi_1 > x)} \sim \mu\lambda h \mathbb{E}\mathrm{e}^{\gamma X} \frac{\mathbb{P}\left(\xi_1(1+\mathrm{e}^{-\lambda h U})/2 > x\right)}{\mathbb{P}(\xi_1 > x)} \stackrel{x\to\infty}{\longrightarrow} 0.$$

□

A.4 Auxiliary results

Lemma A.4.1
Let Y with $Y(t) = \sum_{k=-\infty}^{\infty} f(R_k, t - \Gamma_k) Z_k$ for $t \in \mathbb{R}$ be a stationary mixed Poisson shot noise process satisfying the condition (M1) and the sequence $0 < a_n \uparrow \infty$ of constants is given by (2.0.8). Then

$$\lim_{m\to\infty} \lim_{n\to\infty} n\mathbb{P}\left(\left|\sum_{|k|>m} \mathbf{f}(R_k,\Gamma_k)Z_k\right| > a_n x\right) = 0.$$

Proof.
Let $l > 0$ be fixed. Define $\Omega = \mathbb{R}_+ \times \mathbb{R}$, $\Omega^{(l)} = \{(r,s) \in \mathbb{R}_+ \times \mathbb{R} : |s| > l\}$ and $\mathbf{f}_l(r,t) = \mathbf{f}(r,t)\,\mathbf{1}_{[-l,l]}(t)$ for $t \in \mathbb{R}$, $r \in \mathbb{R}_+$. We decompose the probability

$$\begin{aligned}
\mathbb{P}\left(\left|\sum_{|k|>m} \mathbf{f}(R_k,\Gamma_k)Z_k\right| > 3a_n x\right) &\le \mathbb{P}\left(\left|\int_{\Omega^{(l)}} \mathbf{f}(r,s)\,d\Lambda_1(r,s)\right| > a_n x\right) \\
&+\mathbb{P}\left(\left|\int_{\Omega\setminus\Omega^{(l)}} \mathbf{f}(r,s)\,d\Lambda_1(r,s) - \sum_{k=-m}^{m} \mathbf{f}_l(R_k,\Gamma_k)Z_k\right| > a_n x\right) \\
&+\mathbb{P}\left(\left|\sum_{k=-m}^{m} \mathbf{f}_l(R_k,\Gamma_k)Z_k - \sum_{k=-m}^{m} \mathbf{f}(R_k,\Gamma_k)Z_k\right| > a_n x\right). \qquad \text{(A.4.1)}
\end{aligned}$$

Step 1. For the first summand of (A.4.1) we obtain by an application of (2.2.14)

$$\lim_{n\to\infty} n\mathbb{P}\left(\left|\int_{\Omega^{(l)}} \mathbf{f}(r,s)\,d\Lambda_1(r,s)\right| > a_n x\right) = x^{-\alpha}\int_{\Omega^{(l)}} |\mathbf{f}(r,s)|^{\alpha}\,\pi(dr)\,ds \overset{l\to\infty}{\longrightarrow} 0. \quad \text{(A.4.2)}$$

Step 2. The second summand of (A.4.1) has the upper bound

$$\begin{aligned}
&\mathbb{P}\left(\left|\int_{\Omega\setminus\Omega^{(l)}} \mathbf{f}(r,s)\,d\Lambda_1(r,s) - \sum_{k=-m}^{m} \mathbf{f}_l(R_k,\Gamma_k)Z_k\right| > a_n x\right) \\
&= \sum_{i,j\ge m+1} \mathbb{P}\left(\left|\sum_{k=m+1}^{i} \mathbf{f}(R_k,\Gamma_k)Z_k + \sum_{k=-j}^{-m-1} \mathbf{f}(R_k,\Gamma_k)Z_k\right| > a_n x \,\middle|\, N(l)=i, N(-l)=-j\right) \\
&\qquad \times\mathbb{P}(N(l)=i, N(-l)=-j) \\
&\le \sum_{i=m+1}^{\infty}\sum_{j=m+1}^{\infty} \mathbb{P}\left(f^{+}\sum_{k=1}^{i+j}|Z_k| > a_n x\right)\mathbb{P}(N(l)=i)\mathbb{P}(N(-l)=-j).
\end{aligned}$$

Regarding (2.10) in Chover et al. [43], for any $\epsilon > 0$ there exists a $K > 0$ such that

$$\mathbb{P}\left(f^+ \sum_{k=1}^{i+j} |Z_k| > a_n x\right) \le \mathbb{P}\left(f^+ |Z_k| > a_n x\right) K(1+\epsilon)^{i+j} \quad \text{for } n, i, j \in \mathbb{N}.$$

Applying dominated convergence we obtain

$$\begin{aligned} \lim_{n\to\infty} n\mathbb{P}\left(\left|\int_{\Omega\setminus\Omega^{(l)}} \mathbf{f}(r,s)\, d\Lambda_1(r,s) - \sum_{k=-m}^{m} \mathbf{f}_l(R_k,\Gamma_k) Z_k\right| > a_n x\right) \qquad \text{(A.4.3)} \\ \le x^{-\alpha} K \left[\sum_{i=m+1}^{\infty} (1+\epsilon)^i \mathbb{P}(N(l)=i)\right]^2 \overset{m\to\infty}{\longrightarrow} 0. \end{aligned}$$

Step 3. Considering the last summand of (A.4.1) we obtain by Lemma 2.1.3

$$\begin{aligned} &\lim_{n\to\infty} n\mathbb{P}\left(\left|\sum_{k=-m}^{m} \mathbf{f}(R_k,\Gamma_k) Z_k - \sum_{k=-m}^{m} \mathbf{f}_l(R_k,\Gamma_k) Z_k\right| > a_n x\right) \\ &= x^{-\alpha} \sum_{k=-m}^{m} \mathbb{E}|\mathbf{f}(R_k,\Gamma_k) - \mathbf{f}_l(R_k,\Gamma_k)|^\alpha \\ &\overset{m\to\infty}{\longrightarrow} x^{-\alpha} \sum_{k=-\infty}^{\infty} \mathbb{E}|\mathbf{f}(R_k,\Gamma_k) - \mathbf{f}_l(R_k,\Gamma_k)|^\alpha. \qquad \text{(A.4.4)} \end{aligned}$$

Since $\mathbb{E}|\mathbf{f}(R_k,\Gamma_k) - \mathbf{f}_l(R_k,\Gamma_k)|^\alpha \le \mathbb{E}|\mathbf{f}(R_k,\Gamma_k)|^\alpha$ dominated convergence yields

$$\sum_{k=-\infty}^{\infty} \mathbb{E}|\mathbf{f}(R_k,\Gamma_k) - \mathbf{f}_l(R_k,\Gamma_k)|^\alpha \longrightarrow 0 \quad \text{for } l \to \infty. \qquad \text{(A.4.5)}$$

Thus we get by (A.4.4) and (A.4.5)

$$\lim_{l\to\infty} \lim_{m\to\infty} \lim_{n\to\infty} n\mathbb{P}\left(\left|\sum_{k=-m}^{m} \mathbf{f}(R_k,\Gamma_k) Z_k - \sum_{k=-m}^{m} \mathbf{f}_l(R_k,\Gamma_k) Z_k\right| > a_n x\right) = 0. \qquad \text{(A.4.6)}$$

Hence, by (A.4.1)-(A.4.3) and (A.4.6) the result follows. □

Lemma A.4.2
Let either of the following two conditions hold:

(a) *Let Y be a stationary mixed MA process given by (2.0.1) satisfying condition (M2). Moreover, Y has decomposition $Y = Y_1 + Y_2$ as given in (2.0.6) with $Y_1(t) = \sum_{k=-\infty}^{\infty} f(R_k, t-\Gamma_k)Z_k$ for $t \in \mathbb{R}$.*

(b) *Let Y be a stationary renewal shot noise process given by (2.3.1) satisfying condition (R2) such that $Y(t) = \sum_{k=-\infty}^{\infty} f(t-\Gamma_k)Z_k$ for $t \in \mathbb{R}$.*

Then the sequence $0 < a_n \uparrow \infty$ of constants satisfies

$$n\mathbb{P}(a_n^{-1}Z_1 \in \cdot) \overset{v}{\Longrightarrow} 1/\mu\sigma(\cdot) \quad \text{on } \mathcal{B}(\overline{\mathbb{R}}\backslash\{0\}), \tag{A.4.7}$$

where $\sigma(dx) = p\alpha x^{-\alpha-1}\,\mathbf{1}_{(0,\infty)}(x)\,dx + (1-p)\alpha(-x)^{-\alpha-1}\,\mathbf{1}_{(-\infty,0)}(x)\,dx$ for some $p \in (0,1]$. For some $\eta \in \mathbb{R}$ define

$$I_k = [\eta + (\Gamma_{k-1}+\Gamma_k)/2, \eta + (\Gamma_k + \Gamma_{k+1})/2) \qquad \text{for } k \in \mathbb{N}.$$

If $(1-p)f^- > 0$ assume furthermore there exist independent sequences of i. i. d. positive r. v. s $\{Z_k^{(s)}\}_{k\in\mathbb{Z}}$ for $s = 1, 2$, such that $Z_k = Z_k^{(1)} - Z_k^{(2)}$,

$$\mathbb{P}(Z_k > x) \sim \mathbb{P}(Z_k^{(1)} > x) \quad \text{and} \quad \mathbb{P}(Z_k < x) \sim \mathbb{P}(Z_k^{(2)} > x) \quad \text{for } x \to \infty.$$

In the case $(1-p)f^- = 0$ define $Z_k^{(1)} = Z_k$ and $Z_k^{(2)} = 0$. For $k \in \mathbb{Z}$ let be

$$Y_k^{(1)} = \sum_{j=-\infty}^{\infty} \sup_{h\in I_k}\{f(h-\Gamma_j)\}Z_j^{(1)}, \qquad Y_k^{(2)} = \sum_{j=-\infty}^{\infty} \sup_{h\in I_k}\{-f(h-\Gamma_j)\}Z_j^{(2)}.$$

Then holds for $x > 0$:

(i) $\lim_{n\to\infty} n\mathbb{P}\left(\sup_{h\in I_k} Y(h) \le a_n(x-\epsilon), Y_k^{(1)} + Y_k^{(2)} > a_n(x+\epsilon)\right) = 0.$

(ii) $\lim_{n\to\infty} n\mathbb{P}\left(\sup_{h\in I_k} Y(h) > a_n(x+\epsilon), Y_k^{(1)} + Y_k^{(2)} \le a_n(x-\epsilon)\right) = 0.$

Proof.
Define $Y_k = \sup_{h\in I_k} \sum_{j=-\infty}^{\infty} f(h-\Gamma_j)Z_j$ for $k \in \mathbb{Z}$. First we show for $x > 0$

$$\lim_{n\to\infty} n\mathbb{P}\left(Y_k \le a_n(x-\epsilon), Y_k^{(1)} + Y_k^{(2)} > a_n(x+\epsilon)\right) = 0. \tag{A.4.8}$$

We define

$$\begin{array}{rclrcl} \psi_{2j-1} &=& \sup_{h\in I_k}\{f(R_j,h-\Gamma_j)\}, & \psi_{2j} &=& \sup_{h\in I_k}\{-f(R_j,h-\Gamma_j)\},\\ \widetilde{\psi}_{2j-1}(t) &=& f(R_j,t-\Gamma_j), & \widetilde{\psi}_{2j}(t) &=& -f(R_j,t-\Gamma_j), \end{array}$$

$\widetilde{Z}_{2j-1} = Z_j^{(1)}$, $\widetilde{Z}_{2j} = Z_j^{(2)}$ for $j \in \mathbb{Z}$. Let $x > 0$ be fixed. We have the inequality

$$\begin{aligned} n\mathbb{P}&\left(\sum_{j=-\infty}^{\infty} \psi_j \widetilde{Z}_j > a_n(x+2\epsilon), \sup_{h\in I_k} \sum_{j=-\infty}^{\infty} \widetilde{\psi}_j(h)\widetilde{Z}_j \le a_n(x-2\epsilon)\right)\\ &\le n\mathbb{P}\left(\sum_{|j|>m} \psi_{2j-1}Z_j^{(1)} > a_n\epsilon/2\right) + n\mathbb{P}\left(\sum_{|j|>m} \psi_{2j}Z_j^{(2)} > a_n\epsilon/2\right)\\ &\quad +n\mathbb{P}\left(\sup_{h\in I_k} \sum_{\substack{j>2m\\ j<-2m-1}} \widetilde{\psi}_j(h)\widetilde{Z}_j < -a_n\epsilon\right) \qquad\qquad (A.4.9)\\ &\quad +n\mathbb{P}\left(\sum_{j=-2m-1}^{2m} \psi_j\widetilde{Z}_j > a_n(x+\epsilon), \sup_{h\in I_k} \sum_{j=-2m-1}^{2m} \widetilde{\psi}_j(h)\widetilde{Z}_j \le a_n(x-\epsilon)\right). \end{aligned}$$

The first three summands tend to 0 as $n, m \to \infty$ by Lemma A.4.1, where for the third summand we used additionally

$$\sum_{\substack{j>2m\\ j<-2m-1}} \widetilde{\psi}_j(\Gamma_k)\widetilde{Z}_j \le \sup_{h\in I_k} \sum_{\substack{j>2m\\ j<-2m-1}} \widetilde{\psi}_j(h)\widetilde{Z}_j.$$

We shall show that also the last summand of (A.4.9) tends to 0 as n tends to ∞. Note first that $\widetilde{\psi}_l(h_k) = \psi_l$ for some h_k in the closure of I_k, and $|\widetilde{\psi}_j(h)| \le f^+$ for every $h \in \mathbb{R}$. Then

$$\begin{aligned} \mathbb{P}&\left(\bigvee_{l=-2m-1}^{2m} \psi_l\widetilde{Z}_l > a_n x, \sup_{h\in I_k} \sum_{j=-2m-1}^{2m} \widetilde{\psi}_j(h)\widetilde{Z}_j \le a_n(x-\epsilon)\right) \qquad\qquad (A.4.10)\\ &\le \sum_{l=-2m-1}^{2m} \mathbb{P}\left(\psi_l\widetilde{Z}_l > a_n x, \sum_{\substack{j=-2m-1\\ j\ne l}}^{2m} \widetilde{\psi}_j(h_k)\widetilde{Z}_j + \psi_l\widetilde{Z}_l \le a_n(x-\epsilon)\right)\\ &\le \sum_{l=-2m-1}^{2m} \mathbb{P}\left(\psi_l\widetilde{Z}_l > a_n x, \sum_{\substack{j=-2m-1\\ j\ne l}}^{2m} -f^+|\widetilde{Z}_j| < -a_n\epsilon\right)\\ &\le \sum_{l=-2m-1}^{2m} \mathbb{P}\left(f^+|\widetilde{Z}_l| > a_n x\right)\mathbb{P}\left(\sum_{\substack{j=-2m-1\\ j\ne l}}^{2m} f^+|\widetilde{Z}_j| > \epsilon a_n\right), \end{aligned}$$

where the last inequality follows by independence of $\widetilde{Z}$. We conclude by (A.4.7) and (A.4.10)

$$\lim_{n\to\infty} n\mathbb{P}\left(\bigvee_{l=-2m-1}^{2m} \psi_l \widetilde{Z}_l > a_n x, \sup_{h\in I_k} \sum_{j=-2m-1}^{2m} \widetilde{\psi}_j(h)\widetilde{Z}_j \le a_n(x-\epsilon)\right) = 0. \quad \text{(A.4.11)}$$

Moreover, since $|\psi_j| \le f^+$ for $j \in \mathbb{Z}$,

$$\begin{aligned}
&\mathbb{P}\left(\bigvee_{l=-2m-1}^{2m} \psi_l \widetilde{Z}_l \le a_n x, \sum_{j=-2m-1}^{2m} \psi_j \widetilde{Z}_j > a_n(x+\epsilon)\right) \\
&\le \sum_{l=-2m-1}^{2m} \mathbb{P}\left(\sum_{\substack{j=-2m-1\\ j\ne k}}^{2m} \min\{\psi_j \widetilde{Z}_j, \psi_l \widetilde{Z}_l\} > a_n\epsilon\right) \\
&\le \sum_{l=-2m-1}^{2m} \mathbb{P}\left(f^+ \sum_{\substack{j=-2m-1\\ j\ne l}}^{2m} \min\{|\widetilde{Z}_j|, |\widetilde{Z}_l|\} > a_n\epsilon\right). \quad \text{(A.4.12)}
\end{aligned}$$

By the independence of $\widetilde{Z}_j$ and $\widetilde{Z}_k$ the r. h. s. is bounded above by

$$\sum_{l=-2m-1}^{2m} \sum_{\substack{j=-2m-1\\ j\ne l}}^{2m} \mathbb{P}\left(|\widetilde{Z}_j| > \frac{a_n\epsilon}{4m+2}\right) \mathbb{P}\left(|\widetilde{Z}_l| > \frac{a_n\epsilon}{4m+2}\right) = o\left(\frac{1}{n}\right) \quad \text{(A.4.13)}$$

for $n \to \infty$, where we used (A.4.7). Then (A.4.8) follows from (A.4.9)-(A.4.13).

Let $Y_2(t) = \int_{\mathbb{R}\times\mathbb{R}_+} f(t-s)\, d\Lambda_2(r,s)$ and $M^{(2)}(h) = \sup_{0\le t\le h} |Y_2(t)|$. The sequence $\overline{M}_k = \sup_{k-1\le t\le k} |Y_2(t)|$ is stationary, such that for $x \ge 0$,

$$\mathbb{P}(M^{(2)}(h) > x) \le \mathbb{P}\left(\bigcup_{k=1}^{\lceil h\rceil} \{\overline{M}_k > x\}\right) \le (h+1)\mathbb{P}(M^{(2)}(1) > x).$$

Denote by F_Γ the d. f. of Γ_1. Then we have an uniform bound:

$$\frac{\mathbb{P}(M^{(2)}(\Gamma) > x)}{\mathbb{P}(M^{(2)}(1) > x)} = \int_0^\infty \frac{\mathbb{P}(M^{(2)}(h) > x)}{\mathbb{P}(M^{(2)}(1) > x)} F_\Gamma(dh) \le \int_0^\infty (h+1) F_\Gamma(dh) = \frac{1}{\mu} + 1 \quad \text{(A.4.14)}$$

for all $x > 0$. Taking (1.3.15) into account there exists a $C > 0$, such that

$$\mathbb{P}(\sup_{0\le t\le 1} |Y_2(t)| > x) \le C\mathrm{e}^{-x} \quad \text{for } x > 0. \quad \text{(A.4.15)}$$

Note that (1.3.15) holds also for mixed MA processes, since the Lévy measure of $Y_2(t)$ has bounded support. By the independence of Γ and Λ_2 we obtain by (A.4.14) and (A.4.15)

$$n\mathbb{P}\left(\sup_{t\in I_k}|Y_2(t)| > a_n\epsilon\right) \leq 2nC\,(1/\mu+1)\,\mathrm{e}^{-a_n\epsilon} \stackrel{n\to\infty}{\longrightarrow} 0. \tag{A.4.16}$$

Using decomposition (2.0.4) we have

$$Y_k - \sup_{t\in I_k}(-Y_2(t)) \leq \sup_{t\in I_k} Y(t) \leq Y_k + \sup_{t\in I_k} Y_2(t). \tag{A.4.17}$$

For the proof of *(i)* we apply (A.4.8), (A.4.16) and (A.4.17) such that

$$\begin{aligned} &n\mathbb{P}\left(\sup_{t\in I_k} Y(t) \leq a_n(x-\epsilon), Y_k^{(1)} + Y_k^{(2)} > a_n(x+\epsilon)\right) \\ &\quad \leq n\mathbb{P}\left(\sup_{t\in I_k}|Y_2(t)| > a_n\epsilon\right) + n\mathbb{P}\left(Y_k \leq a_n x, Y_k^{(1)} + Y_k^{(2)} > a_n(x+\epsilon)\right) \stackrel{n\to\infty}{\longrightarrow} 0. \end{aligned}$$

In *(ii)* we have by (A.4.17)

$$\begin{aligned} &n\mathbb{P}\left(\sup_{t\in I_k} Y(t) > a_n(x+\epsilon), Y_k^{(1)} + Y_k^{(2)} \leq a_n(x-\epsilon)\right) \\ &\quad \leq n\mathbb{P}\left(\sup_{t\in I_k}|Y_2(t)| > a_n\epsilon\right) + n\mathbb{P}(Y_k^{(1)} + Y_k^{(2)} \leq a_n(x-\epsilon), Y_k > a_n x). \end{aligned} \tag{A.4.18}$$

In the case $(1-p)f^+ > 0$ the second summand of (A.4.18) is 0, such that by (A.4.16) statement *(ii)* holds. In the case $(1-p)f^+ = 0$ we have with g as in (2.0.12), (2.3.20), respectively,

$$\mathbb{P}(Y_k^{(1)} \leq a_n(x-\epsilon), Y_k > a_n x) \leq \mathbb{P}\left(\sum_{k=-\infty}^{\infty} g(\Gamma_k) Z_k^- > a_n\epsilon\right) = o(1/n) \text{ for } n\to\infty.$$

Hence also the second summand of (A.4.18) tends to 0. □

Bibliography

[1] Albin, J. M. P. On extremes of infinitely divisible Ornstein-Uhlenbeck processes. Preprint, available at `http://www.math.chalmers.se/~palbin/`.

[2] Albin, J. M. P. (1987) *On Extremal Theory for Nondifferentiable Stationary Processes*. Ph.D. thesis, Lund University.

[3] Albin, J. M. P. (1990) On extremal theory for stationary processes. *Ann. Probab.* **18** (1), 92–128.

[4] Albin, J. M. P. (1993) Extremes of totally skewed stable motion. *Statist. Probab. Lett.* **16** (3), 219–224.

[5] Albin, J. M. P. (1997) Extremes for non-anticipating moving averages of totally skewed α-stable motion. *Statist. Probab. Lett.* **36**, 289–297.

[6] Albin, J. M. P. (1999) Extremes of totally skewed α-stable processes. *Stochastic Process. Appl.* **79** (2), 185–212.

[7] Albin, J. M. P. (2000) Extremes and upcrossing intensities for P-differentiable stationary processes. *Stochastic Process. Appl.* **87** (2), 199–234.

[8] Anh, V. V., Heyde, C. C., and Leonenko, N. N. (2002) Dynamic models of long-memory processes driven by Lévy noise. *J. Appl. Probab.* **39** (4), 1–18.

[9] Applebaum, D. (2004) *Lévy processes and stochastic calculus*. Cambridge University Press, New York.

[10] Balkema, A. A. and de Haan, L. (1988) Almost sure continuity of stable moving average processes with index less than one. *Ann. Probab.* **16** (1), 333–343.

[11] Baltrunas, A., Daley, D. J., and Klüppelberg, C. (2002) Tail behaviour of the busy period of a GI/GI/1 queue with subexponential service times. *To appear in Stochastic Process. Appl.* .

[12] Barndorff-Nielsen, O. E. (1998) Processes of normal inverse Gaussian type. *Finance Stoch.* **2**, 41–68.

[13] Barndorff-Nielsen, O. E. (2001) Superposition of Ornstein–Uhlenbeck type processes. *Theory Probab. Appl.* **45** (2), 175–194.

[14] Barndorff-Nielsen, O. E. and Shephard, N. (2001) Modelling by Lévy processes for financial econometrics. In: O. E. Barndorff-Nielsen, T. Mikosch, and S. I. Resnick (Eds.), *Lévy Processes: Theory and Applications*, pp. 283–318. Birkhäuser, Boston.

[15] Barndorff-Nielsen, O. E. and Shephard, N. (2001) Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics (with discussion). *J. Roy. Statist. Soc. Ser. B* **63** (2), 167–241.

[16] Barndorff-Nielsen, O. E. and Shephard, N. (2002) Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *J. Roy. Statist. Soc. Ser. B* **64**, 253–280.

[17] Barndorff-Nielsen, O. E. and Shephard, N. (2004) *Financial Volatility in Continuous Time*. Cambridge University Press. To appear.

[18] Basrak, B. (2000) *The Sample Autocorrelation Function of Non-Linear Time Series*. Ph.D. thesis, Rijksuniversteit Groningen.

[19] Basrak, B., Davis, R. A., and Mikosch, T. (2002) A characterization of multivariate regular variation. *Ann. Appl. Probab.* **12**, 908–920.

[20] Basrak, B., Davis, R. A., and Mikosch, T. (2002) Regular variation of GARCH processes. *Stochastic Process. Appl.* **99**, 95–115.

[21] Bauer, H. (1992) *Maß- und Integrationstheorie*. 2nd ed. Walter de Gruyter, Berlin.

[22] Benhabib, J. and Rustichini, A. (1991) Vintage capital, investment and growth. *J. Econom. Theory* **55**, 323–329.

[23] Berman, S. M. (1964) Limit theorems for the maximum term in stationary sequences. *Ann. Math. Statist.* **35**, 502–516.

[24] Berman, S. M. (1964) Limiting distribution of the maximum of a diffusion process. *Ann. Math. Statist.* **35**, 319–329.

[25] Berman, S. M. (1982) Sojourns and extremes of stationary processes. *Ann. Probab.* **10** (1), 1–46.

[26] Berman, S. M. (1992) *Sojourns and Extremes of Stochastic Processes*. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove.

[27] Bertoin, J. (1996) *Lévy Processes*. Cambridge University Press, Cambridge.

[28] Billingsley, P. (1999) *Convergence of Probability and Measures*. 2nd ed. Wiley, New York.

[29] Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1987) *Regular Variation*. Cambridge University Press, Cambridge.

[30] Bondesson, L. (1988) Shot-noise processes and distributions. In: *Encyclopedia of Statistical Science*, vol. 8, pp. 448–452. Wiley, New York.

[31] Bondesson, L. (1992) *Generalized Gamma Convolutions and Related Classes of Distributions and Densities*, vol. 76 of *Lecture Notes in Statistics*. Springer, New York.

[32] Borkovec, M. and Klüppelberg, C. (1998) Extremal behavior of diffusions models in finance. *Extremes* **1** (1), 47–80.

[33] Braverman, M. and Samorodnitsky, G. (1995) Functionals of infinitely divisible stochastic processes with exponential tails. *Stochastic Process. Appl.* **56** (2), 207–231.

[34] Breiman, L. (1965) On some limit theorems similar to the arc-sin law. *Theory Probab. Appl.* **10** (2), 323–331.

[35] Brockwell, P. J. (2000) Heavy-tailed and non-linear continuous-time ARMA models for financial time series. In: W. S. Chan, W. K. Li, and H. Tong (Eds.), *Statistics and Finance*. Imperial College Press, London.

[36] Brockwell, P. J. (2001) Lévy-driven CARMA processes. *Ann. Inst. Statist. Math.* **53** (1), 113–123.

[37] Brockwell, P. J., Chadraa, E., and Lindner, A. M. Continuous time GARCH processes of higher order. In preparation.

[38] Brockwell, P. J. and Davis, R. A. (1991) *Time Series: Theory and Methods*. 2nd ed. Springer, New York.

[39] Brockwell, P. J. and Marquardt, T. (2003) Levy-driven and fractionally integrated ARMA processes with continuous time parameter. Preprint, available at `http://www.ma.tum.de/stat/`.

[40] Buchmann, B. and Klüppelberg, C. Extreme value theory for stochastic processes driven by fractional Brownian motion. In preparation.

[41] Buchmann, B. and Klüppelberg, C. (2004) Fractional integral equation and state space transforms. Preprint, available at `http://www.ma.tum.de/stat/`.

[42] Chatterji, S. D. (1969) An L^p-convergence theorem. *Ann. Math. Stat.* **40** (3), 1068–1070.

[43] Chover, J., Ney, P., and Wainger, S. (1973) Degeneracy properties of subcritical branching processes. *Ann. Probab.* **1** (4), 663–673.

[44] Cline, D. B. H. (1986) Convolution tails, product tails and domains of attraction. *Probab. Theory Related Fields* **72**, 529–557.

[45] Cline, D. B. H. (1987) Convolution of distributions with exponential and subexponential tails. *J. Austral. Math. Soc. Ser. A* **43** (3), 347–365.

[46] Cline, D. B. H. and Samorodnitsky, G. (1994) Subexponentiality of the product of independent random variables. *Stochastic Process. Appl.* **49** (1), 75–98.

[47] Cont, R. and Tankov, P. (2004) *Financial Modelling with Jump Processes*. Chapman & Hall, Boca Raton.

[48] Cramér, H. (1965) A limit theorem for the maximum values of certain stochastic processes. *Theory Probab. Appl.* **10** (126-128).

[49] Cramér, H. (1966) On intersections between the trajectories of normal stationary sequences and a high level. *Ark. Mat.* **6** (337-349).

[50] Daley, D. J. and Vere-Jones, D. (2003) *An Introduction to the Theory of Point Processes*, vol. I: Elementary Theory and Methods. 2nd ed. Springer, New York.

[51] Davis, R. (1982) Maximum and minimum of one-dimensional diffusions. *Stochastic Process. Appl.* **13**, 1–9.

[52] Davis, R. and Hsing, T. (1995) Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Probab.* **23** (2), 879–917.

[53] Davis, R. and Mikosch, T. (1998) The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Ann. Statist.* **26** (5), 2049–2080.

[54] Davis, R. and Resnick, S. (1985) Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.* **13** (2), 179–195.

[55] Davis, R. and Resnick, S. (1988) Extremes of moving averages of random variables from the domain of attraction of the double exponential distribution. *Stochastic Process. Appl.* **30** (1), 41–68.

[56] De Haan, L. (1970) *On Regular Variation and its Application to the Weak Convergence of Sample Extremes*, vol. 32 of *Mathematical Centre Tracts*. Mathematical Centre, Amsterdam.

[57] De Haan, L., Resnick, S. I., Rootzén, H., and Vries, C. G. (1989) Extremal behavior of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Process. Appl.* **32**, 213–224.

[58] Doney, R. A. and O'Brien, G. (1991) Loud shot noise. *Ann. Appl. Probab.* **1** (1), 88–103.

[59] Drost, F. and Werker, B. J. M. (1996) Closing the GARCH gap: continuous time GARCH modelling. *J. Econometrics* **74**, 31–57.

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

[61] Engle, R. F. (1995) ARCH: selected readings. Oxford University Press, Oxford.

[62] Fasen, V. (2004) Extremes of regular varying mixed moving average processes. In preparation.

[63] Fasen, V. (2004) Extremes of subexponential Lévy driven moving average processes. Preprint, available at `http://www.ma.tum.de/stat/`.

[64] Fasen, V., Klüppelberg, C., and Lindner, A. (2004) Extremal behavior of stochastic volatility models. Preprint, available at `http://www.ma.tum.de/stat/`.

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

[66] 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 Philos. Soc.* **24**, 180–190.

[67] Gnedenko, B. V. (1943) Sur la distribution limité du terme d'une série aléatoire. *Ann Math.* **44**, 423–453.

[68] Gänssler, P. and Stute, W. (1977) *Wahrscheinlichkeitstheorie*. Springer, Berlin.

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

[70] Goldie, C. M. and Klüppelberg, C. (1998) Subexponential distributions. In: R. J. Adler and R. E. Feldman (Eds.), *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, pp. 435–459. Birkhäuser, Boston.

[71] Goldie, C. M. and Resnick, S. (1988) Distributions that are both subexponential and in the domain of attraction of an extreme-value distribution. *Adv. Appl. Probability* **20** (4), 706–718.

[72] Goldie, C. M. and Resnick, S. (1988) Subexponential distribution tails and point processes. *Commun. Statist.-Stochastic Models* **4** (2), 361–372.

[73] Gumbel, E. J. (1958) *Statistics of Extremes*. Columbia Univ. Press, New York.

[74] Gushchin, A. A. and Küchler, U. (2000) On stationary solutions of delay differential equations driven by a Lévy process. *Stochastic Process. Appl.* **88** (2), 195–211.

[75] Homble, P. and McCormick, W. P. (1993) Weak limit results for the extremes of a class of shot noise processes. TR#93–10, Department of Statistics, University Georgia, Athen.

[76] Homble, P. and McCormick, W. P. (1995) Weak limit results for the extremes of a class of shot noise processes. *J. Appl. Probab.* **32** (3), 707–726.

[77] Hsing, T. (1993) On some estimates based on sample behavior near high level excursions. *Probab. Theory Relat. Fields* **95**, 331–356.

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

[79] Hult, H. and Lindskøg, F. (2004) Extremal behavior for regularly varying stochastic processes. *To appear in Stoch. Process. Appl.* .

[80] Kallenberg, O. (1982) *Random Measures*. Akademie-Verlag, Berlin.

[81] Kallenberg, O. (1997) *Foundations of Modern Probability*. Springer, New York.

[82] Karlin, A. and Taylor, H. M. (1975) *A First Course in Stochastic Processes*. 2nd ed. Acad. Press.

[83] Klüppelberg, C. (1989) Subexponential distributions and characterizations of related classes. *Probab. Theory Relat. Fields* **82**, 259–269.

[84] Klüppelberg, C. (2004) Risk management with extreme value theory. In: B. Finkenstädt and H. Rootzén (Eds.), *Extreme Values in Finance, Telecommunication and the Environment*, pp. 101–168. Chapman & Hall/CRC, Boca Raton.

[85] Klüppelberg, C. and Kühn, C. (2003) Fractional Brownian motion as a weak limit of Poisson shot noise processes – with applications to finance. *To appear in Stoch. Process. Appl.* .

[86] Klüppelberg, C. and Lindner, A. (2003) Extreme value theory for moving average processes with light-tailed innovations. Preprint, available at `http://www.ma.tum.de/stat/`.

[87] Klüppelberg, C., Lindner, A., and Maller, R. (2004) A continuous time GARCH process driven by a Lévy process: stationarity and second order behaviour. *J. Appl. Probab.* **41** (3), 601–622.

[88] Klüppelberg, C., Lindner, A., and Maller, R. (2004) Continuous time volatility modelling: COGARCH versus Ornstein-Uhlenbeck models. In: *Invited paper for the Bachelier Colloquium 2005*.

[89] Klüppelberg, C. and Mikosch, T. (1995) Delay in claim settlement and ruin probability approximations. *Scand. Actuar. J.* **2**, 154–168.

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

[91] Kokoszka, P. S. and Taqqu, M. S. (1995) Fractional ARIMA with stable innovations. *Stochastic Process. Appl.* **60**, 19–47.

[92] Kwapień, S. and Woyczyzński, W. A. (1992) *Random Series and Stochastic Integrals: Single and Multiple*. Birkhäuser, Boston.

[93] Leadbetter, M. R. (1974) On extreme values in stationary sequences. *Z. Wahrsch. Verw. Gebiete* **28**, 289–303.

[94] Leadbetter, M. R. (1983) Extremes and local dependence in stationary sequences. *Z. Wahrsch. Verw. Gebiete* **65**, 291–306.

[95] Leadbetter, M. R., Lindgren, G., and Rootzén, H. (1983) *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York.

[96] Leadbetter, M. R. and Rootzén, H. (1982) Extreme value theory for continuous parameter stationary processes. *Z. Wahrsch. Verw. Gebiete* **60** (1), 1–20.

[97] Leadbetter, M. R. and Rootzén, H. (1988) Extremal theory for stochastic processes. *Ann. Probab.* **16** (2), 431–478.

[98] LEBEDEV, A. V. (2000) Extremes of subexponential shot noise. *Math. Notes* **71** (2), 206–210.

[99] LINDSKØG, F. (2004) *Multivariate Extremes and Regular Variation for Stochastic Processes*. Ph.D. thesis, ETH Zürich.

[100] LINDVALL, T. (1973) Weak convergence of probability measures and random functions in the function space $\mathbb{D}[0,\infty)$. *J. Appl. Probab.* **10**, 109–121.

[101] LOYNES, R. M. (1965) Extreme values in uniformly mixing stationary stochastic processes. *Ann. Math. Statist.* **36**, 993–999.

[102] MACKEY, M. (1989) Commodity price fluctuations: price dependent delays and nonlinearities as explanatory factor. *J. Econom. Theory* **48**, 497–509.

[103] MADAN, D. B. AND SENETA, E. (1990) The variance gamma (VG) model for share market returns. *J. Bus.* **63**, 511–524.

[104] MANDELBROT, B. B. (1963) The variation of certain speculative prices. *J. Business* **36**, 394–419.

[105] MANDL, P. (1968) *Analytical Treatment of One-Dimensional Markov Processes*. Springer, New York.

[106] MARUYAMA, G. (1970) Infinitely divisible processes. *Theory Probab. Appl.* **15** (1), 1–22.

[107] MCCORMICK, W. AND SEYMOUR, L. (2000) Extreme values for a class of shot-noise processes. In: I. V. Basawa, C. C. Heyde, and R. L. Taylor (Eds.), *Selected Proceedings of the Symposium on Inference for Stochastic Processes*, vol. 37 of *Lecture Notes-Monograph Series*, pp. 33–45. Beachwood, Institute of Mathematical Statistics.

[108] MCCORMICK, W. P. (1997) Extremes for shot noise processes with heavy tailed amplitudes. *J. Appl. Probab.* **34** (2), 643–656.

[109] MCDONALD, N. (1978) *Time Lags in Biological Models*. Lecture Notes in Biomathematics. Springer, Berlin.

[110] MIKOSCH, T. (1999) Regular variation, subexponentiality and their applications in probability theory. Heavy-Tailed Queuing Workshop in Eindhoven, University of Groningen.

[111] Mikosch, T. (2004) *Non-Life Insurance Mathematics*. Springer.

[112] Mikosch, T. and Samorodnitsky, G. (2000) The supremum of a negative drift random walk with dependent heavy-tailed steps. *Ann. Appl. Probab.* **10** (3), 1025–1064.

[113] Mikosch, T. and Stărică, C. (2000) Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. *Ann. Statist.* **28**, 1427–1451.

[114] Newell, G. F. (1962) Asymptotic extreme value distribution for one dimensional diffusion processes. *J. Math. Mech.* pp. 481–469.

[115] Pakes, A. G. (2004) Convolution equivalence and infinite divisibility. *J. Appl. Probab.* **41** (2), 407–424.

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

[117] Pedersen, J. (2003) The Lévy-Ito decomposition of an independently scattered random measure. Tech. Rep. 2003-2, Centre for Mathematical Physics and Stochastics, University of Aarhus.

[118] Pickands, J. (1969) Asymptotic properties of the maximum in a stationary gaussian process. *Trans. Amer. Math. Soc.* **145**, 75–86.

[119] Pickands, J. (1969) Upcrossing probabilities for stationary Gaussian processes. *Trans. Amer. Math. Soc.* **145**, 52–73.

[120] Pollard, D. (1984) *Convergence of Stochastic Processes*. Springer, New York.

[121] Rajput, B. S. and Rosinski, J. (1989) Spectral representations of infinitely divisible processes. *Probab. Theory Related Fields* **82** (3), 453–487.

[122] Reiss, M. (2002) *Nonparametric Estimation for Stochastic Delay Differential Equations*. Ph.D. thesis, Humboldt–Universität zu Berlin.

[123] Resnick, S. and Willekens, E. (1991) Moving averages with random coefficients and random coefficient autoregressive models. *Commun. Statist.-Stochastic Models* **7** (4), 511–525.

[124] Resnick, S. I. (1986) Point processes, regular variation and weak convergence. *Adv. Appl. Probability* **18** (6), 66–138.

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

[126] Resnick, S. I. (2002) *Adventures in Stochastic Processes*. 3rd ed. Birkhäuser, Boston.

[127] Rice, S. O. (1939) The distribution of the maximum of a random curve. *Amer. J. Math.* **61**, 409–416.

[128] Rice, S. O. (1944) Mathematical analysis of random noise. *Bell. System Techn. J.* **23**, 282–332.

[129] Rice, S. O. (1945) Mathematical analysis of random noise. *Bell. System Techn. J.* **24**, 46–156.

[130] Rootzén, H. (1978) Extremes of moving averages of stable processes. *Ann. Probab.* **6** (5), 847–869.

[131] Rootzén, H. (1986) Extreme value theory for moving average processes. *Ann. Probab.* **14** (2), 612–652.

[132] Rosinski, J. and Samorodnitsky, G. (1993) Distributions of subadditive functionals of sample paths of infinitely divisible processes. *Ann. Probab.* **21** (2), 996–1014.

[133] Samorodnitsky, G. Extreme value theory, ergodic theory, and the boundary between short memory and long memory for stationary stable processes. Preprint, available at `http://www.orie.cornell.edu/~gennady/`.

[134] Samorodnitsky, G. Maxima of continuous time stationary stable processes. Preprint, available at `http://www.orie.cornell.edu/~gennady/`.

[135] Samorodnitsky, G. (1996) A class of shot noise models for financial applications. In: C. C. Heyde, Y. V. Prohorov, R. Pyke, and S. T. Rachev (Eds.), *Proceedings of Athens International Conference on Applied Probability and Time Series Analysis*, 1, pp. 332–351.

[136] Samorodnitsky, G. (1998) Tail behavior of some shot noise processes. In: R. Adler and R. E. Feldman (Eds.), *A Practical Guide to Heavy Tails: Statistical techniques and applications*, pp. 473–486. Birkhäuser, Boston.

[137] Samorodnitsky, G. and Taqqu, M. S. (1994) *Stable Non-Gaussian Random Processes*. Chapman & Hall, New York.

[138] Sato, K. I. (1999) *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge.

[139] Shephard, N. (2004) Stochastic volatility: selected readings. Oxford University Press, Oxford. To appear.

[140] Urbanic, K. and Woyczynski, W. A. (1967) Random integrals and Orlicz spaces. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* pp. 161–169.

[141] Urbanic, K. and Woyczynski, W. A. (1968) Random measures and harmonizable sequences. *Studia Math.* pp. 61–88.

[142] Vervaat, W. (1979) On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. Appl. Probability* **11**, 750–783.

[143] Watson, G. S. (1954) Extreme values in samples from m-dependent stationary stochastic processes. *Ann. Math. Statist.* **25**, 798–800.

[144] Willekens, E. (1986) *Hogere Orde Theorie voor Subexponentiële Verdelingen*. Ph.D. thesis, Katholieke Universiteit Leuven.

Index

Notation

The notation follows the usual conventions, nevertheless some standard abbreviations are gathered in the first table. General mathematical symbols that are used are presented in the second table. The third table summarizes frequently used quantities from the text.

Abbreviations

a. s.	almost surely
ARMA process	autoregressive moving average process
BDLP	background driving Lévy process
CARMA process	continuous time ARMA process
d. f.	distribution function
d. f. s	distribution functions
e. g.	for example (exempli gratia)
FICARMA process	fractional integrated CARMA process
GCIR model	generalized Cox-Ingersoll-Ross model
i. d.	infinitely divisible
i. d. i. s. r. m.	infinitely divisible independently scattered random measure
i. e.	that is (id est)
i. i. d	independent identical distributed
LLN	law of large numbers
MA process	moving average process
OU process	Ornstein-Uhlenbeck process
r. h. s.	right hand side
r. v.	random variable
r. v. s	random variables

SDDE	stochastic delay differential equation
SDE	stochastic differential equation
supOU	superposition of Ornstein-Uhlenbeck processes
w. l. o. g.	without loss of generality

General Symbols

$A := B$	A is defined by B
$[a,b]$, (a,b), $(a,b]$, $[a,b)$	closed, open, half-open interval from a to b
$\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$	$\{1,2,\ldots\}$, $\{0,1,2,\ldots\}$, $\{\ldots,-1,0,1,\ldots\}$
$\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_-$, $\overline{\mathbb{R}}$	$(-\infty,\infty)$, $[0,\infty)$, $(-\infty,0]$, $[-\infty,\infty]$
$\mathbb{C}$	complex numbers
$\Re(z)$, $\Im(z)$	real part and imaginary of $z \in \mathbb{C}$
$\lfloor x \rfloor$	largest integer smaller or equal to $x \in \mathbb{R}$
$\lceil x \rceil$	smallest integer larger or equal to $x \in \mathbb{R}$
$a \vee b$, $a \wedge b$	maximum, minimum of a and b
a^+	$a^+ = 0 \vee a$
a^-	$a^- = 0 \vee -a$
$\log^+(a)$	$\log^+(a) = \log(\max\{a,1\})$
$a \ll b$, $a \approx b$	a is much smaller than, approximately equal to b
$\lvert\mathbf{x}\rvert$	Euclidean norm of $x \in \mathbb{R}^d$
$\mathrm{card}(S)$	cardinality of the set S
$\mathrm{supp}(f)$	support of f
$A \subseteq B$	A is contained in B or $A = B$
∂A	boundary of the set A
f', f'', $f^{(m)}$	first, second, m-fold derivative of f
$f\rvert_A$	function f restricted to the set A
$\mathbf{1}_A$	indicator function of the set A
log, exp	natural logarithm, exponential function
$\mathbb{P}$, $\mathbb{E}$, Var, Cov	probability, expected value, variance and covariance
$\mathbf{0}$	$(0,\ldots,0) \in \mathbb{R}^d$
$\mathbf{e}_j$	the j^{th} unit vector
$\mathbf{x}^t$	the transposed of the vector $\mathbf{x}$
$\lVert A \rVert$	row-sum norm of matrix A
ε_x	Dirac measure at x
$\mathbb{S}^{d-1}$	$\{\mathbf{x} \in \mathbb{R}^d : \lvert\mathbf{x}\rvert = 1\}$
$\mathbb{L}^\delta$	space of δ-integrable functions

$\mathbb{L}^\delta(\pi)$	space of functions $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ measurable with $\int_{\mathbb{R}_+\times\mathbb{R}} \|f(r,s)\|^\delta ds\,\pi(dr) < \infty$
$C(I)$	space of continuous functions $f : I \to \mathbb{R}$
$C_c(I)$	space of continuous functions $f : I \to \mathbb{R}$ with compact support
$C^1(I)$	space of continuous differentiable functions $f : I \to \mathbb{R}$
$\mathbb{D}(I)$	space of function $f : I \to \mathbb{R}$, which are right continuous with left hand limits
$\mathbb{S}^{d-1}$	unit sphere in $\mathbb{R}^d$

Specific Symbols

$X \stackrel{d}{=} Y$	the distribution of X coincides with the distribution of Y
$B = \{B(t)\}_{t\geq 0}$	Brownian motion
D	differential operator
$\overline{F}$	right tail of the distribution function F
F^{*2}	convolution $F * F$ of the distribution function F
F_Z	distribution function of random variable Z
$\stackrel{v}{\Longrightarrow}$	vague convergence
$\stackrel{w}{\Longrightarrow}$	weak convergence
$\stackrel{n\to\infty}{\Longrightarrow}$	weak convergence for $n \to \infty$
$\mathcal{F}_s$	collection of bounded non negative step functions with bounded support on $\overline{\mathbb{R}}^d\backslash\{\mathbf{0}\}$
$g(t) \sim h(t)$	$g(t)/h(t) \stackrel{t\to\infty}{\longrightarrow} 1$
$\sum \varepsilon_{\mathbf{x}_k}$	point process
$\mathcal{L}(\gamma)$	class of long tailed distributions
$L = \{L(t)\}_{t\geq 0}$	Lévy process
(m, σ^2, ν)	generating triplet of a Lévy process
(m, σ^2, ν, π)	generating quadruple of an i. d. i. s. r. m.
$\mathrm{MDA}(G)$	maximum domain of attraction of G
$M(T)$	$M(T) = \sup_{0\leq t\leq T} Y(t)$ for a stochastic process $\{Y(t)\}_{t\geq 0}$
M_n	$M_n = \max_{k=1,\dots,n} X_k$ for a stochastic process $\{X_k\}_{k\in\mathbb{N}}$
$M_P(S)$	class of integer-valued Radon measures on S
$\mathrm{PRM}(\vartheta)$	Poisson random measure with mean measure ϑ
$\mathcal{R}_{-\infty}$	space of rapidly varying functions
$\mathcal{R}_\alpha$	space of regularly varying functions of index α
$\mathcal{S}$	class of subexponential distributions

$\mathcal{S}(\gamma)$	class of convolution equivalent tails
$S_\alpha(c, \beta, \tau)$	α-stable distribution with dispersion c, skewness β and location τ
$\nu(A)$	Lévy measure of the set A