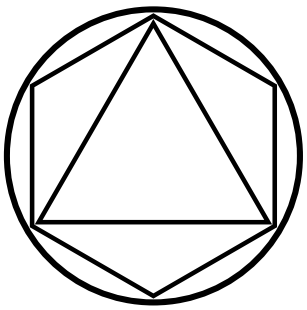
Extremal Behavior of Multivariate Mixed Moving Average Processes and of Random Walks with Dependent Increments

Martin Moser



Zentrum Mathematik
Technische Universität München
D-85748 Garching

2012

Technische Universität München
Lehrstuhl für Mathematische Statistik

Extremal Behavior of Multivariate Mixed Moving Average Processes and of Random Walks with Dependent Increments

Martin Moser

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

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Matthias Scherer
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Robert Stelzer
	2.	Univ.-Prof. Dr. Nina Gantert
	3.	Prof. Dr. Gennady Samorodnitsky Cornell University, USA

Die Dissertation wurde am 15.12.2011 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 21.02.2012 angenommen.

To Michaela

Zusammenfassung

In dieser Arbeit werden verschiedene stochastische Modelle im Hinblick auf ihr extremales Verhalten untersucht. Der Begriff extremales Verhalten meint hier insbesondere die reguläre Variation. Das Konzept der regulären Variation ist weit verbreitet und kann sowohl auf eindimensionale, als auch mehrdimensionale und funktionale Zufallsvariablen angewandt werden.

Zunächst werden die maximalen Zuwächse von zufälligen Irrfahrten analysiert. Es wird gezeigt, dass die Verteilungsfunktion dieser maximalen Zuwächse gegen eine Fréchet Verteilung konvergiert, wenn die Sprünge der zugehörigen Irrfahrt regulär variierend sind. Dieses Resultat wird sowohl für Irrfahren mit linear abhängigen Sprüngen bewiesen, als auch für Irrfahren mit allgemein abhängigen Sprüngen, wenn zusätzliche Mischungsbedingungen erfüllt sind.

Weiterhin wird auch die allgemeine Klasse der mehrdimensionalen Mixed Moving Average (MMA) Prozesse behandelt. Diese Klasse beinhaltet viele interessante Prozesse, wie zum Beispiel CARMA Prozesse, Ornstein-Uhlenbeck (OU) Prozesse und Überlagerungen von Ornstein-Uhlenbeck (supOU) Prozessen. Es wird gezeigt, dass die endlich dimensionalen Verteilungen solcher MMA Prozesse regulär variierend sind, wenn die treibende Lévy Basis regulär variierend ist und die Kernfunktion eine Integrierbarkeitsbedingung erfüllt. Es wird außerdem bewiesen, dass ein MMA Prozess sogar regulär variierend in einem funktionalen Sinn ist, wenn die Kernfunktion einige zusätzliche Bedingungen erfüllt.

Als Spezialfälle von MMA Prozessen werden auch die multivariaten supOU Prozesse beleuchtet und im Hinblick auf ihr extremales Verhalten untersucht. Die Bedingungen für die reguläre Variation, die für MMA Prozesse hergeleitet wurden, werden in diesem Spezialfall verifiziert und somit können geeignetere Bedingungen angegeben werden. Außerdem analysieren wir das extremale Verhalten des stochastischen Volatilitätsmodells vom Typ supOU, bei dem die Volatilität durch einen positiv semidefiniten supOU Prozess beschrieben wird.

Abstract

In this thesis, several different stochastic models are analyzed with respect to their tail behavior. Here, tail behavior is understood in terms of regular variation. The concept of regular variation is well known and can be applied to univariate, multivariate and functional settings.

The first model to be analyzed is the maximum increment of a random walk. It is shown that the distribution of the maximum increment converges to a Fréchet distribution if the jump times of the random walk are regularly varying. The result is proved for random walks with linearly dependent increments and for general dependent increments if additional mixing conditions hold.

Furthermore, we consider the general class of multivariate mixed moving average (MMA) processes. This class includes many interesting processes such as CARMA processes, Ornstein-Uhlenbeck (OU) processes and superpositions of Ornstein-Uhlenbeck (supOU) processes. It is shown that the finite-dimensional distributions of an MMA processes are regularly varying if the driving Lévy basis is regularly varying and the kernel function satisfies an integrability condition. Moreover if the kernel function satisfies some additional conditions, it can be proved that the MMA process is also regularly varying in a functional sense.

As a special case of MMA processes, multivariate supOU processes are also analyzed in detail with respect to their tail behavior. The regular variation conditions derived for MMA processes are verified and more suitable conditions are given for supOU processes. Furthermore, the tail behavior of the supOU type stochastic volatility model, where the volatility is modeled by a positive semi-definite supOU process, is also considered.

Acknowledgments

At this place, I would like to take the space and time to thank all those people, without whom this thesis would not have been possible.

The first person I want to thank is my advisor, Prof. Dr. Robert Stelzer. I am very grateful to him for having given me the opportunity to write this thesis and to present my work at renowned international conferences. Furthermore, I want to express my special thanks to him for all the helpful advice and support he gave me throughout the years.

I also want to thank Prof. Dr. Thomas Mikosch for all the support he gave me and for all the fruitful discussions we had during my stay at the University of Copenhagen. It was a pleasure to work with him. I am also very grateful to him and his colleagues at the Department of Mathematics for all the hospitality they offered me and that made my stay so enjoyable.

My sincere thanks goes also to my colleagues and friends at the Chair for Mathematical Statistics at TUM for all the support, advice, inspiring debates, encouragement, patience, company and entertainment they offered me during the years, without ever expecting anything in exchange.

Moreover, I gratefully acknowledge the financial support of the TUM Institute for Advanced Study (IAS) in cooperation with the International Graduate School of Science and Engineering (IGSSE). They made not only this thesis possible, but their support also allowed me to participate in conferences, to visit the University of Copenhagen for several months and to attend several trainings in different fields.

Finally, I want thank the most important people in my life: my family and friends, without whom life wouldn't be worth half as much. Thank you for all your time, your patience, your unconditional support and your helpful advice.

And last, but not least, I thank you, Michaela, for all your love, your infinite support and your implicit confidence in me. You make every day of my life a better day!

Contents

Chapter 1

Introduction

In recent decades, there has been an increasing demand for heavy tailed models, especially in applications such as insurance, finance, meteorology and others. This led to the development of extreme value theory and related fields. However, the notion of an "extreme value" is not very precise. Intuitively, an extreme value of a time series is a value which is relatively high compared to the others. Then, a heavy tailed random time series is one where extreme values occur with a relatively high probability. Before making this idea mathematically precise, let us consider several examples showing the importance of extreme value models.

In meteorology, statistical methods are used to measure and model weather impacts such as rainfall, air pressure, wind, temperature etc. The results are used to get a better understanding of the interrelations between the different factors affecting the weather and, as a consequence, to be able to make better forecasts. Although this is already a very important field of research itself, the real threat is posed by extreme weather events. Events of that kind can be, for example, huge storms, heavy rainfalls or long dry spells and they cause great damages to the affected area and its inhabitants. Thus, a thorough understanding of such extreme events is indispensable and can lead to prevention programs, early warning systems and evacuation plans that help minimizing the damage.

Closely related are insurance companies where customers are insured against weather damages and other events. In particular, reinsurance companies work with so-called excess-of-loss contracts, where they agree to cover all losses of the primary insurer excessing a certain threshold. Hence, the reinsurance companies are mostly interested in the extreme values of the loss distribution and not in characteristics like means or standard deviations describing the center of the distribution. The need of appropriate models is apparent, see also Embrechts et al. (1997).

As a final example we have a look at data networks, where an increasing demand for

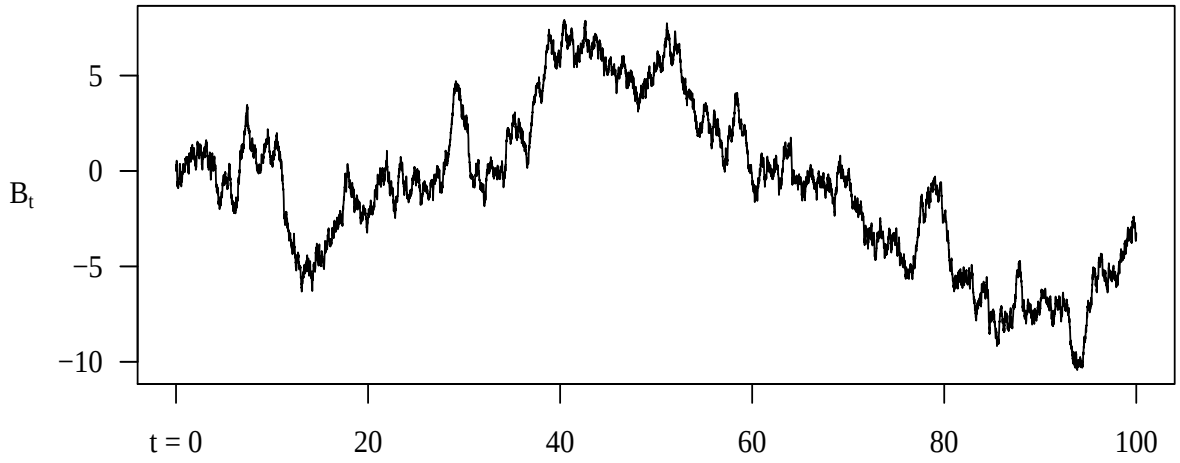


Figure 1.0.1: Simulation of one path of a Brownian motion B_t

heavy tailed models has occurred in recent years. In such networks, data packets of variable size arrive at the nodes of the network at different times. The transmission duration of a data packet at a certain node depends mainly on the packet size as well as on the transmission rate of the node. If too many packets of a certain size arrive in a small time interval, it can cause the network to collapse. Thus, in order to avoid collapses, the capacity of a network has to be adjusted according to the expected maximal data traffic. Since empirical measurements show that characteristics such as file sizes, transmission durations etc. are often heavy tailed (cf. Maulik et al. (2002) and Resnick and Rootzén (2000)), the application of appropriate extreme value models is well motivated. For details see also Resnick (2007), Section 8.

Next we show the importance of extreme value modeling in the example of several well-known time series models. Figure 1.0.1 shows a Brownian motion (B_t). A Brownian motion is a stochastic process where the increments $B_t - B_s$ are independent and follow a normal (Gaussian) distribution. Since all moments of the normal distribution are finite, the movements of the Brownian motion can be considered as light tailed, i.e. no extreme movements occur with a high probability. The Brownian motion and Gaussian distributions are well understood and data modeling with related distributions is relatively easy. In contrast, the increments of an α-stable Lévy motion (L_t) are heavy tailed for values $\alpha < 2$. Figure 1.0.2 shows such a Lévy motion which is α-stable with $\alpha = 1.5$. Comparing it to the Brownian motion, we see that the Lévy motion has some additional and relatively

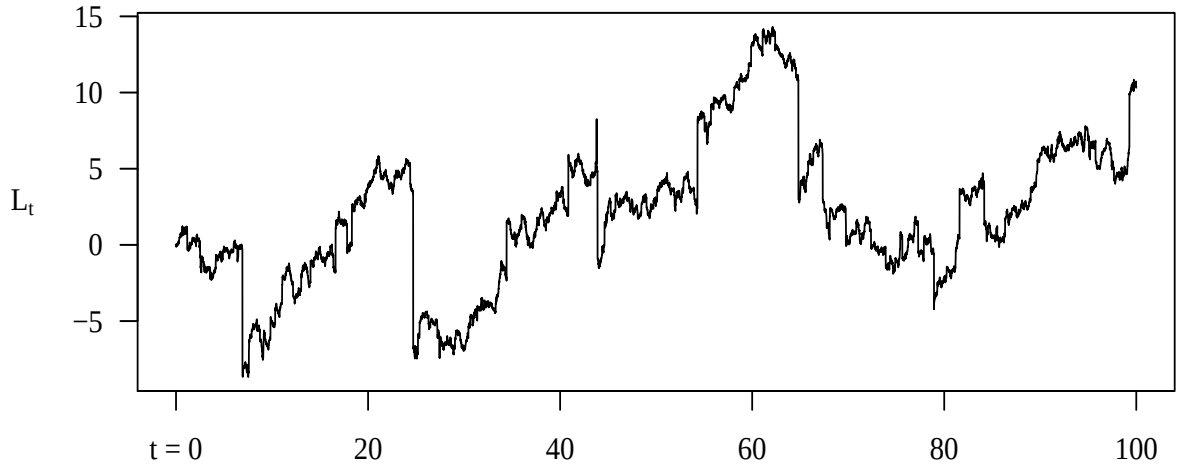


Figure 1.0.2: Simulation of one path of an α-stable Lévy motion L_t with $\alpha = 1.5$

large jumps. Although these jumps are relatively rare, they cannot be neglected in the study of these processes. This gets even more obvious for $\alpha = 1$, which is shown in Figure 1.0.3. In that case, the movement of the process L_t is more or less only determined by one extremely large jump at approximately $t = 52$. This means that extreme values can massively influence a time series, even if they are rare events. Consequently, there is a need for heavy tailed models in places where Gaussian and other light tailed distributions seem to be an inappropriate choice. On the other hand, the major problem of extreme value statistics is also obvious: in general, as extremes are rare events, there are only very few data points that can be used for the statistical analysis.

The most basic extremal characteristic of a distribution is the maximum. For a sequence of random variables (X_i), $i \in \mathbb{N}$, the sample maximum is given by

$$M_n = \max\{X_1, \ldots, X_n\}, \qquad n \in \mathbb{N}.$$

The properties of M_n have been thoroughly studied and are well understood. In particular, if (X_i) is a sequence of independent and identically distributed random variables, it follows from the classical result of the Fisher-Tippett theorem (cf. Embrechts et al. (1997), Theorem 3.2.3) that the limit distribution of the properly normalized and centered sample maxima M_n can only be one of the following three distributions: a Fréchet distribution Φ_α, a Weibull distribution Ψ_α or a Gumbel distribution Λ. These are considered as the

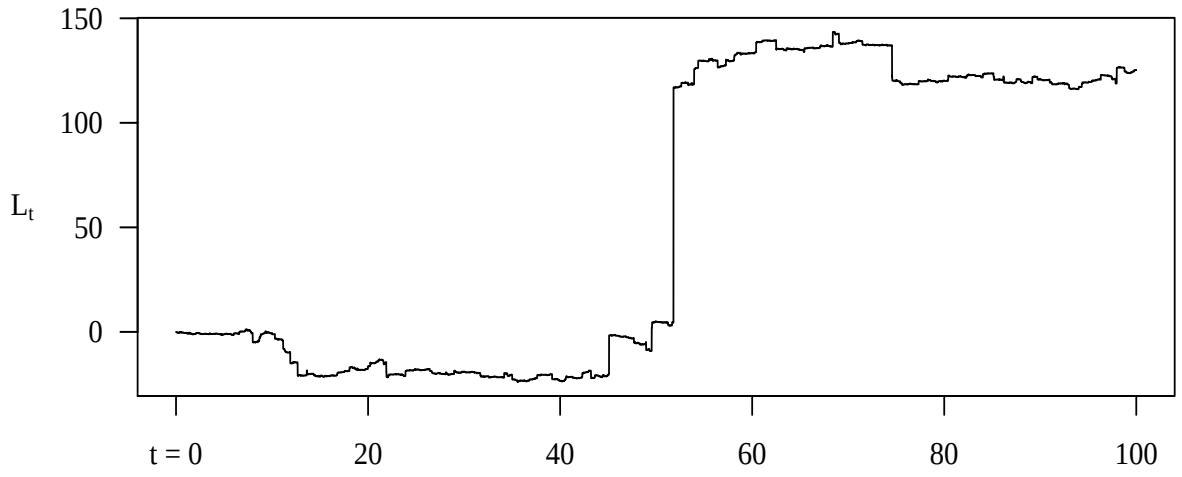


Figure 1.0.3: Simulation of one path of an α-stable Lévy motion L_t with $\alpha = 1$

standard extreme value distributions. In the study of the domains of attraction, one is interested in the following question: which conditions on the distribution function F of X_1 are necessary and sufficient for the maxima M_n to converge to a Fréchet, Weibull or Gumbel distribution? The answer for the Fréchet and Gumbel distribution can be given in terms of regularly varying functions, i.e. functions $f : \mathbb{R}^+ \to \mathbb{R}^+$ satisfying

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

for some index $\alpha > 0$. Then F belongs to the maximum domain of attraction of the Fréchet distribution Φ_α if and only if the function $P(X > x) = 1 - F(x)$ is regularly varying with index α (c.f. Embrechts et al. (1997), Theorem 3.3.7). Furthermore, F belongs to the maximum domain of attraction of the Weibull distribution Ψ_α if and only if the function $P(X > x_F - 1/x) = 1 - F(x_F - 1/x)$ is regularly varying with index α, where x_F is the finite right endpoint (c.f. Embrechts et al. (1997), Theorem 3.3.12). This is the beginning of the long success story of regular variation in extreme value theory.

One of the big advantages of regular variation is the fact that it can be implemented in basically any proper space. It can be defined for real-valued random variables, random vectors, stochastic processes and even for random elements of a general Banach space (cf. Hult and Lindskog (2006b)). Following the above discussion on maxima, a univariate random variable is regularly varying if the tail of the distribution function is a regularly

varying function. For more general spaces, regular variation is set up as convergence to limit measures. A random variable X with values in E is regularly varying if

$$nP\left(a_n^{-1}X \in \cdot\right) \xrightarrow{pro} \mu(\cdot),$$

where the limit measure μ is a nonzero Radon measure with no mass at infinity and (a_n) is an increasing sequence of positive real numbers. Here, $\xrightarrow{pro}$ is understood as convergence in a proper topology depending on the concrete state space E, e.g. vague convergence in finite-dimensional spaces. It turns out that the limit measure μ is homogeneous, i.e. in every direction, the distribution of extreme values behaves like a regularly varying function. On the other hand, μ contains the relevant information about the structure (or directions) of the extremes in space. Therefore, the theory of regular variation is an appropriate tool for the description of extreme values. The concept is well known and applied in many areas of probability theory and statistics. For further details we refer to Resnick (2007) and Hult and Lindskog (2006b).

The Maximum Increment of a Random Walk

In this thesis, the concept of regular variation is applied to several different time series models. One of the most basic models in discrete time is the random walk. For a sequence of increments (X_n), the corresponding random walk (S_n) is given by

$$S_0 = 0 \quad \text{and} \quad S_n = X_1 + \cdots + X_n \quad \text{for} \quad n \geq 1.$$

Hence, the movement of the random walk at time n is described by the random variable X_n. A survey on the general limit theory and (functional) central limit theorems of random walks can be found in Embrechts et al. (1997), Section 2. Random walks play an important role in the modeling of insurance claims, where the ruin probabilities

$$P\Big(\sup_{n\in\mathbb{N}}(S_n - c\,n) > u\Big)$$

are considered for $c > 0$ and $u \to \infty$. See Mikosch and Samorodnitsky (2000a), Mikosch and Konstantinides (2005) and Mikosch and Samorodnitsky (2000b) for the study of these probabilities in several different settings. In this thesis we focus on one extremal characteristic of the random walk: the maximum increment given by

$$\widetilde{M}_n = \max_{1\leq l\leq n}(f(l))^{-1} \max_{0\leq k\leq n-l} |S_{k+l} - S_k|,$$

where $n \in \mathbb{N}$ and f is a proper normalizing function. Quantities of that kind appear as test statistics for the detection of epidemic change points in a time series. See also Csörgö and

Horváth (1997) for a general survey on change point analysis and Račkauskas and Suquet (2004) and Račkauskas and Suquet (2006) for the related limit theory.

Mikosch and Račkauskas (2010) have shown that $\widetilde{M}_n$ converges to a Fréchet limit distribution Φ_α if the jumps (X_i) of the random walk are independent, identically distributed and regularly varying with index α. However, independence is not always an appropriate assumption in real-life applications. Thus, it is always helpful to have similar results for the case of dependent increments. In this thesis we prove limit theorems for the maximum increment of a random walk in several cases of stationary, but dependent jumps which are regularly varying.

First, we study the case of linear dependent jumps, where the increment sequence (X_i) constitutes a finite moving average, i.e.

$$X_t = X_t^{(q)} = \sum_{i=1}^{q} \psi_i Z_{t-i}, \tag{1.0.1}$$

where $(\psi_i)_{i\in\mathbb{N}}$ is a real-valued sequence of coefficients and $(Z_i)_{i\in\mathbb{Z}}$ is an iid noise sequence. We assume that Z_1 is again regularly varying with index α and consequently the same holds true for (X_i). In that case, we can show that $\widetilde{M}_n$ converges to a transformed Fréchet distribution $\Phi_\alpha^{m_q^\alpha}(x)$. Here, m_q is a characteristic based on the moving average coefficients (ψ_i), i.e. the linear structure appears in the limit distribution.

The same result holds for linear processes or moving averages of infinite order. A linear process (X_t) is given by (1.0.1) with $q = \infty$. Assuming almost sure convergence of the linear process, the limit distribution of $\widetilde{M}_n$ is shown to be $\Phi_\alpha^{m_\infty^\alpha}(x)$, where m_∞ is again a constant depending on the linear structure of the process described by the coefficients (ψ_i).

Furthermore, convergence to a (transformed) Fréchet distribution is also proved for a general dependent setting. In this setting, we use the theory of Davis and Hsing (1995), who provided results for the point process convergence of dependent regularly varying sequences that satisfy certain mixing conditions. The general result is then also applied to two examples of heavy tailed time series with multiplicative noise: GARCH processes and stochastic volatility models. In both cases, the conditions of the general setting are verified and thus convergence to a (transformed) Fréchet limit distribution holds.

Multivariate Mixed Moving Average Processes

Another important class of processes being studied in this thesis are multivariate mixed moving average (MMA) processes. They have been first introduced by Surgailis et al.

(1993) and can be given in the integral representation

$$X_t = \int_{M_d} \int_{\mathbb{R}} f(A, t-s)\Lambda(dA, ds),$$

where $f : M_d \times \mathbb{R} \mapsto M_{n,d}$ is a general measurable kernel function and Λ is an $\mathbb{R}^d$-valued Lévy basis. Here, M_d denotes the set of all $d \times d$ matrices and $M_{n,d}$ is the set of all $n \times d$ matrices. Thus, (X_t) is $\mathbb{R}^n$-valued. The class of MMA processes is a rich class including many continuous-time processes used in various areas of application. Examples include Ornstein-Uhlenbeck (OU) processes, superpositions of Ornstein-Uhlenbeck (supOU) processes (cf. Barndorff-Nielsen (2001)), CARMA processes, fractionally integrated CARMA processes (cf. Brockwell (2004) and Marquardt (2007)) and increments of fractional Lévy processes (cf. Marquardt (2006) and Bender et al. (2011)).

The tail behavior of real-valued MMA processes has already been studied by Fasen (2005) and Jacobsen et al. (2009). We extend their results to the multivariate setting. Given the condition that the driving Lévy basis (or driving Lévy process or driving Lévy measure resp.) is regularly varying with some index α, we show that the finite-dimensional distributions of the MMA process are regularly varying with the same index α. Sufficient conditions therefore are, besides existence, an integrability condition $f \in \mathbb{L}^\alpha$ on the kernel function and a non-degeneracy condition. We also derive necessary conditions on the kernel function f, which are very close to the sufficient ones. In fact, in the univariate case both conditions coincide and thus we get necessary and sufficient conditions.

Furthermore, we apply the definition of functional regular variation to MMA processes. Hult and Lindskog (2005) have introduced this notion for processes with càdlàg sample paths and showed that regular variation of the finite-dimensional distributions implies functional regular variation if several relative compactness criteria are satisfied. We verify these conditions for MMA processes under some additional integrability and continuity assumptions on the kernel function f. This shows that MMA processes are also regularly varying in the space of càdlàg functions.

As an example we also study the tail behavior of multivariate supOU processes. The introduction of these processes is motivated by the problem that the well-known OU processes do not account for long memory effects. To overcome this problem, Barndorff-Nielsen (2001) defined univariate supOU processes as the superposition of infinitely many OU processes. This concept has been extended to the multivariate setup by Barndorff-Nielsen and

Stelzer (2011a). Multivariate supOU processes can be given in their integral representation

$$X_t = \int_{M_d^-} \int_{-\infty}^{t} e^{A(t-s)} \Lambda(dA, ds),$$

where Λ is again an $\mathbb{R}^d$-valued Lévy basis and M_d^- is the set of all $d \times d$ matrices with eigenvalues having strictly negative real part. It is clear from this representation that supOU processes are MMA processes with kernel function $f(A, s) = e^{As} \mathbb{1}_{[0,\infty)}(s)$. Hence, the regular variation results derived for MMA processes can be applied.

In particular, we verify the MMA conditions for both, finite-dimensional and functional regular variation, in the special case of supOU processes. This leads to some more accessible sufficient conditions on the supOU kernel function which needs to decay in a proper way. Additionally, we also study the sufficient conditions for regular variation of supOU processes.

One of the main areas of application of OU and supOU processes are stochastic volatility models. In these models, the variance (or volatility) process is modeled by an OU or supOU process. Then, the logarithmic stock price process (X_t) is given by

$$\begin{aligned} dX_t &= a_t dt + \Sigma_t^{1/2} dW_t + \Psi(dL_t) \\ X_0 &= 0, \end{aligned}$$

where a is predictable process, W is the standard Brownian motion, L is the Lévy process associated with Λ and Ψ is a linear operator. Of course, one can also consider more basic versions of the model. Models of that type have been introduced by Barndorff-Nielsen and Shephard (2001) in the univariate OU case and have been extended to the multivariate OU setting by Pigorsch and Stelzer (2009).

In this thesis, we study in detail the tail behavior of the multivariate supOU type stochastic volatility model introduced by Barndorff-Nielsen and Stelzer (2011b), where the volatility process (Σ_t) is modeled by a positive semi-definite supOU process. Based on the preceding results on multivariate MMA and supOU processes, we show that positive-semidefinite supOU processes are regularly varying with index α under similar conditions. As a consequence, it can then be proved that the volatility part $\Sigma_t^{1/2} dW_t$ of the logarithmic stock prices process is regularly varying with index 2α. As regular variation of the drift term $a_t dt$ and the leverage term $\Psi(dL_t)$ is in general easy to obtain for the common choices of these terms, this result yields a good intuition for the understanding of the tail behavior of the stock price process (X_t).

General Outline

The thesis is subdivided into four main chapters.

In Chapter 2 we introduce the relevant theory that is used repeatedly in the later parts of the thesis. Section 2.1 is a survey on regular variation of functions, real-valued random variables, multivariate random vectors and stochastic processes. In Section 2.2 we implement point processes as well as the related weak convergence theory and discuss their use in extreme value theory. Then, in Section 2.3, we review the theory of Lévy processes and infinitely divisible distributions and analyze their tail behavior.

In Chapter 3 which is based on the paper Mikosch and Moser (2012) we examine the limit behavior of the maximum increment of a random walk with regularly varying jumps. After the introduction of some important preliminaries in Section 3.2, we distinguish three different settings of the dependence structure of the random walk increments: the independent setting is reviewed in Section 3.3, the linear dependent setting is analyzed in Section 3.4 and the general dependent setting is dealt with in Section 3.5.

In Chapter 4 which is based on the paper Moser and Stelzer (2011) we study finite-dimensional regular variation of multivariate MMA processes. The essential preliminaries and the related integration theory for MMA processes are introduced in Section 4.2. In Section 4.3 we analyze necessary as well as sufficient conditions for regular variation of MMA processes. The conditions are then verified in the special case of multivariate supOU processes in Section 4.4. Finally, in Section 4.5, we apply the results to the multivariate supOU type stochastic volatility model.

In Chapter 5 we prove functional regular variation of multivariate MMA processes. After the statement of some useful preliminaries in Section 5.2, we first analyze the sample path behavior of MMA processes in Section 5.3. In Section 5.4 we introduce the notion of regular variation for processes with càdlàg sample paths and show that MMA processes are regularly varying in that sense. The results are then applied to supOU processes in Section 5.5. Finally, we discuss the benefit of functional regular variation of MMA processes in view of point process convergence and their extremal behavior in Section 5.6.

General Notation

Given the real numbers $\mathbb{R}$ we use the notation $\mathbb{R}^+$ for the positive real numbers and $\mathbb{R}^-$ for the negative real numbers, both without 0. $\mathbb{N}$ are the natural numbers, $\mathbb{Z}$ the integers and $\mathbb{Q}$ the rational numbers. The respective sets including 0 are denoted by $\mathbb{R}_0^+$, $\mathbb{R}_0^-$ and $\mathbb{N}_0$. The integer part $\lfloor a \rfloor$ rounds a real number $a \in \mathbb{R}$ to the next lowest $z \in \mathbb{Z}$. S_d is the

unit sphere on $\mathbb{R}^d$. By $B_r(x) := \{y \in \mathbb{R}^d : \|y - x\| \leq r\}$ we denote the closed ball of radius r centered at x. The σ-algebra of Borel sets on a space E is given by $\mathcal{B}(E)$ and $\mathcal{B}_b(E)$ is the respective restriction to bounded Borel sets. For a set $A \in \mathcal{B}(E)$, ∂A is the boundary of A and $\overline{A}$ is its closure. $\mathbb{1}$ is the standard indicator function.

$\mathbb{D}$ is the space of càdlàg (right-continuous with left limits) functions $x : [0, 1] \to \mathbb{R}^d$ and $S_{\mathbb{D}} = \{x \in \mathbb{D} : \sup_{t\in[0,1]} \|x_t\| = 1\}$ is the unit sphere in $\mathbb{D}$.

For matrices, $M_{n,d}$ is the set of all $n \times d$ matrices and M_d the set of all $d \times d$ matrices. M_d^- is the set of all $d \times d$ matrices with eigenvalues having strictly negative real part. I_d is the $d \times d$ identity matrix, $\mathbb{S}_d$ denotes the symmetric $d \times d$ matrices and $\mathbb{S}_d^+$ the positive semidefinite $d \times d$ matrices. We write A^T for the transposed of a matrix A and $\|A\|$ for an arbitrary matrix norm. Since all norms on finite-dimensional spaces are equivalent, the type of norm is not important for the results, but if we make no further specifications, we use the operator norm induced by the Euclidean norm.

If random variables, vectors, processes, measures etc. are considered, they are given as measurable mappings with respect to a complete probability space $(\Omega, \mathcal{F}, P)$. Furthermore, λ denotes the univariate Lebesgue measure and $\mathbb{E}\, X$ is the expectation of the random variable X.

Chapter 2

Regular Variation, Point Processes and Lévy Processes

This thesis touches a wide range of topics in probability theory and statistics. Just to name a few, throughout this thesis we make use of extreme value theory, regular variation, vague topology, weak topology, $\hat{w}$-topology, point processes, random measures, Poisson processes, Lévy processes, Lévy bases, random walks, linear processes, stochastic volatility models, GARCH-, MMA-, supOU- and CARMA processes and many more.

Although the general theory of most of these topics is well known and can be found in the standard literature, it seems appropriate to introduce the relevant notation and main results in view of a clear representation of this thesis. Wherever possible, we incorporate such notations, definitions and results into the sections of this thesis where they are used. However, there are three subjects which appear repeatedly throughout this thesis: regular variation, point processes and Lévy processes. Hence, it might be beneficial for the reader to have a more general introduction to these topics.

The three topics correspond to the three sections of this chapter. In each, we first introduce the necessary notation and give the relevant definitions. Moreover, we also state the main results that help in the understanding of the theory and also those that we use in the later parts of this thesis. Most of them are well known and thus we only refer to the relevant literature for proofs. Some results are also proved directly, since the proofs are not to be found in the literature.

The chapter is organized as follows. In Section 2.1 we introduce the concept of regular variation starting with regularly varying functions. Then, step by step, we extend it to real-valued random variables, then to multivariate random variables and, finally, to stochastic processes with càdlàg sample paths. In Section 2.2 we give a survey on the theory of (weak convergence of) point processes and its relevance to extreme value theory. Finally,

the definition, the main properties and the tail behavior of Lévy processes and infinitely divisible distributions are given and analyzed in Section 2.3.

2.1 Regular Variation

The relevance of the concept of regular variation for the characterization of the tail behavior of random variables, vectors, processes etc. has already been discussed in the introduction in Chapter 1. Regular variation can be defined on many different spaces, e.g. for random variables with values in $\mathbb{R}$, $\mathbb{R}^d$ or in a functional space. Even regular variation of random variables with values in a general Banach space can be defined.

The origin of regular variation can be found in the theory of regularly varying real-valued functions, which is introduced in Section 2.1.1. This method is then extended to random variables. In Section 2.1.2 we analyze regular variation of real-valued random variables, where regular variation of the random variable is understood as regular variation of the tail of the distribution function. In Section 2.1.3 multivariate regular variation is formulated in terms of vague convergence of measures. Furthermore, we state some alternative definitions and several important results characterizing multivariate regular variation. In Section 2.1.4 we review the work of Hult and Lindskog (2005), who extended regular variation to the space of stochastic processes with càdlàg sample paths.

2.1.1 Regularly Varying Functions

In the field of extreme value theory, the use of regularly varying functions if essential in many areas. Regularly varying functions are functions acting like a power law in the limit. The theory has its origin in the work of Karamata (1930) and it has been brought to probability theory by Feller (1966). For a detailed survey on the topic, see also Bingham et al. (1987).

Definition 2.1.1 (Regular Variation of Functions). Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be a Lebesgue measurable function. Then f is *regularly varying with index* $\alpha \in \mathbb{R}$ if

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

for every $x > 0$. We write $f \in RV_\alpha$ and α is called the *index of regular variation* or *tail index*. A Lebesgue measurable function $l : \mathbb{R}^+ \to \mathbb{R}^+$ is *slowly varying* if $l \in RV_0$.

The above definition describes regular (and slow) variation at ∞. Similarly, regular variation can also be defined at any point $a \in \mathbb{R}$. If nothing else is specified, we always

consider regular variation at ∞ throughout this thesis. We start with a few easy results that characterize regular varying functions.

Lemma 2.1.2. *Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be a Lebesgue measurable function.*

(i) If there exists a function h such that for all $x > 0$,

$$\lim_{t\to\infty} \frac{f(tx)}{f(t)} = h(x),$$

then $h(x) = x^{-\alpha}$ for some index $\alpha \in \mathbb{R}$ and $f \in RV_\alpha$.

(ii) $f \in RV_\alpha$ if and only if there exists $l \in RV_0$ such that

$$f(x) = x^{-\alpha} l(x).$$

(iii) If $f \in RV_\alpha$, then

$$\lim_{x\to\infty} \frac{log(f(x))}{log(x)} = \alpha$$

and consequently

$$\lim_{x\to\infty} f(x) = \begin{cases} 0 & \text{if } \alpha > 0, \\ \infty & \text{if } \alpha < 0. \end{cases}$$

Proof. For (i) see Proposition 2.3(i) in Resnick (2007). (ii) is a direct consequence of $l(x) := f(x)/x^{-\alpha} \in RV_0$ for every $f \in RV_\alpha$. The proof of (iii) is based on Karamata's representation and can be found in Resnick (2007), Proposition 2.6(i).

□

Example 2.1.3. Examples of slowly varying functions are constant functions, logarithms and iterated logarithms and all functions asymptotic to any of these. By Lemma 2.1.2(ii), functions of the class RV_α can be given as a product of the power law $x^{-\alpha}$ and a slowly varying function, e.g.

$$\begin{aligned} &f(x) = x^{-\alpha}, \\ &f(x) = c\, x^{-\alpha}, \quad c \in \mathbb{R}\backslash\{0\}, \\ &f(x) = x^{-\alpha} \ln(1+x), \\ &f(x) = x^{-\alpha} \ln(\ln(e+x)) \quad \text{and} \\ &f(x) = (x + \ln(1+x))^{-\alpha}. \end{aligned}$$

One of the most essential results for regular variation is Karamata's theorem. It links regular variation of functions to regular variation of integrals.

Theorem 2.1.4 (Karamata's Theorem). *Let f be locally integrable. Then the following statements hold:*

(i) Let $f \in RV_\alpha$ and $\alpha \leq 1$. Then

$$\int_0^x f(t)\,dt \in \mathrm{RV}_{\alpha-1} \quad \text{and} \quad \lim_{x\to\infty} \frac{xf(x)}{\int_0^x f(t)\,dt} = 1-\alpha.$$

(ii) Let $f \in RV_\alpha$ and $\alpha > 1$. Then

$$\int_x^\infty f(t)\,dt \in \mathrm{RV}_{\alpha-1} \quad \text{and} \quad \lim_{x\to\infty} \frac{xf(x)}{\int_x^\infty f(t)\,dt} = \alpha-1.$$

(iii) $f \in RV_{1-\alpha}$ for $\alpha > 0$ if

$$\lim_{x\to\infty} \frac{xf(x)}{\int_0^x f(t)\,dt} = \alpha.$$

(iv) $f \in RV_{1+\alpha}$ for $\alpha > 0$ if

$$\int_x^\infty f(t)\,dt < \infty \quad \text{and} \quad \lim_{x\to\infty} \frac{xf(x)}{\int_x^\infty f(t)\,dt} = \alpha.$$

Proof. See for example Resnick (2007), Theorem 2.1.

□

Karamata's theorem and the related Karamata representation are very powerful tools for the analysis of regularly varying functions. Many results can be derived from these two. As a final example, we present the Potter bounds giving concrete bounds for regularly varying functions.

Lemma 2.1.5 (Potter Bounds). *Let $f \in RV_\alpha$ for $\alpha \in \mathbb{R}$ and let $\varepsilon > 0$. Then there exists t_0 such that for $x \geq 1$ and $t \geq t_0$,*

$$(1-\varepsilon)x^{-\alpha-\varepsilon} < \frac{f(tx)}{f(t)} < (1+\varepsilon)x^{-\alpha+\varepsilon}.$$

Proof. See Resnick (2007), Proposition 2.6 (ii).

□

2.1.2 Real-Valued Random Variables

The notion of regular variation for real-valued random variables is closely linked to regular variation of functions. It is motivated by the classical problem of extreme value theory, the distributional convergence of maxima (cf. Chapter 3 in Embrechts et al. (1997)). For a short introduction to the problem, assume that $(X_i)_{i\in\mathbb{N}}$ is an iid sequence of non-degenerate random variables with common distribution function F and denote by

$$M_n := \max\{X_1, \ldots, X_n\}$$

the *sample maxima* of (X_i). Now one is interested in the question, whether there are sequences of norming constants $c_n > 0$ and $d_n \in \mathbb{R}$ and a (non-degenerate) limiting distribution function H such that, as $n \to \infty$,

$$c_n^{-1}\,(M_n - d_n) \xrightarrow{d} H. \tag{2.1.1}$$

If the limit exists, then by the classical result, the Fisher-Tippett Theorem, H can only be of one of the following three types of distribution functions:

$$\textit{Fréchet distribution}: \qquad \Phi_\alpha(x) = \begin{cases} 0 & \text{if } x \le 0, \\ \exp(-x^{-\alpha}) & \text{if } x > 0 \end{cases} \quad \text{and } \alpha > 0.$$

$$\textit{Weibull distribution}: \qquad \Psi_\alpha(x) = \begin{cases} \exp(-(-x)^{-\alpha}) & \text{if } x \le 0, \\ 1 & \text{if } x > 0 \end{cases} \quad \text{and } \alpha > 0.$$

$$\textit{Gumbel distribution}: \qquad \Lambda(x) = \exp(-e^{-x}) \quad \text{for all } x \in \mathbb{R}.$$

Therefore, the Fréchet-, Weibull- and Gumbel distribution are referred to as the *standard extreme value distributions*. In particular, we say that X (or its distribution function F) belongs to the *maximum domain of attraction* of the standard extreme value distribution H, written $X \in \mathrm{MDA}(H)$ (or $F \in \mathrm{MDA}(H)$), if (2.1.1) is satisfied for H and constants $c_n > 0$ and $d_n \in \mathbb{R}$. In order to describe the maximum domains of attraction of the three standard extreme value distributions, we need the concept of regularly varying functions introduced in Section 2.1.1. By $\overline{F}(x) := 1 - F(x) = P(X > x)$ we denote the *tail* of the distribution F and by

$$x_F := \sup\{x \in \mathbb{R} : F(x) < 1\}$$

we denote the *right endpoint* of F. We start with the result for the Fréchet distribution.

Theorem 2.1.6 (Embrechts et al. (1997), Theorem 3.3.7). *Let $\alpha > 0$. Then $F \in MDA(\Phi_\alpha)$ if and only if $\overline{F}(x) \in RV_\alpha$.*

Examples of the class $\text{MDA}(\Phi_\alpha)$ are Pareto distributions, Cauchy distributions, Burr distributions and α-stable distributions with $\alpha < 2$. A similar result can be obtained for the Weibull distribution.

Theorem 2.1.7 (Embrechts et al. (1997), Theorem 3.3.12). *Let $\alpha > 0$. Then $F \in MDA(\Psi_\alpha)$ if and only if $x_F < \infty$ and $\overline{F}(x_F - 1/x) \in RV_\alpha$.*

Examples of the class $\text{MDA}(\Psi_\alpha)$ are uniform and Beta distributions. Although the conditions for $F \in \text{MDA}(\Lambda)$ (including e.g. normal and lognormal distributions) cannot be described by regular variation (instead they can be derived using von Mises functions, see Section 3.3.3 in Embrechts et al. (1997)), we now see the importance of the concept of regular variation for real-valued random variables.

Definition 2.1.8 (Regular Variation of Random Variables in $\mathbb{R}$). A real-valued random variable X with distribution function F is *regularly varying* with index $\alpha > 0$ if there exists $p, q \in [0, 1]$ with $p + q = 1$ such that

$$\lim_{t\to\infty} \frac{P(X > tx)}{P(|X| > t)} = p\, x^{-\alpha} \qquad \text{and} \qquad \lim_{t\to\infty} \frac{P(X < -tx)}{P(|X| > t)} = q\, x^{-\alpha}, \qquad x > 0. \tag{2.1.2}$$

Condition (2.1.2) is referred to as the *tail balance condition* and α is the *tail index*. Note that if X is regularly varying with index $\alpha > 0$, then

$$P(|X| > x) \in RV_\alpha.$$

The tail coefficients p and q are the univariate equivalent to the so-called spectral measure that we use in the following section in order to introduce the more involved concept of regular variation in $\mathbb{R}^d$.

2.1.3 Multivariate Random Variables

The extension of regular variation to the multivariate setting is not straightforward, since it is not clear, what an extreme value is in a multivariate space. However, Definition 2.1.8 already gives us a good idea if we interpret the coefficients p and q as a probability measure on the univariate unit sphere $S_1 = \{-1, 1\}$. Thus univariate regular variation can be given in terms of vague convergence of measures and this concept can be extended to the multivariate case.

A measure μ on $\mathcal{B}(\mathbb{R}^d)$ is called a *Radon or locally finite measure* if $\mu(B) < \infty$ for all bounded Borel sets $B \in \mathcal{B}_b(\mathbb{R}^d)$. A sequence of Radon measures (μ_n) *converges vaguely* to the Radon measure μ if

$$\int_{\mathbb{R}^d} f(x)\mu_n(dx) \to \int_{\mathbb{R}^d} f(x)\mu(dx)$$

for all continuous functions f with bounded support. We write $\mu_n \xrightarrow{v} \mu$. Convergence in the vague topology is closely linked to weak convergence of measures, which implies vague convergence. On the other hand, if the measures (μ_n) and μ have uniformly bounded total mass, then the notions of vague and weak convergence coincide (cf. Kallenberg (1983), 15.7.6).

Let $\mathcal{K}$ denote the compact subsets of $\mathbb{R}^d$, $\mathcal{G}$ the open subsets of $\mathbb{R}^d$ and let

$$\mathcal{B}_\mu := \left\{B \in \mathcal{B}_b(\mathbb{R}^d) : \mu(\partial B) = 0\right\}.$$

Then we can state the following characterization of vague convergence of measures.

Theorem 2.1.9 (Resnick (2007), Theorem 3.2). *Let μ and (μ_n) be positive Radon measures on $\mathcal{B}(\mathbb{R}^d)$. Then the following statements are equivalent:*

(i) $\mu_n \xrightarrow{v} \mu$.

(ii) $\mu_n(B) \to \mu(B)$ *for all* $B \in \mathcal{B}_\mu$ *that are relatively compact.*

(iii) For all $K \in \mathcal{K}$ *we have*

$$\limsup_{n\to\infty} \mu_n(K) \leq \mu(K)$$

and for all $G \in \mathcal{G}$ *that are relatively compact we have*

$$\liminf_{n\to\infty} \mu_n(G) \geq \mu(G).$$

In terms of regular variation and extreme value theory, one is usually interested in the "big" values of a random variable. This leads to a focus on sets of the form

$$V_r := \{x \in \mathbb{R}^d : \|x\| > r\}, \qquad r > 0,$$

which are bounded away from the origin. However, the sets V_r are in general not relatively compact in the vague topology if we define it on $\mathbb{R}^d$. To overcome this problem, we consider regular variation on the space $\overline{\mathbb{R}}^d\backslash\{0\} = \mathbb{R}^d \cup \{-\infty, \infty\}\backslash\{0\}$, which assures that the sets $B \subseteq V_r$ can be referred to as the relatively compact sets in the vague topology.

This procedure of excluding the origin from the original space is called the *one-point uncompactification* (c.f. Resnick (2007), Section 6.1.3).

We mention that vague convergence can similarly be defined on any other complete, separable metric space E which is locally compact. For further reading on vague convergence we refer to Kallenberg (1983). Very good introductions to vague convergence in view of regular variation can also be found in Resnick (1987), Resnick (2007), Lindskog (2004) and Basrak (2000).

Now we can state multivariate regular variation in terms of vague convergence of measures, where several different but equivalent definitions exist. We start with one that is widely used and can be easily extended to other spaces than $\mathbb{R}^d$.

Definition 2.1.10 (Multivariate Regular Variation). A random vector $X \in \mathbb{R}^d$ is *regularly varying* if there exists a sequence $(a_n)_{n\in\mathbb{N}}$, $0 < a_n \nearrow \infty$, and a nonzero Radon measure μ on $\mathcal{B}(\overline{\mathbb{R}}^d\backslash\{0\})$ such that $\mu(\overline{\mathbb{R}}^d\backslash\mathbb{R}^d) = 0$ and, as $n \to \infty$,

$$nP(a_n^{-1}X \in \cdot) \xrightarrow{v} \mu(\cdot)$$

on $\mathcal{B}(\overline{\mathbb{R}}^d\backslash\{0\})$. Similarly, we call a Radon measure ν regularly varying if (a_n) and μ exist as above such that, as $n \to \infty$,

$$n\,\nu(a_n^{-1}\cdot) \xrightarrow{v} \mu(\cdot).$$

From Theorem 1.14 in Lindskog (2004) we know that the limit measure μ is homogeneous, i.e. there exists $\alpha > 0$ such that

$$\mu(uB) = u^{-\alpha}\mu(B)$$

for every $u > 0$ and $B \in \mathcal{B}(\overline{\mathbb{R}}^d\backslash\{0\})$. Thus we say that X (or ν resp.) is *regularly varying with (tail) index* α and we write $X \in RV(\alpha, (a_n), \mu)$ (or $\nu \in RV(\alpha, (a_n), \mu)$ resp.). The sequence (a_n) can always be chosen as

$$a_n = \inf\{x \geq 0 : P(\|X\| \leq x) \geq 1 - n^{-1}\}\,, \quad n \geq 1\,,$$

and is regularly varying with index $1/\alpha$, i.e. for every $x > 0$ the convergence

$$\lim_{n\to\infty} \frac{a_{\lfloor nx \rfloor}}{a_n} = x^{1/\alpha}$$

holds (cf. Lindskog (2004), Remark 1.13). In analogy to the univariate case we also have

$$P(\|X\| \geq x) \in RV_\alpha.$$

If $d = 1$ and there exists $p \in [0, 1]$ such that

$$\lim_{x\to\infty} \frac{P(X > x)}{P(|X| > x)} = p,$$

then regular variation as given in Definition 2.1.10 coincides with univariate regular variation of Definition 2.1.8. Next we give a very illustrative characterization of multivariate regular variation.

Theorem 2.1.11 (Lindskog (2004), Theorem 1.15). *The random vector $X \in \mathbb{R}^d$ is regularly varying with index $\alpha > 0$ if and only if there exists a probability measure σ on $\mathcal{B}(S_d)$ such that, as $t \to \infty$,*

$$\frac{P\left(\|X\| > tu, X/\|X\| \in \cdot\right)}{P\left(\|X\| > t\right)} \xrightarrow{v} u^{-\alpha}\,\sigma(\cdot)$$

for every $u > 0$.

The probability measure σ in the previous theorem is called the *spectral measure (of regular variation)* of X. This characterization shows nicely, how multivariate regular variation is established. If we decompose the regularly varying vector X into its polar coordinates $(\|X\|, X/\|X\|)$, we see that the radial part and the spherical part behave asymptotically independent. The radial part $\|X\|$ can be described as a regularly varying random variable on the positive real numbers, whereas the spherical part $X/\|X\|$ is described by the probability measure σ.

The following theorem states two more characterizations of multivariate regular variation which can often be found in the literature.

Theorem 2.1.12. *Let X be an $\mathbb{R}^d$-valued random variable. Then the following statements are equivalent.*

(i) $X \in RV(\alpha, (a_n), \mu)$.

(ii) There exists a Radon measure μ on $\overline{\mathbb{R}}^d\backslash\{0\}$ with $\mu(\overline{\mathbb{R}}^d\backslash\mathbb{R}^d) = 0$ and a relatively compact set $E \subset \overline{\mathbb{R}}^d\backslash\{0\}$ such that $tE \in \mathcal{B}_\mu$, $t \in T$, for a dense set $T \subset (0, \infty)$ and, as $t \to \infty$,

$$\mu_t(\cdot) := \frac{P(X \in t\cdot)}{P(X \in tE)} \xrightarrow{v} \mu(\cdot).$$

(iii) There exists an $\alpha > 0$, a slowly varying function l and a nonzero Radon measure μ defined on $\mathcal{B}(\overline{\mathbb{R}}^d\backslash\{0\})$ with $\mu(\overline{\mathbb{R}}^d\backslash\mathbb{R}^d) = 0$ such that, as $u \to \infty$,

$$u^\alpha L(u) P(u^{-1}X \in \cdot) \xrightarrow{v} \mu(\cdot)$$

on $\mathcal{B}(\overline{\mathbb{R}}^d\backslash\{0\})$.

Proof. (i) $\Leftrightarrow$ (ii)
Theorem 2.1.11 in combination with Basrak (2000), Theorem 2.1.8.
(ii) $\Rightarrow$ (iii)
Let X be regularly varying in the sense of (ii). According to Basrak (2000), p. 27, if the condition in (ii) holds for one $E \in \mathcal{B}_\mu$, it also holds for all $E \in \mathcal{B}_\mu$. Choose

$$E := \{x : \|x\| \geq 1\}.$$

Then by definition, there exists some Radon measure μ with $\mu(\overline{\mathbb{R}}^d \backslash \mathbb{R}^d) = 0$ such that, as $t \to \infty$,

$$\frac{P(X \in t\cdot)}{P(X \in tE)} \xrightarrow{v} \mu(\cdot).$$

From Basrak (2000), p. 27, we know that $tE \in \mathcal{B}_\mu$ for all $t > 0$. Furthermore, we define

$$f(x) := P(\|X\| \geq x).$$

We have

$$\frac{f(tx)}{f(x)} = \frac{P(\|X\| \geq tx)}{P(\|X\| \geq xE)} \frac{P(\|X\| \geq xE)}{P(\|X\| \geq x)} \xrightarrow{x\to\infty} \frac{\mu(t\{\|X\| \geq 1\})}{\mu(\{\|X\| \geq 1\})} = t^{-\alpha},$$

since $\mu(tB) = t^{-\alpha}\mu(B)$. It follows that $f(x)$ is regularly varying as a function of x and thus

$$P(X \in tE) = P(\|X\| \geq t) = t^{-\alpha} l(t)$$

for some slowly varying function l. It is easy to show that

$$\widetilde{l}(t) := \frac{1}{l(t)}$$

is again slowly varying. Putting things together, we obtain

$$u^\alpha \widetilde{l}(u) P(u^{-1}X \in \cdot) = \frac{P(u^{-1}X \in \cdot)}{u^{-\alpha} l(u)} = \frac{P(X \in u\cdot)}{P(X \in uE)} \xrightarrow{v} \mu(\cdot).$$

(iii) $\Rightarrow$ (ii)
Let α be positive, l be a slowly varying function and μ a nonzero Radon measure defined on $\mathcal{B}(\overline{\mathbb{R}}^d \backslash \{0\})$ with $\mu(\overline{\mathbb{R}}^d \backslash \mathbb{R}^d) = 0$ and, as $u \to \infty$,

$$u^\alpha l(u) P(u^{-1}X \in \cdot) \xrightarrow{v} \mu(\cdot)$$

on $\mathcal{B}(\overline{\mathbb{R}}^d \backslash \{0\})$. As μ is nonzero and locally finite, there exists a constant $r > 0$ such that, for $E := \{x : \|x\| > r\}$, we have $0 < \mu(E) < \infty$ and $\mu(\partial E) = 0$. It follows

$$\frac{P(X \in u\cdot)}{P(X \in uE)} = \frac{u^{-\alpha} l(u) P(u^{-1}X \in \cdot)}{u^{-\alpha} l(u) P(u^{-1}X \in E)} \xrightarrow{v} \frac{\mu(\cdot)}{\mu(E)} =: \widetilde{\mu}(\cdot).$$

Obviously, $\widetilde{\mu}$ is again a Radon measure with $\widetilde{\mu}(\overline{\mathbb{R}}^d \backslash \mathbb{R}^d) = 0$.

□

Note that both limiting measures μ of (ii) and (iii) in the theorem above are again homogeneous with index α, where the index of regular variation coincides with the index specified in (i). Furthermore, we can see from the proofs that the sequence (a_n), the slowly varying function l and the set E can be chosen such that the limiting measures μ coincide in all three cases. Since the characterization of Theorem 2.1.12(iii) is very convenient in some proofs of this thesis, we write $X \in RV(\alpha, l, \mu)$ if we want to emphasize the usage of this definition.

A direct consequence of multivariate regular variation can be obtained by observing that for a multivariate regularly varying random variable X, the norm $\|X\|$ is univariate regularly varying and thus $P(\|X\| > x)$ is a regularly varying function.

Lemma 2.1.13. *Let $X \in \mathbb{R}^d$ be a random vector such that $X \in RV(\alpha, (a_n), \mu)$. Then for $\delta > 0$*

$$\mathbb{E}\,\|X\|^\delta \begin{cases} < \infty & \text{if } \delta < \alpha, \\ = \infty & \text{if } \delta > \alpha. \end{cases}$$

No general statement can be given in the case $\delta = \alpha$.

If (a_n) is a sequence corresponding to a random vector $X \in RV(\alpha, (a_n), \mu)$, i.e. (a_n) is regularly varying with index $1/\alpha$, then we have an interesting result for random variables with an existing δ moment, where $\delta > \alpha$.

Lemma 2.1.14. *Let (a_n) be a sequence that is regularly varying with index $1/\alpha$. If $Y \in \mathbb{R}^d$ is a random vector such that $\mathbb{E}\,\|Y\|^\delta < \infty$ for some $\delta > \alpha$, then*

$$\lim_{n\to\infty} nP(a_n^{-1}Y \in B) = 0$$

for every relatively compact $B \in \mathcal{B}(\overline{\mathbb{R}}^d \backslash \{0\})$.

Proof. The result follows directly from the proof of Lemma 1.32 in Lindskog (2004). □

Next we consider sums of independent random vectors which are regularly varying. In the following theorem, the definition of regular variation is weakened in the sense that degenerate limit measures are also allowed.

Theorem 2.1.15 (Lindskog (2004), Theorem 1.28). *Let $X_1, X_2 \in \mathbb{R}^d$ be two independent random vectors such that there exists a positive sequence (a_n), $a_n \nearrow \infty$, and two Radon measures μ_1 and μ_2 on $\mathcal{B}(\overline{\mathbb{R}}^d \backslash \{0\})$ with $\mu_1(\overline{\mathbb{R}}^d \backslash \mathbb{R}^d) = \mu_2(\overline{\mathbb{R}}^d \backslash \mathbb{R}^d) = 0$ such that, as $n \to \infty$,*

$$nP(a_n^{-1}X_1 \in \cdot) \xrightarrow{v} \mu_1(\cdot) \quad \text{and} \quad nP(a_n^{-1}X_2 \in \cdot) \xrightarrow{v} \mu_2(\cdot)$$

on $\mathcal{B}(\overline{\mathbb{R}}^d \backslash \{0\})$. *Then*

$$nP(a_n^{-1}(X_1 + X_2) \in \cdot) \xrightarrow{v} \mu_1(\cdot) + \mu_2(\cdot).$$

Combining Theorem 2.1.15 and Lemma 2.1.14 we obtain an immediate corollary.

Corollary 2.1.16. *Let* $X_1, X_2 \in \mathbb{R}^d$ *be two independent random vectors such that* $X_1 \in RV(\alpha, (a_n), \mu)$ *and* $\mathbb{E}\,\|X_2\|^\delta < \infty$ *for some* $\delta > \alpha$. *Then*

$$nP(a_n^{-1}(X_1 + X_2) \in \cdot) \xrightarrow{v} \mu(\cdot),$$

i.e. $(X_1 + X_2) \in RV(\alpha, (a_n), \mu)$.

Note that the previous theorem also holds in case of dependent random variables.

Since one is often interested in sums of random vectors which are regularly varying with a different index of regular variation, we can also obtain an easy corollary for that special case.

Corollary 2.1.17. *Let* $X_1, X_2 \in \mathbb{R}^d$ *be two independent random vectors such that* $X_1 \in RV(\alpha_1, (a_n), \mu_1)$ *and* $X_2 \in RV(\alpha_2, (b_n), \mu_2)$, *where* $\alpha_1 < \alpha_2$. *Then*

$$nP(a_n^{-1}(X_1 + X_2) \in \cdot) \xrightarrow{v} \mu_1(\cdot)$$

and hence $(X_1 + X_2) \in RV(\alpha_1, (a_n), \mu_1)$.

Proof. Set $\delta := (\alpha_1 + \alpha_2)/2$. Then $\delta \in (\alpha_1, \alpha_2)$ and thus by Lemma 2.1.13 $\mathbb{E}\,\|X_2\|^\delta < \infty$. The result follows by Corollary 2.1.16.

□

In addition to sums of random variables, products can also be considered. The following Theorem is the multivariate version of Breiman's Lemma.

Theorem 2.1.18 (Basrak et al. (2002), Proposition A.1). *Let* $X \in \mathbb{R}^d$ *be a random vector such that* $X \in RV(\alpha, (a_n), \mu)$ *and let* A *be a random* $q \times d$ *matrix independent of* X. *If* $\mathbb{E}\,\|A\|^\delta \in (0, \infty)$ *for one* $\delta > \alpha$, *then, as* $n \to \infty$,

$$nP(a_n^{-1}AX \in \cdot) \xrightarrow{v} \mathbb{E}\left(\mu \circ A^{-1}(\cdot)\right)$$

on $\mathcal{B}(\overline{\mathbb{R}}^d \backslash \{0\})$.

The results of Corollary 2.1.17 and Theorem 2.1.18 can be interpreted as follows. In sums (and products resp.) of independent random variables, the tail behavior is always dominated by the summand (and factor resp.) with the lowest index of regular variation. On the other hand, the summands (factors) with lighter tails do not contribute to the tail behavior of the sum (product).

2.1.4 Processes with Càdlàg Sample Paths

We follow Hult and Lindskog (2005) to introduce the notion of regular variation on $\mathbb{D}$. Let $\mathbb{D}$ be the space of càdlàg (right-continuous with left limits) functions $x : [0,1] \to \mathbb{R}^d$ equipped with the J_1 metric (equivalent to the d_0 metric of Billingsley (1968)) such that $\mathbb{D}$ is a complete and separable metric space. Using the supremum norm $\|x\|_\infty = \sup_{t\in[0,1]} \|x_t\|$ we can then introduce the unit sphere $S_\mathbb{D} = \{x \in \mathbb{D} : \|x\|_\infty = 1\}$, equipped with the relativized topology of $\mathbb{D}$. Next, we equip $(0,\infty]$ with the metric $\rho(x,y) = |1/x - 1/y|$ which makes it a complete separable metric space. Then also the space $\overline{\mathbb{D}}_0 = (0,\infty] \times S_\mathbb{D}$, equipped with the metric $\max\{\rho(x^*,y^*), d_0(\widetilde{x},\widetilde{y})\}$, is a complete separable metric space.

If we use the polar coordinate transformation $T : \mathbb{D}\backslash\{0\} \to \overline{\mathbb{D}}_0$, $x \mapsto (\|x\|_\infty, x/\|x\|_\infty)$, we see that the spaces $\mathbb{D}\backslash\{0\}$ and $(0,\infty) \times S_\mathbb{D}$ are homeomorphic. Thus, the Borel sets $\mathcal{B}(\overline{\mathbb{D}}_0)$ of interest can be viewed as the infinite dimensional extension of the one-point uncompactification that is used to introduce finite dimensional regular variation (cf. Bingham et al. (1987), Embrechts et al. (1997) and Resnick (1987)).

Regular Variation on $\mathbb{D}$ can then be introduced in terms of the so-called $\hat{w}$-convergence of boundedly finite measures on $\overline{\mathbb{D}}_0$. A measure μ on a complete separable metric space E is said to be *boundedly finite* if $\mu(B) < \infty$ for every bounded set $B \in \mathcal{B}(E)$. Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of boundedly finite measures on E. Then (μ_n) converges to μ in the *$\hat{w}$-topology* if $\mu_n(B) \to \mu(B)$ for all bounded Borel sets $B \in \mathcal{B}(E)$ with $\mu(\partial B) = 0$. We write $\mu_n \xrightarrow{\hat{w}} \mu$. Note that for locally compact spaces E the boundedly finite measures are called Radon measures and the notions of $\hat{w}$-convergence and vague convergence coincide. See Daley and Vere-Jones (1988) and Kallenberg (1983) for details on $\hat{w}$-convergence and vague convergence.

We are now able to formulate regular variation for stochastic processes with sample paths in $\mathbb{D}$.

Definition 2.1.19 (Functional Regular Variation). A stochastic process (X_t) with sample paths in $\mathbb{D}$ is said to be *regularly varying* if there exists a positive sequence (a_n), $n \in \mathbb{N}$, with $a_n \nearrow \infty$ and a nonzero boundedly finite measure μ on $\mathcal{B}(\overline{\mathbb{D}}_0)$ with $\mu(\overline{\mathbb{D}}_0\backslash\mathbb{D}) = 0$ such that, as $n \to \infty$,

$$nP(a_n^{-1}X \in \cdot) \xrightarrow{\hat{w}} \mu(\cdot) \quad \text{ on } \mathcal{B}(\overline{\mathbb{D}}_0).$$

As in the finite dimensional case, direct calculation shows that the measure μ is homogeneous, i.e. there exists a positive index $\alpha > 0$ such that $\mu(uB) = u^{-\alpha}\mu(B)$ for all $u > 0$ and for every $B \in \mathcal{B}(\overline{\mathbb{D}}_0)$. Thus, it makes sense to say that the process (X_t) is regularly varying with (tail) index α. For short, we write $X \in RV_{\overline{\mathbb{D}}_0}(\alpha, (a_n), \mu)$.

Several alternative definitions of regular variation exist. For example, a process (X_t) in $\mathbb{D}$ is regularly varying if and only if there exists an index $\alpha > 0$ and a probability measure σ on $\mathcal{B}(S_{\mathbb{D}})$ such that for every positive $x > 0$, as $u \to \infty$,

$$\frac{P(\|X\|_\infty > ux, X/\|X\|_\infty \in \cdot)}{P(\|X\|_\infty > u)} \xrightarrow{\hat{w}} x^{-\alpha}\sigma(\cdot) \quad \text{on } \mathcal{B}(S_{\mathbb{D}}). \tag{2.1.3}$$

In this definition, the probability measure σ is called the *spectral measure* of X. See Theorem 4 in Hult and Lindskog (2005) for the equivalence between (2.1.3) and Definition 2.1.19.

Now we recall several useful results from Hult and Lindskog (2005) related to regular variation on $\mathbb{D}$. Since it is often of interest, how the regular variation property is preserved under mappings, we look at two different continuous mapping theorems. Therefore, for any function h from a metric space E to a metric space E' we introduce the set $disc(h)$ which consists of all discontinuities of h.

Theorem 2.1.20 (Hult and Lindskog (2005), Theorem 6). *Let (X_t) be a stochastic process with sample paths in $\mathbb{D}$ and let E' be a complete separable metric space. Assume that $X \in RV_{\overline{\mathbb{D}}_0}(\alpha, (a_n), \mu)$ and $h : \mathbb{D} \to E'$ is a measurable mapping such that $\mu(disc(h)) = 0$ and $h^{-1}(B)$ is bounded in $\overline{\mathbb{D}}_0$ for every bounded $B \in \mathcal{B}(E')$. Then, as $n \to \infty$,*

$$nP(h(a_n^{-1}X) \in \cdot) \xrightarrow{\hat{w}} \mu \circ h^{-1}(\cdot) \quad \text{on } \mathcal{B}(E').$$

There is a different version of the previous theorem for the special case of positively homogeneous mappings of order $\gamma > 0$, i.e. mappings $h : \mathbb{D} \to \mathbb{D}$ with $h(\lambda x) = \lambda^\gamma h(x)$ for all $\lambda \geq 0$ and $x \in \mathbb{D}$.

Theorem 2.1.21 (Hult and Lindskog (2005), Theorem 8). *Let (X_t) be a stochastic process with sample paths in $\mathbb{D}$ and let $X \in RV_{\overline{\mathbb{D}}_0}(\alpha, (a_n), \mu)$. Furthermore, suppose that $h : \mathbb{D} \to \mathbb{D}$ is a measurable mapping which is positively homogeneous of order $\gamma > 0$ such that $\mu(disc(h) \cap \overline{\mathbb{D}}_0) = 0$ and $h^{-1}(B)$ is bounded in $\overline{\mathbb{D}}_0$ for every bounded $B \in \mathcal{B}(\overline{\mathbb{D}}_0) \cap \mathbb{D}$. Then, as $n \to \infty$,*

$$nP(a_n^{-\gamma}h(X) \in \cdot) \xrightarrow{\hat{w}} \mu \circ h^{-1}(\cdot \cap \mathbb{D}) \quad \text{on } \mathcal{B}(\overline{\mathbb{D}}_0).$$

The next theorem states some necessary and sufficient conditions for regular variation on $\mathbb{D}$. In the theorem, we use the notation

$$w(x, T_0) := \sup_{t_1, t_2 \in T_0} \|x_{t_1} - x_{t_2}\| \quad \text{and}$$

$$w''(x, \delta) := \sup_{t_1 \leq t \leq t_2;\, t_2 - t_1 \leq \delta} \min\{\|x_t - x_{t_1}\|, \|x_{t_2} - x_t\|\}$$

for $x \in \mathbb{D}$, $T_0 \subseteq [0, 1]$ and $\delta \in [0, 1]$.

Theorem 2.1.22 (Hult and Lindskog (2005), Theorem 10). *Let (X_t) be a stochastic process with sample paths in $\mathbb{D}$. Then the following statements are equivalent.*

(i) $X \in RV_{\overline{\mathbb{D}}_0}(\alpha, (a_n), \mu)$.

(ii) There exists a set $T \subseteq [0,1]$ containing 0, 1 and all but at most countably many points of $[0,1]$, a positive sequence $a_n \nearrow \infty$ and a collection $\{\mu_{t_1,\dots,t_k} : t_i \in T, k \in \mathbb{N}\}$ of Radon measures on $\mathcal{B}(\overline{\mathbb{R}}^{dk}\backslash\{0\})$ with $\mu_{t_1,\dots,t_k}(\overline{\mathbb{R}}^{dk}\backslash\mathbb{R}^{dk}) = 0$ and μ_t is nonzero for some $t \in T$ such that

$$nP(a_n^{-1}(X_{t_1},\dots,X_{t_k}) \in \cdot) \xrightarrow{v} \mu_{t_1,\dots,t_k}(\cdot) \quad \text{on } \mathcal{B}(\overline{\mathbb{R}}^{dk}\backslash\{0\}) \tag{2.1.4}$$

holds for all $t_1,\dots,t_k \in T$. Furthermore, for every $\varepsilon, \eta > 0$, there exist $\delta \in (0,1)$ and $n_0 \in \mathbb{N}$ such that, for $n \geq n_0$,

$$nP(a_n^{-1} w(X, [0,\delta)) \geq \varepsilon) \leq \eta, \tag{2.1.5}$$

$$nP(a_n^{-1} w(X, [1-\delta,1)) \geq \varepsilon) \leq \eta, \tag{2.1.6}$$

$$nP(a_n^{-1} w''(X, \delta) \geq \varepsilon) \leq \eta. \tag{2.1.7}$$

Remark 2.1.23. The theorem links regular variation of the process $(X_t)_{t\in[0,1]}$ with sample paths in $\mathbb{D}$ to regular variation of the finite dimensional distributions $(X_{t_1},\dots,X_{t_k})$ of the the process. Key to that connection are the relative compactness criteria (2.1.5), (2.1.6) and (2.1.7) which restrict the oscillation of the process (X_t) in small areas. See Hult and Lindskog (2005), Example 11, for a process satisfying conditions (2.1.5) and (2.1.6), but not (2.1.7).

Some more accessible sufficient conditions for regular variation in $\mathbb{D}$ can be given in case of a strong Markov process with sample paths in $\mathbb{D}$. In that case we denote by $P_{s,t}(x,B)$ the transition functions and define

$$\alpha_r(u) := \sup\left\{P_{s,t}(x, B_r(x)^c) : x \in \mathbb{R}^d,\ s,t \in [0,1] \text{ and } t-s \in [0,u]\right\}.$$

Theorem 2.1.24 (Hult and Lindskog (2005), Theorem 13). *Let (X_t), $t \in [0,1]$, be a strong Markov process with sample paths in $\mathbb{D}$ that satisfies $\lim_{r\to\infty} \alpha_r(1) = 0$. Assume that X_t is regularly varying for every fixed $t \in [0,1]$, i.e. there exists a set $T \subseteq [0,1]$ with $0,1 \in T$ containing all but at most countably many points of $[0,1]$, a positive sequence $a_n \nearrow \infty$ and a collection $\{\mu_t : t \in T\}$ of Radon measures on $\mathcal{B}(\overline{\mathbb{R}}^d\backslash\{0\})$ with $\mu_t(\overline{\mathbb{R}}^d\backslash\mathbb{R}^d) = 0$ and μ_1 is nonzero such that, as $n \to \infty$,*

$$nP(a_n^{-1}(X_t) \in \cdot) \xrightarrow{v} \mu_t(\cdot) \quad \text{on } \mathcal{B}(\overline{\mathbb{R}}^d\backslash\{0\})$$

holds for all $t \in T$*. If, additionally, for every* $\varepsilon > 0$ *and* $\eta > 0$ *there exists* $\delta > 0$ *such that* $\delta \in T$, $1 - \delta \in T$ *as well as*

$$\mu_\delta(B_\varepsilon(0)^c) - \mu_0(B_\varepsilon(0)^c) \leq \eta \quad \text{and} \quad \mu_1(B_\varepsilon(0)^c) - \mu_{1-\delta}(B_\varepsilon(0)^c) \leq \eta,$$

then $(X_t) \in RV_{\overline{\mathbb{D}}_0}(\alpha, (a_n), \mu)$*, where* μ *is uniquely determined by* $\{\mu_t : t \in T\}$*.*

A similar theory can also be established for stochastic processes with continuous sample paths. See e.g. de Haan and Lin (2001) for details and limit results.

2.2 Weak Convergence of Point Processes

Point processes (or random point measures resp.) play an important role in probability theory and statistics. They are especially useful in extreme value theory, as they are closely related to regular variation. Hence, in a regularly varying setting, point processes are a very helpful tool in proofs. Furthermore, point processes of maxima, exceedances etc. play an important role in the analysis of the structure of the extremes. Interesting aspects of that area can be, for example, extremal clustering, long memory effects etc.

In Section 2.2.1 we introduce the general theory of point processes. We define point processes and introduce the notion of weak convergence of point processes. Furthermore, we also have a closer look at the special case of the Poisson process. Then, in Section 2.2.2, we state the classic result that links regular variation to point process convergence. Moreover, we introduce several other (point) processes of interest, e.g. the point process of exceedances, the point process of record times and extremal processes. Finally, we state a result for a functional setting.

2.2.1 General Theory

In this section we shortly introduce point processes and the relevant theory of weak convergence. For more detailed and very good introductions to the topic, see also Daley and Vere-Jones (1988) and Daley and Vere-Jones (2008). Since point processes are a special case of random measures, see also Kallenberg (1983) for more general reading on random measures. Introductions to point processes with extreme value theory in view can also be found in Resnick (2007) and Embrechts et al. (1997). The statistical aspect of point processes is examined in Illian et al. (2008).

Let E be a locally compact space and $\mathcal{E}$ the corresponding σ-algebra. A measure $m : \mathcal{E} \to \mathbb{N}_0 \cup \{\infty\}$ is called *counting measure*. A counting measure m is a *point measure* if

m is Radon, i.e. $m(K) < \infty$ for every compact set $K \in \mathcal{E}$. A point measure m is called *simple* if $m(\{x\}) \in \{0, 1\}$ for all $x \in E$. Defining the *Dirac measure* ε_x by

$$\varepsilon_x(A) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A, \end{cases}$$

for any $x \in E$ and $A \in \mathcal{E}$, every point measure m can be written in the form

$$m(A) = \sum_{i \in \mathbb{N}} \varepsilon_{x_i}(A),$$

where $x_i \in E$, $i \in \mathbb{N}$, are called the *atoms* of m. Illustrative, $m(A)$ denotes the number of points in the set A. Denote by $M_p(E)$ the set of all point measures on E and by $\mathcal{M}_p(E)$ the corresponding σ-algebra. Now a point process can be defined as a random point measure.

Definition 2.2.1 (Point Process). A measurable mapping

$$N : (\Omega, \mathcal{F}, P) \to [M_p(E), \mathcal{M}_p(E)]$$

is called *point process* (or *random point measure*).

A point process N is called *simple* if N is a simple point measure with probability 1. Likewise to point measures, point processes can be written in the form

$$N(A) = \sum_{i \in \mathbb{N}} \varepsilon_{X_i}(A),$$

where $A \in \mathcal{E}$ and $(X_i)_{i \in \mathbb{N}}$ are random elements of E. Very useful for the characterization of point processes are Laplace functionals.

Definition 2.2.2 (Laplace Functional). Let N be a point process. Then the *Laplace functional* of N is given by

$$\begin{aligned} \Psi_N(f) &= \mathbb{E}\left(\exp\left(-\int_E f \, dN\right)\right) \\ &= \int_{M_p(E)} \exp\left(-\int_E f(x) \, m(dx)\right) P_N(dm), \end{aligned}$$

where $f : E \to [0, \infty)$ is a measurable function.

Note that the definition of the Laplace functional can be easily extended to a more general space of random measures. The importance of the Laplace functional can be derived from the following fact.

Lemma 2.2.3. *The Laplace functional Ψ_N determines the distribution of N uniquely.*

The limit theory for point processes is described by the concept of weak convergence. Weak convergence is in general defined as convergence of the infinite-dimensional distributions, but by Daley and Vere-Jones (1988), Theorem 9.1.IV, it is sufficient to consider only the finite-dimensional distributions.

Definition 2.2.4 (Weak Convergence). A sequence of point processes $(N_n)_{n\in\mathbb{N}}$ on the space $[E, \mathcal{E}]$ *converges weakly* to a point process N (on the same space) in $M_p(E)$ if

$$P\left(N_n(A_1) = k_1, \ldots, N_n(A_m) = k_m\right) \to P\left(N(A_1) = k_1, \ldots, N(A_m) = k_m\right)$$

for all $m, k_1, \ldots, k_m \in \mathbb{N}$ and $A_1, \ldots, A_m \in \mathcal{E}$ such that $P(N(\partial A_i) = 0) = 1$ for $i = 1, \ldots, m$. In that case we write $N_n \xrightarrow{d} N$.

There exists an important characterization of this convergence concept which is very useful in many proofs: weak convergence of point processes is equivalent to convergence of the corresponding Laplace functionals. Therefore denote by $C_K^+(E)$ the space of continuous functions $f : E \to [0, \infty)$ with compact support.

Theorem 2.2.5 (Daley and Vere-Jones (1988), Theorem 9.1.VII(ii)). *Let N and $(N_n)_{n\in\mathbb{N}}$ be point processes on $[E, \mathcal{E}]$. Then $N_n \xrightarrow{d} N$ if and only if*

$$\Psi_{N_n}(f) \to \Psi_N(f)$$

for all $f \in C_K^+(E)$.

Another important tool for the work with point processes are continuous mapping arguments. They ensure that weak convergence is preserved under continuous mappings. Therefore, consider a second locally compact space E_2 with corresponding σ-algebra $\mathcal{E}_2$ and let $T : E \to E_2$ be a continuous function. For a point measure $m = \sum_{i\in\mathbb{N}} \varepsilon_{x_i}$ define the transformed measure $T(m)$ by

$$T(m) = m \circ T^{-1} = \sum_{i\in\mathbb{N}} \varepsilon_{T(x_i)}.$$

Theorem 2.2.6 (Resnick (2007), Proposition 5.5). *Let N and $(N_n)_{n\in\mathbb{N}}$ be point processes on $[E, \mathcal{E}]$. Suppose that $T : E \to E_2$ is a continuous function such that the preimages of compact sets are again compact. If $N_n \xrightarrow{d} N$ in $M_p(E)$, then*

$$T(N_n) \xrightarrow{d} T(N)$$

in $M_p(E_2)$.

Finally, we introduce as an example one of the most important point processes: the Poisson random measure. It has two properties: the number of points in a fixed set A follows a Poisson distribution and the numbers of points in disjoint sets are independent.

Definition 2.2.7 (Poisson Random Measure). Let μ be a Radon measure on $\mathcal{E}$. A point process N is called *Poisson random measure* or *Poisson process* with mean measure μ (write PRM(μ)) if the following two conditions are satisfied:

(i) For any $A \in \mathcal{E}$ and $k \in \mathbb{N}$

$$P(N(A) = k) = \begin{cases} e^{-\mu(A)} \ \frac{\mu(A)^k}{k!} & \text{if } \mu(A) < \infty, \\ 0 & \text{if } \mu(A) = \infty. \end{cases}$$

(ii) For any $m \in \mathbb{N}$ and for any collection of disjoint sets $A_1, \ldots, A_m \in \mathcal{E}$ the random variables $N(A_1), \ldots, N(A_m)$ are independent.

The Laplace functional which determines the distribution of the Poisson random measure completely can be given explicitly.

Theorem 2.2.8 (Resnick (2007), Theorem 5.1). *A point process N is a PRM(μ) if and only if the Laplace functional is of the form*

$$\Psi_N(f) = \exp\left(-\int_E \left(1 - e^{-f(x)}\right) \ \mu(dx)\right)$$

for measurable functions $f : E \to [0, \infty)$.

A result similar to Theorem 2.2.6 does also hold for Poisson random measures: continuous transformations of Poisson random measures are again Poisson random measures.

Theorem 2.2.9 (Resnick (2007), Proposition 5.2). *Let $T : E \to E_2$ be a function such that the preimages of compact sets are again compact. If N is a PRM(μ) on E, then $T(N)$ is a PRM($T(\mu)$) on E_2.*

Poisson random measures are very important for extreme value theory. They appear as a weak limit of point processes with points that are independent and regularly varying. The complete theory is introduced in the following section.

2.2.2 Point Processes and Extreme Value Theory

The relevance of point processes for extreme value theory is part of basically any monograph on the topic. In this section we introduce the basic idea and show the main results. For details we refer to Resnick (1987), Resnick (2007), Embrechts et al. (1997), de Haan and Ferreira (2006) and Leadbetter et al. (1983).

We start with the main theorem that links regular variation to point process convergence. As we have defined multivariate regular variation on the space $\mathbb{R}^d$ in Section 2.1.3, we state the result for the case $E = \mathbb{R}^d$ and $\mathcal{E} = \mathcal{B}(\mathbb{R}^d)$.

Theorem 2.2.10 (Resnick (2007), Theorem 5.3). *Let $(X_i)_{i\in\mathbb{N}}$ be a collection of iid random variables with values in $\mathbb{R}^d$. Then $X_1 \in RV(\alpha, (a_n), \mu)$ if and only if*

$$\sum_{i=1}^{n} \varepsilon_{a_n^{-1} X_i} \xrightarrow{d} N,$$

where N is a PRM(μ).

This basic result can now be used in combination with the continuous mapping theorem to obtain numerous other results. As an example we present a result for point processes of vectors $(X_k, X_{k-1}, \ldots, X_{k-d+1})$, where the single entries are given by an iid sequence of real-valued random variables. Results of that kind are used for example in the analysis of the extremal behavior of discrete time moving averages and linear processes. See also Section 3.2.2 and Section 3.4.3.

Theorem 2.2.11 (Davis and Resnick (1985), Theorem 2.2). *Let $(X_i)_{i\in\mathbb{Z}}$ be a sequence of iid real-valued random variables which are regularly varying with index α, i.e. the tail balance condition of Definition 2.1.8 is satisfied. Let (a_n) be the positive sequence satisfying $nP(a_n^{-1}|X_1| > x) \to x^{-\alpha}$ for all $x > 0$. For fixed $d \in \mathbb{N}$ define the point processes*

$$N_n := \sum_{k=1}^{\infty} \varepsilon_{a_n^{-1}(X_k, X_{k-1}, \ldots, X_{k-d+1})}$$

in $M_p(\overline{\mathbb{R}}^d \backslash \{0\})$. Then

$$N_n \xrightarrow{d} N := \sum_{k=1}^{\infty} \sum_{i=1}^{d} \varepsilon_{J_k e_i},$$

where (J_k) are the points of a PRM(μ) on $\overline{\mathbb{R}}^d \backslash \{0\}$ with mean measure

$$\mu(dx) = \alpha |x|^{\alpha-1} \left(p \, \mathbb{1}_{(0,\infty]}(x) + q \, \mathbb{1}_{[-\infty,0)}(x) \right) dx$$

and e_i is the ith unit vector in $\mathbb{R}^d$.

Another interesting characteristic is the point process of exceedances, see also Embrechts et al. (1997), Section 5.3. It is given by

$$N_n(\cdot) = \sum_{i=1}^{n} \varepsilon_{i/n}(\cdot)\, \mathbb{1}_{\{X_i > u_n\}}, \tag{2.2.1}$$

where $n \in \mathbb{N}$, $(X_i)_{i\in\mathbb{N}}$ is a sequence of real-valued random variables and (u_n) is a sequence of real thresholds. From the following observation, we see the importance of such processes for extreme value theory. Let $X_{k:n}$ be the kth biggest value (order statistic) of the sample $X_1, \ldots, X_n$ and denote by $M_n = X_{1:n}$ the sample maximum. Then direct calculation shows (cf. Embrechts et al. (1997), Remark 5.1.3)

$$\{N_n((0,1]) < k\} = \{X_{k:n} \leq u_n\}$$

and in particular

$$\{N_n((0,1]) = 0\} = \{M_n \leq u_n\}.$$

The results for weak convergence of point processes of exceedances to Poisson random measures are also well known.

Theorem 2.2.12 (Embrechts et al. (1997), Theorem 5.3.2). *Suppose $(X_i)_{i\in\mathbb{N}}$ is a sequence of iid real-valued random variables with common distribution function F and (u_n) is s sequence of real thresholds. Assume that the condition*

$$n\overline{F}(u_n) = \mathbb{E}\left(\sum_{i=1}^{n} \mathbb{1}_{\{X_i > u_n\}}\right) \to \tau$$

holds for some $\tau \in (0, \infty)$. Then the point process of exceedances given by (2.2.1) *satisfies*

$$N_n \xrightarrow{d} N,$$

where N is a PRM$(\tau\lambda(\cdot))$ on $(0,1]$.

The conditions of the theorem can be relaxed in several ways. First, in the definition of N_n, the number of summands n can be replaced by a random Number $N'(n)$ and the times i by T_i. Then the point process of exceedances is given by

$$N_n(\cdot) = \sum_{i=1}^{N'(n)} \varepsilon_{T_i/n}(\cdot)\, \mathbb{1}_{\{X_i > u_n\}},$$

where $T_n = Y_1 + \cdots + Y_n$ are the points of a renewal counting process, $(Y_i)_{i\in\mathbb{N}}$ is an iid sequence of positive random variables independent of (X_i) and $N'(n) = \#\{i \in \mathbb{N} : T_i \leq n\}$.

Then by Embrechts et al. (1997), Theorem 5.3.4, the result of Theorem 2.2.12 still holds. Furthermore, the result of Theorem 2.2.12 still holds if the iid sequence (X_i) is replaced by a strictly stationary sequence $(\widetilde{X}_t)$ (cf. Embrechts et al. (1997), Theorem 5.3.6). Results for limit probabilities and distributions of upper order statistics as well as weak convergence of the number of exceedances can be obtained likewise (cf. Embrechts et al. (1997), Theorem 5.3.8, Theorem 5.3.9 and Corollary 5.3.10).

Next, we analyze the convergence of sample maxima to extremal processes. Recall that $M_n = \max(X_1, \ldots, X_n)$ denotes the sample maximum of an iid real-valued sequence (X_i). If N given by

$$N = \sum_{i \in \mathbb{N}} \varepsilon_{(t_i, j_i)}$$

is PRM$(\lambda \times \mu)$ for the univariate Lebesgue measure λ and a Radon measure μ, define the corresponding *extremal process* by

$$Y_\mu(t) = \max_{k \in \mathbb{N}:\, t_k \leq t} j_k. \tag{2.2.2}$$

Weak convergence of Y_μ is equivalent to regular variation.

Theorem 2.2.13 (Resnick (2007), Proposition 7.2). *Let $(X_i)_{i \in \mathbb{N}}$ be an iid sequence of real-valued nonnegative random variables. Then $X_1 \in RV(\alpha, (a_n), \mu)$ if and only if*

$$Y_n(\cdot) := a_n^{-1} M_{\lfloor n \cdot \rfloor} \xrightarrow{d} Y_\mu(\cdot)$$

on the space of càdlàg functions $y : [0, \infty) \to [0, \infty)$.

Note that the theorem also holds for random variables in $\mathbb{R}^d$ if maxima are taken componentwise.

Extremal processes can also be defined for general (not regularly varying) random variables. For an iid real-valued sequence (X_i) of random variables with common distribution function F the finite-dimensional distributions of the maxima M_n are given by

$$P\left(M_{t_1} \leq x_1, \ldots, M_{t_m} \leq x_m\right) = F^{t_1}\left(\max_{i=1}^{m} x_i\right) F^{t_2 - t_1}\left(\max_{i=2}^{m} x_i\right) \cdots F^{t_m - t_{m-1}}\left(x_m\right), \tag{2.2.3}$$

where $m, t_1, \ldots, t_m \in \mathbb{R}$ such that $t_1 < t_2 < \ldots < t_m$ and $x_1, \ldots, x_m \in \mathbb{R}$ (cf. Embrechts et al. (1997), Section 5.4.2).

Definition 2.2.14 (F-Extremal Process). The process $Y = (Y(t))_{t \in \mathbb{R}^+}$ which is uniquely determined by the finite-dimensional distributions (2.2.3) is called *F-extremal process*.

F-extremal processes are the natural extension of the discrete-time sequences of sample maxima (M_n) to a continuous-time setting. By Resnick (1987), Proposition 4.7, F-extremal processes are stochastically continuous, càdlàg, nondecreasing (almost surely) and they are Markov jump processes. Furthermore, it follows from Embrechts et al. (1997), Proposition 5.4.4, that F-extremal processes have a point process representation, i.e. any F-extremal process $(Y(t))$ can be given in the form

$$Y(t) = Y_\mu(t),$$

where Y_μ is the extremal process specified by (2.2.2) and μ is given by $\mu((a,b]) = \ln(F(b)) - \ln(F(a))$ for $a < b$. The jump times of an F-extremal process constitute another Poisson random measure.

Theorem 2.2.15 (Embrechts et al. (1997), Theorem 5.4.7). *Let $Y = (Y(t))$ be an F-extremal process and let (τ_i), $i \in \mathbb{N}$, be the jump times of Y. If F is continuous, then the point process N_∞ of jump times given by*

$$N_\infty = \sum_{i \in \mathbb{N}} \varepsilon_{\tau_i}$$

is a PRM(μ) on $\mathbb{R}^+$ with

$$\mu((a,b]) = \ln(b) - \ln(a).$$

F-extremal processes are closely related to records and record times. Intuitively, for a sequence (X_i) of random variables, X_n is a *record* if $X_n > M_{n-1}$. Denote by (L_i) the *record times*, i.e. the times where the records occur given by the recursion

$$L_1 = 1 \quad \text{and} \quad L_n = \min\{k > L_{n-1} : X_k > X_{L_{n-1}}\} \quad \text{for } n > 1.$$

The sequence of records is then given by $(X_{L_i})_{i \in \mathbb{N}}$ and these are the points of a Poisson random measure.

Theorem 2.2.16 (Embrechts et al. (1997), Theorem 5.4.1). *Let F be a distribution function with left and right endpoint given by*

$$x_F^l = \inf\{x \in \mathbb{R} : \ F(x) > 0\} \quad \text{and} \quad x_F^r = \sup\{x \in \mathbb{R} : \ F(x) < 1\}$$

and let $(X_i)_{i \in \mathbb{N}}$ be an iid sequence with common distribution function F. Then the records (X_{L_i}) are the points of a PRM(μ) with mean measure μ given by

$$\mu((a,b]) = \ln\left(\overline{F}(a)\right) - \ln\left(\overline{F}(b)\right)$$

for $a, b \in \mathbb{R}$ such that $x_F^l < a \leq b < x_F^r$.

From the definition of F-extremal processes one would intuitively hope that the record times (L_i) and the jump times (τ_i) of the corresponding F-extremal process $(Y(t))$ behave similarly, as $(M_n) \stackrel{d}{=} (Y(n))$ for $n \in \mathbb{N}$. The relation holds in the limit as shown in the following theorem.

Theorem 2.2.17 (Embrechts et al. (1997), Theorem 5.4.9). *Let (L_i) be the record times of an iid sequence (X_i) and let (τ_i) be the jump times of the corresponding F-extremal process $(Y(t))$. Then*

$$N_n(\cdot) := \sum_{i \in \mathbb{N}} \varepsilon_{L_i/n}(\cdot) \xrightarrow{d} N_\infty(\cdot) = \sum_{i \in \mathbb{N}} \varepsilon_{\tau_i}(\cdot)$$

on $M_p(\mathbb{R}^+)$.

There exist many other results related to records and extremal processes. For example, limits results for the frequency and growth of records can be derived accordingly. For further details and many interesting applications we refer to Embrechts et al. (1997), Chapter 5, and Resnick (1987), Chapter 4.

To conclude this section, we present a result for the point process convergence of stochastic processes with càdlàg sample paths. Therefore, recall the definitions of $\mathbb{D}$ and $\overline{\mathbb{D}}_0$ from Section 2.1.4 and let $M_p(\overline{\mathbb{D}}_0)$ denote the space of all point measures on $\overline{\mathbb{D}}_0$ equipped with the $\hat{w}$-topology.

The following theorem is the extension of the classical result of Theorem 2.2.10 to the space $\mathbb{D}$, i.e. to a state space which is not locally compact. Similar results have also been proved by de Haan and Lin (2001), Theorem 2.4, in the case of real-valued processes which are regularly varying with index 1 and by Davis and Mikosch (2008) for $\mathbb{D}$-valued random fields.

Theorem 2.2.18. *Let $(X_i)_{i \in \mathbb{N}}$ be an iid sequence of stochastic processes with values in $\mathbb{D}$. Then $X_1 \in RV_{\overline{\mathbb{D}}_0}(\alpha, (a_n), \mu)$ if and only if*

$$N_n = \sum_{i=1}^{n} \varepsilon_{a_n^{-1} X_i} \xrightarrow{d} N$$

in $M_p(\overline{\mathbb{D}}_0)$, where N is a PRM(μ).

Proof. The proof can be obtained by changing from vague-topology to the $\hat{w}$-topology in the proof of Proposition 3.21 in Resnick (1987). This change of topology does not affect the proof which is based on the Laplace functionals of the point processes involved (cf. Davis and Mikosch (2008), Proof of Lemma 2.2).

□

2.3 Lévy Processes and Infinite Divisibility

Lévy processes are a very general class of stochastic processes which have stationary and independent increments. They cover a wide range of commonly used examples such as the Brownian Motion (or Gaussian processes), Poisson and compound Poisson processes. Lévy processes are the natural extension of random walks with independent increments to a continuous time setting and they are closely linked to infinite divisibility.

In Section 2.3.1 we introduce the general theory of Lévy processes and infinite divisibility. We give definitions and the results that link the two. Furthermore, we also analyze the corresponding characteristic functions and introduce the Lévy-Itô decomposition. In Section 2.3.2 we analyze the tail behavior of Lévy processes in terms of regular variation which is closely linked to the driving Lévy measure. A functional result is also given.

2.3.1 General Theory

We follow the relevant literature to give a short introduction to Lévy processes and infinitely divisible distributions. For details and very good monographs on the topic we refer to Sato (2002), Applebaum (2004) and Kyprianou (2006).

Definition 2.3.1 (Lévy Process). A stochastic process (L_t), $t \in \mathbb{R}^+$, with values in $\mathbb{R}^d$ is called a *Lévy process* if it satisfies the following conditions:

(1) (L_t) has independent increments, i.e. for every $n \in \mathbb{N}$ and for every increasing sequence $0 \leq t_0 \leq t_1 \leq \ldots \leq t_n$ the increments L_{t_0}, $L_{t_1} - L_{t_0}$, ..., $L_{t_n} - L_{t_{n-1}}$ are independent.

(2) the increments of (L_t) are stationary, i.e. the distribution of $L_{t+s} - L_t$ does not depend on the choice of t.

(3) $L_0 = 0$ almost surely.

(4) (L_t) is stochastically continuous, i.e. for every $\varepsilon > 0$ and $t \geq 0$

$$\lim_{t \to s} P(\|L_t - L_s\| > \varepsilon) = 0.$$

(5) (L_t) has càdlàg sample paths almost surely.

Note that two-sided Lévy processes (L_t), $t \in \mathbb{R}$, can be defined similarly. There are several other processes closely related to the conditions above. An *additive process* is a process where all of the conditions of a Lévy process except (2) are satisfied. If condition (5) is not required, we speak of a Lévy process (additive process) *in law*.

Theorem 2.3.2 (Sato (2002), Theorem 11.5). *Let (L_t) be a Lévy (resp. additive) process in law. Then there exists a modification of (L_t) which is a Lévy (resp. additive) process.*

The theorem shows, why condition (5) in the definition of Lévy processes does not play an important role in the theory of such processes. This is also the reason, why one often finds definitions of Lévy processes without this condition in the literature.

Closely related to Lévy processes are infinitely divisible distributions.

Definition 2.3.3 (Infinite Divisibility). A random vector $X \in \mathbb{R}^d$ is said to be *infinitely divisible* if for each $n \in \mathbb{N}$ there exist iid random vectors $Y_1^{(n)}, \dots, Y_n^{(n)}$ such that

$$X \stackrel{d}{=} Y_1^{(n)} + \cdots + Y_n^{(n)}.$$

Infinite divisibility can also be defined in terms of probability measures. For any probability measure μ denote by

$$\mu^n := \underbrace{\mu * \cdots * \mu}_{n\text{-times}}$$

the n-fold convolution of μ. Then a probability measure μ is said to be infinitely divisible if for any $n \in \mathbb{N}$ there is a probability measure μ_n such that $\mu = \mu_n^n$. The measure μ_n is the so-called *n-th root of μ*. By Applebaum (2004), Proposition 1.2.6, both definitions of infinite divisibility coincide if μ is the law of X.

The concrete form of the characteristic function of an infinitely divisible random variable (distribution) is also known. It is called the *Lévy-Khintchine representation* or *Lévy-Khintchine formula* an can be found for example in Sato (2002), Theorem 8.1.

Theorem 2.3.4 (Lévy-Khintchine Formula). *Let the random variable $X \in \mathbb{R}^d$ be infinitely divisible. Then the characteristic function of X is given by*

$$\mathbb{E}\left(e^{iu^T X}\right) = e^{\varphi(u)}$$

for all $u \in \mathbb{R}^d$ and the characteristic exponent φ has the form

$$\varphi(u) = iu^T\gamma - \frac{1}{2}u^T\Sigma u + \int_{\mathbb{R}^d}\left(e^{iu^T x} - 1 - iu^T x \mathbb{1}_{[-1,1]}(\|x\|)\right)\nu(dx), \tag{2.3.1}$$

where $\gamma \in \mathbb{R}^d$, Σ is a nonnegative-definite $d \times d$ matrix and the measure $\nu : \mathbb{R}^d \to \mathbb{R}$ satisfies

$$\nu(\{0\}) = 0 \quad \text{and} \quad \int_{\mathbb{R}^d}(\|x\|^2 \wedge 1)\,\nu(dx) < \infty.$$

Conversely, for any γ, Σ and ν as above there exists an infinitely divisible distribution with characteristic function $e^{\varphi(u)}$.

The matrix Σ in the above theorem is called the *Gaussian covariance matrix* and ν is the *Lévy measure*. As the choice of γ, Σ and ν is unique for every infinitely divisible random variable X (cf. Sato (2002), Theorem 8.1(ii)), we call (γ, Σ, ν) the *characteristic triplet* (or *generating triplet*) of X.

We have already mentioned that there is a close link between infinite divisibility and Lévy processes.

Theorem 2.3.5. *The following statements hold.*

(i) If (L_t) is a Lévy process, then L_t is infinitely divisible for each $t \geq 0$ and

$$\mathbb{E}\left(e^{iu^T L_t}\right) = e^{t\varphi(u)},$$

where φ is the characteristic exponent of L_1 given by (2.3.1).

(ii) For every infinitely divisible distribution μ on $\mathbb{R}^d$ there is a Lévy process (L_t) such that $P_{L_1} = \mu$. Furthermore, the choice of the Lévy process is unique apart from modifications.

Proof. See Applebaum (2004), Proposition 1.3.1 and Theorem 1.3.3, and Sato (2002), Corollary 11.6.

□

The theorem shows that Lévy processes and infinitely divisible distributions can be considered as equivalent. If (L_t) is a Lévy process and (γ, Σ, ν) is the characteristic triplet of L_1, then we observe that $(t\gamma, t\Sigma, t\nu)$ is the characteristic triplet of L_t for all $t \geq 0$. Hence, the Lévy process is uniquely determined by (γ, Σ, ν) and we say that (γ, Σ, ν) is the characteristic triplet of (L_t).

Next we recall an interesting result describing the jump structure of Lévy processes, the Lévy-Itô decomposition. The main statement is that any Lévy process can be decomposed into a deterministic drift, a continuous Brownian motion and a jump part. The jump part is given by the big jumps described by a Poisson random measure with mean measure $\nu \times \lambda$, where λ is the univariate Lebesgue measure, and the small jumps are represented by the corresponding compensated Poisson measure. Details and proofs of the Lévy-Itô decomposition can be found in Sato (2002), Chapter 4, Applebaum (2004), Section 2.4, and Kallenberg (2002), Chapter 15.

Theorem 2.3.6 (Lévy-Itô Decomposition). *Let (L_t) be a Lévy process with characteristic triplet (γ, Σ, ν). Then there exists a drift $b \in \mathbb{R}^d$, a Brownian motion (B_t) with*

covariance matrix Σ *and a Poisson random measure* N *independent of* (B_t) *and with mean measure* $\nu \times \lambda$ *such that for each* $t \geq 0$

$$L_t = b\,t + B_t + \int_0^t \int_{\|x\| \leq 1} x\,(N(dx, ds) - ds\,\nu(dx)) + \int_0^t \int_{\|x\|>1} x\,N(dx, ds). \tag{2.3.2}$$

The theorem shows that the sample paths of a Lévy process are described by a Brownian motion and the jumps of a (compensated) Poisson process. Note that by definition (or Theorem 2.3.2 resp.) Lévy processes have càdlàg sample paths. The following theorem gives necessary and sufficient conditions for the sample paths of a Lévy process to be of finite variation.

Theorem 2.3.7 (Sato (2002), Theorem 21.9). *A Lévy process* (L_t) *is of finite variation if and only if* $\Sigma = 0$ *and*

$$\int_{\|x\| \leq 1} \|x\|\,\nu(dx) < \infty.$$

As the Brownian motion has paths of infinite variation on any compact interval, the condition $\Sigma = 0$ is obvious. The additional condition $\int_{\|x\|\leq 1} \|x\|\,\nu(dx) < \infty$ ensures that the integral for the small jumps with $\|x\| \leq 1$ in (2.3.2) exists without a compensator and thus the small jumps are of bounded variation.

2.3.2 Regular Variation of Lévy Processes

In this section we analyze the tail behavior of Lévy processes. Therefore, we give conditions for multivariate and functional regular variation of Lévy processes. We start with necessary and sufficient conditions for Lévy processes to have finite moments.

Theorem 2.3.8 (Sato (2002), Corollary 25.8). *Suppose* (L_t) *is a Lévy process with characteristic triplet* (γ, Σ, ν) *and let* $\delta > 0$. *Then* $\mathbb{E}\,(\|L_t\|^\delta) < \infty$ *for all* $t > 0$ *if and only if*

$$\int_{\|x\| \geq 1} \|x\|^\delta\,\nu(dx) < \infty.$$

Combining Theorem 2.3.8 with the Lévy-Itô decomposition of Theorem 2.3.6 we obtain an immediate consequence.

Corollary 2.3.9. *If* (L_t) *is a Lévy process with bounded jumps, then* $\mathbb{E}\,(\|L_t\|^\delta) < \infty$ *for all* $\delta > 0$.

It is well known that also all moments of the Brownian motion are finite. Consequently, in view of the Lévy-Itô decomposition and Corollary 2.1.16 we see that the heavy tails of a Lévy process can only be caused by the big jumps of it. The big jumps are described by the Poisson random measure N which has an intensity driven by the Lévy measure ν. Hence, the extreme values of the Lévy process must come from the extreme values of ν.

Theorem 2.3.10 (Hult and Lindskog (2006a), Proposition 3.1). *Let $X \in \mathbb{R}^d$ be an infinitely divisible random variable with characteristic triplet (γ, Σ, ν). Then we have $X \in RV(\alpha, (a_n), \mu)$ if and only if $\nu \in RV(\alpha, (a_n), \mu)$. In particular, a Lévy process (L_t) with characteristic triplet (γ, Σ, ν) satisfies $L_t \in RV(\alpha, (a_n), t\mu)$ for all $t > 0$ if and only if $\nu \in RV(\alpha, (a_n), \mu)$.*

Combining this theorem with the results of Lemma 2.1.13 and Theorem 2.3.8, we obtain the following.

Corollary 2.3.11. *Let (L_t) be a Lévy process and suppose that (L_t) (or ν resp.) is regularly varying with index $\alpha > 0$. Then*

$$\mathbb{E}\,\|L_t\|^\delta < \infty \quad \text{and} \quad \int_{\|x\|\geq 1} \|x\|^\delta \nu(dx) < \infty$$

for all $\delta < \alpha$ and $t \geq 0$.

Functional regular variation as introduced in Section 2.1.4 follows similarly. If the underlying Lévy measure of the Lévy processes (L_t) is regularly varying, then (L_t), restricted to $t \in [0, 1]$, is also regularly varying in a functional sense.

Theorem 2.3.12. *Let $(L_t)_{t\in[0,1]}$ be a Lévy process and assume that L_1 has Lévy measure $\nu \in RV(\alpha, (a_n), \mu_1)$. Then $(L_t) \in RV_{\mathbb{D}_0}(\alpha, (a_n), \mu)$, where μ is uniquely determined by the measures $\{\mu_t = t\mu_1, t \in [0, 1]\}$.*

Proof. By definition (or Theorem 2.3.2 resp.) (L_t) has sample paths in $\mathbb{D}$. Furthermore, (L_t) is also a strong Markov process (cf. Sato (2002), Theorem 10.5 and Corollary 40.11). Moreover, Theorem 2.3.10 yields $L_t \in RV(\alpha, (a_n), t\mu)$. Then the result follows by Theorem 2.1.24. See also Hult and Lindskog (2005), Example 17.

□

Note that a similar result for the bigger class of additive processes does not hold in general (cf. Hult and Lindskog (2005), Example 12).

Chapter 3

The Limit Distribution of the Maximum Increment of a Random Walk with Dependent Regularly Varying Jump Sizes[1]

In this chapter, we investigate the maximum increment of a random walk with heavy-tailed jump size distribution. Here heavy-tailedness is understood as regular variation of the finite-dimensional distributions. The jump sizes constitute a strictly stationary sequence. Using a continuous mapping argument acting on the point processes of the normalized jump sizes, we prove that the maximum increment of the random walk converges in distribution to a Fréchet distributed random variable. Several different settings of the dependence structure between jumps sizes are considered.

3.1 Introduction

For several decades, the interplay between heavy tails and serial dependence in a strictly stationary sequence (X_t) of random variables with common distribution F has attracted a lot of attention. One of the main goals of this research is to investigate the distributional behavior of suitable functions acting on (X_t) and to compare it with the corresponding behavior of an iid sequence with the same marginal distribution F. In a dependent sequence (X_t), high/low level exceedances typically appear in clusters and significantly determine the distributional behavior of functions of (X_t).

[1]The content of this chapter is based on Mikosch and Moser (2012).

Early on, the asymptotic behavior of the maximum functional $M_n = \max_{i=1,\dots,n} X_i$, $n \geq 1$, was studied for various classes of heavy-tailed stationary sequences. Not surprisingly, linear processes (such as ARMA) with heavy-tailed innovations were the first objects of interest: they constitute a major class of stationary processes in time series analysis. The extremes of such processes were studied first by Rootzén (1978) in the case of infinite variance stable innovations and later by Davis and Resnick (1985) for general innovation processes with regularly varying tails. Another class of a heavy-tailed stationary processes constitute solutions to stochastic recurrence equations. Pioneering work by Kesten (1973) and Goldie (1991) showed that the marginal distributions of such a process have power law tails. The extreme value theory of such processes was studied in detail in de Haan et al. (1989). We mention that the class of GARCH processes can be embedded in a natural way in some stochastic recurrence equation and therefore, under general conditions, these processes have power law tails and the asymptotic behavior of their extremes can be treated by similar methods. See for example de Haan et al. (1989) , Mikosch and Stărică (2000), Basrak et al. (1999) and Basrak et al. (2002) in the ARCH(1), GARCH$(1,1)$ and general GARCH cases. Motivated by applications in time series analysis, Davis and Resnick (1986) studied the asymptotic behavior of the sample autocovariance and sample autocorrelation functions of linear processes with innovations with regularly varying tails and infinite 4th moment. The sample autocovariance and sample autocorrelation functions of GARCH and heavy-tailed bilinear processes were studied in the papers Davis and Resnick (1996), Basrak et al. (1999), Davis and Mikosch (1998), Mikosch and Stărică (2000) and Basrak et al. (2002).

Due to very distinct extremal clustering behavior, the results for GARCH, heavy-tailed linear and iid processes differ significantly. The extremal clustering behavior of these sequences is well described by the point processes of the scaled points X_t. The results for the extremes and the sample autocovariances of a heavy-tailed sequence are then a consequence of a continuous mapping argument acting on the weakly converging sequence of these point processes. Under weak dependence conditions on the heavy-tailed stationary process (X_t), the seminal paper Davis and Hsing (1995) developed a general theory for the weak convergence of these point processes.

Different distributional behavior of functions acting on heavy-tailed stationary sequences can also be observed on the magnitude of the ruin probabilities $P(\sup_{n\geq 1}(S_n - c\,n) > u)$ as $u \to \infty$, for a constant $c > 0$ and assuming that the random walk $S_n = X_1 + \cdots + X_n$, $n \geq 1$, is driftless. These probabilities were studied for linear processes with regularly varying innovations, solutions to stochastic recurrence equations and infinite variance stable processes. See for example Mikosch and Samorodnitsky (2000a), Mikosch and Konstan-

tinides (2005) and Mikosch and Samorodnitsky (2000b). In this context, the study of large deviation probabilities $P(S_n > x_n) \to 0$ for suitable choices of $x_n \to \infty$ is crucial for the understanding of the very different ruin probabilities of distinct heavy-tailed processes (X_t).

Distinct distributional behavior of functions acting on stationary sequencees (X_t) sheds light on the dependence structure of these sequences. In a sense, these results yield qualitative and quantitative indicators which measure certain aspects of the dependence structure, far beyond covariances and correlations. This aspect is particularly important in the case of heavy-tailed sequences, where covariances and correlations are less meaningful.

As a final example of a functional acting on a heavy-tailed sequence (X_t) we mention the maximum increment of a random walk (S_n). For an iid sequence with regularly varying tails, Mikosch and Račkauskas (2010) studied the asymptotic behavior of the quantities

$$\max_{1\le l\le n} (f(l))^{-1} \max_{0\le k\le n-l} |S_{k+l} - S_k| \,, \quad n \ge 1 \,, \tag{3.1.1}$$

for suitable choices of non-decreasing sequences $(f(l))$. They gave conditions under which the distributions of (3.1.1) converge weakly to a Fréchet distribution, i.e. one of the extreme value distributions. The purpose of this chapter is to investigate (3.1.1) and related functionals for dependent stationary sequences (X_t). As in the case of maxima, sample autocovariances, ruin and large deviation probabilities, a natural candidate of a heavy-tailed stationary sequence (X_t) is given by a linear process

$$X_t = \sum_{j=0}^{\infty} \psi_j \, Z_{t-j} \,, \quad t \in \mathbb{Z} \,, \tag{3.1.2}$$

for an iid regularly varying sequence (Z_t), i.e. a generic element Z of this sequence[2] satisfies the tail balance condition

$$P(Z > x) = \widetilde{p}\, x^{-\alpha} \, L(x) \quad \text{and} \quad P(Z \le -x) = \widetilde{q}\, x^{-\alpha} \, L(x) \,, \quad x \to \infty \,, \tag{3.1.3}$$

where L is a slowly varying function, $\alpha > 0$ is the index of regular variation and $\widetilde{p}, \widetilde{q} \ge 0$, $\widetilde{p} + \widetilde{q} = 1$. We refer to Z as a regularly varying random variable. It is well known (see e.g. Embrechts et al. (1997)) that, under conditions on (ψ_j) ensuring the a.s. convergence of the infinite series in (3.1.2), the relations $P(X > x) \sim c_+ P(Z > x)$ and $P(X \le -x) \sim c_- P(Z \le -x)$ as $x \to \infty$ hold for constants c_- and c_+ depending on α and (ψ_j); see also Lemma 3.2.2 below. Using a truncation of the infinite series in (3.1.2) and the techniques

[2]Here and in what follows, we denote by X a generic element of any strictly strictly stationary sequence (X_t).

from Mikosch and Račkauskas (2010), in Section 3.4 we derive a Fréchet limit distribution for the maximum increment functional (3.1.1). For the derivation of this result we make heavy use of the linear structure of (X_t). For linear processes, this approach is similar to the investigation of maxima, sample autocovariances, ruin and large deviation probabilities.

On the other hand, Davis and Hsing (1995) provided some extreme value theory for very general strictly stationary sequences with regularly varying finite-dimensional distributions. This theory is less suited for linear processes but it is applicable to sequences which satisfy certain mixing conditions. In Section 3.5 we exploit this theory together with a continuous mapping argument to derive an asymptotic theory for the functional (3.1.1) for general strictly stationary regularly varying sequences. The limit of (3.1.1) will again be a Fréchet distribution. We apply these results to two standard financial time series models: the GARCH and heavy-tailed stochastic volatility models.

The chapter is organized as follows. In Section 3.2.1 we introduce notions such as regular variation of a random vector and a regularly varying sequence. In Section 3.2.2 we give some background on linear processes, including conditions for their existence as well as some results on their limiting behavior in case of a regularly varying noise sequence. In Section 3.3 we recall the results from Mikosch and Račkauskas (2010). They describe the distributional limit behavior of the maximum increment and of related quantities of a random walk with iid regularly varying jump sizes. This result serves as a benchmark in the case of dependent jump sizes. In Section 3.4 we treat the maximum increment of a random walk with jump sizes given by a linear process with regularly varying innovations. In Section 3.5 we consider the maximum increment of a random walk for a general strictly stationary sequence, but we assume certain mixing conditions.

3.2 Preliminaries and Notation

3.2.1 Regular Variation of Random Elements and Stationary Sequences

In this chapter, we describe the heavy distributional tails of a sequence of random variables by the notion of regular variation. For a real-valued random variable Z, this notion was made precise in (3.1.3). However, we will also need regular variation of random vectors: An $\mathbb{R}^d$-valued random vector Z is said to be *regularly varying* with index $\alpha > 0$ if there exist a non-null Radon measure μ on the Borel σ-field $\mathcal{B}_0$ of $\overline{\mathbb{R}}_0^d = \overline{\mathbb{R}}^d \backslash \{0\}$, where $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$

and a non-decreasing sequence (a_n) of positive numbers such that

$$n\,P(a_n^{-1}Z \in \cdot) \xrightarrow{v} \mu(\cdot), \tag{3.2.1}$$

where $\xrightarrow{v}$ denotes vague convergence. A detailed introduction to vague convergence and regular variation has been given in Section 2.1. See also Kallenberg (1983) and Resnick (1987) for details on vague convergence and Resnick (1987), Hult and Lindskog (2005) and Basrak and Segers (2009) for more reading on regular variation. The sequence (a_n) can always be chosen as

$$a_n = \inf\{x \geq 0 : P(|Z| \leq x) \geq 1 - n^{-1}\}\,, \quad n \geq 1\,,$$

and the measure μ necessarily satisfies the relation $\mu(t\cdot) = t^{-\alpha}\mu(\cdot)$ for any $t > 0$.

The defining property (3.2.1) can be equivalently expressed in spherical coordinates:

$$\frac{P(|Z| > x\,u, Z/|Z| \in \cdot)}{P(|Z| > x)} \xrightarrow{w} u^{-\alpha}P(\Theta \in \cdot), \quad u > 0\,, \quad x \to \infty, \tag{3.2.2}$$

where the random vector Θ assumes values in the unit sphere $\mathbb{S}^{d-1} = \{x \in \mathbb{B} : \|x\| = 1\}$ of $\mathbb{R}^d$. The distribution of Θ is called the spectral measure of Z. For $d = 1$, (3.2.2) coincides with (3.1.3).

The strictly stationary sequence (X_t) of real-valued random variables is said to be regularly varying with index $\alpha > 0$ if for every $d \geq 1$ the vector $Z_d = (X_1, \ldots, X_d)$ is regularly varying with index α and limiting measure μ_d, where for all $d \geq 1$ the sequence (a_n) is chosen such that $nP(|X| > a_n) \to 1$. A simple example of a regularly varying sequence with index $\alpha > 0$ is an iid sequence (X_t), where X is regularly varying with index α. The limiting measures μ_d are then concentrated on the axes. More complicated examples of regularly varying sequences (X_t) will be considered in the following sections.

3.2.2 Linear Processes

In this section, we introduce linear processes and state some useful results regarding their existence and limit behavior.

Definition 3.2.1 (Linear Process). Let (ψ_j) be a sequence of real numbers and let (Z_t) be a noise sequence of iid real-valued random variables. If the process (X_t) given by

$$X_t = \sum_{j=0}^{\infty} \psi_j\, Z_{t-j}$$

exists for all $t \in \mathbb{Z}$, it is called a *linear process* (or *Moving Average of infinite order*).

In this chapter we will especially consider the case where the generic element Z of the iid noise sequence (Z_t) is regularly varying with index $\alpha > 0$. If X_t is a genuine infinite series one needs to verify whether the series (3.1.2) converges a.s., and this condition has to be reconciled with the regular variation of Z. In the following lemma we give sufficient conditions for the a.s. convergence and regular variation of X_t, see e.g. Embrechts et al. (1997), Section A3.3. Similar conditions have also been given by Cline (1983), but they can be weakened by condition (3.2.3) below, which is taken from Mikosch and Samorodnitsky (2000a). This condition is close to those dictated by the 3-series theorem.

Lemma 3.2.2. *Let (Z_t) be an iid sequence of regularly varying random variables with index $\alpha > 0$ satisfying the tail balance condition* (3.1.3) *and, if $\alpha > 1$, $\mathbb{E}\, Z = 0$. Moreover, assume*

$$\sum_{i=0}^{\infty} |\psi_i|^p < \infty, \tag{3.2.3}$$

where $p = 2$ for $\alpha > 2$ and $p = \alpha - \delta$ for some $\delta > 0$ for $\alpha \leq 2$. Then the series (3.1.2) *converges a.s., X_t is regularly varying with index $\alpha > 0$ and the following relation holds*

$$\lim_{x\to\infty} \frac{P(X > x)}{P(|Z| > x)} = \sum_{i=1}^{\infty} \left(\widetilde{p}\,(\psi_j)_+^{\alpha} + \widetilde{q}\,(\psi_j)_-^{\alpha} \right).$$

We mention that the linear process (X_t) with regularly varying innovation sequence is then also regularly varying (see e.g. Hult and Samorodnitsky (2008), where this is proved for multivariate linear processes with random coefficients). However, this property will not be crucial in the sequel.

The limit theory for linear processes with a regularly varying noise sequence has been developed by Davis and Resnick (1985). It is based on a point process result adapted to the special setting of regularly varying linear processes. Therefore, denote by $M_p(E)$ the space of all point measures on the space E and let ϑ be the measure on $(0,\infty) \times \mathbb{R}\backslash\{0\}$ given by

$$\vartheta(dt, dx) = dt \times (\alpha\, \widetilde{p}\, x^{-\alpha-1} \mathbb{1}_{(0,\infty)}(x)\, dx + \alpha\, \widetilde{q}\, (-x)^{-\alpha-1} \mathbb{1}_{(-\infty,0)}(x)\, dx),$$

where $\widetilde{p}$ and $\widetilde{q}$ are given by the tail balance condition (3.1.3). For details on point processes, we refer to Section 2.2.

Theorem 3.2.3 (Davis and Resnick (1985), Theorem 2.4). *Let (X_t), $t \in \mathbb{Z}$, be a linear process that exists with respect to Lemma 3.2.2 and let the noise sequence (Z_t) be*

regularly varying, i.e. the tail balance condition (3.1.3) *is satisfied and there exists a positive sequence* (a_n) *such that* (3.2.1) *holds. Additionally, assume that* (3.2.3) *holds for* $p = 1$ *and that* $\{(t_k, j_k)\}$, $k \in \mathbb{Z}$, *are the points of a PRM*(ϑ) *on* $(0,\infty) \times \mathbb{R}\backslash\{0\}$.

(i) In $M_p((0,\infty) \times \mathbb{R}\backslash\{0\})$ *it holds, as* $n \to \infty$,

$$\sum_{k=1}^{\infty} \varepsilon_{(k/n,\, X_k/a_n)} \xrightarrow{d} \sum_{i=0}^{\infty} \sum_{k=1}^{\infty} \varepsilon_{(t_k,\, j_k \psi_i)}.$$

(ii) In $M_p((0,\infty) \times \mathbb{R}^{l+1}\backslash\{(0,0,\ldots,0)\})$ *we have, as* $n \to \infty$,

$$\sum_{k=1}^{\infty} \varepsilon_{(k/n,\, (X_k, X_{k-1}, \ldots, X_{k-l})/a_n)} \xrightarrow{d} \sum_{i=0}^{\infty} \sum_{k=1}^{\infty} \varepsilon_{(t_k,\, j_k(\psi_i, \psi_{i-1}, \ldots, \psi_{i-l}))}$$

for any $l \in \mathbb{N}$.

The theorem can now be used to derive several properties of the linear process. In particular, the joint limit distribution function of the sample maxima $M_n = \max\{X_1, \ldots, X_n\}$ and the sample minima $W_n = \min\{X_1, \ldots, X_n\}$ can be given. Set

$$\psi_+ = \max_{j \geq 0}\{\psi_j \vee 0\} \quad \text{and} \quad \psi_- = \max_{j \geq 0}\{-\psi_j \vee 0\}.$$

Theorem 3.2.4 (Davis and Resnick (1985), Theorem 3.2). *Again, let* (X_t), $t \in \mathbb{Z}$, *be a linear process that exists with respect to Lemma 3.2.2 and let the noise sequence* (Z_t) *be regularly varying, i.e. the tail balance condition* (3.1.3) *is satisfied and there exists a positive sequence* (a_n) *such that* (3.2.1) *holds. Then for all* $x, y \in \mathbb{R}$, *as* $n \to \infty$,

$$P\left(a_n^{-1} M_n \leq x,\ a_n^{-1} W_n \leq y\right) \to G^{\widetilde{p}}(x,\infty)\, G^{\widetilde{q}}(\infty, x) - G^{\widetilde{p}}(x, -y)\, G^{\widetilde{q}}(-y, x),$$

where

$$G(x,y) = \begin{cases} \exp(-\psi_+^{\alpha}\, x^{-\alpha}) \wedge \exp(-\psi_-^{\alpha}\, y^{-\alpha}) & \text{if } x > 0 \text{ and } y > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 3.2.4 directly leads to a result for a limit distribution of M_n.

Corollary 3.2.5. *Under the setting of Theorem 3.2.4, we have*

$$P\left(a_n^{-1} M_n \leq x\right) \to \Phi_{\alpha}\left(\frac{x}{\psi_+^{\alpha}\, \widetilde{p} + \psi_-^{\alpha}\, \widetilde{q}}\right),$$

where $\Phi_{\alpha}(x) = \exp\{-x^{-\alpha}\}$, $x > 0$, *denotes the* Fréchet distribution function .

Proof. The result follows from the observation

$$P\left(a_n^{-1}M_n \le x\right) = P\left(a_n^{-1}M_n \le x,\ a_n^{-1}W_n \le \infty\right).$$

□

Similarly, maxima of absolute values can be considered. Therefore, define the absolute sample maxima by $\widetilde{M}_n = \max\{|X_1|, \ldots, |X_n|\}$ and set $\psi = \max_{j\ge 0}|\psi_j|$.

Corollary 3.2.6. *Suppose that the conditions of Theorem 3.2.4 hold. Then*

$$P\left(a_n^{-1}\widetilde{M}_n \le x\right) \to \Phi_\alpha\left(\frac{x}{\psi}\right) = \Phi_\alpha^{\psi^\alpha}(x),$$

where Φ *again denotes the Fréchet distribution function.*

Proof. We observe that $\psi = \psi_+ \vee \psi_-$ and calculate

$$\begin{aligned}
P\left(a_n^{-1}\widetilde{M}_n \le x\right) &= P\left(a_n^{-1}M_n \le x,\ a_n^{-1}W_n \ge -x\right)\\
&= P\left(a_n^{-1}M_n \le x\right) - P\left(a_n^{-1}M_n \le x,\ a_n^{-1}W_n < -x\right)\\
&= G^{\widetilde{p}}(x,x)\, G^{\widetilde{q}}(x,x)\\
&= \left(e^{-\widetilde{p}\,\psi_+^\alpha\, x^{-\alpha}} \wedge e^{-\widetilde{p}\,\psi_-^\alpha\, x^{-\alpha}}\right)\left(e^{-\widetilde{q}\,\psi_+^\alpha\, x^{-\alpha}} \wedge e^{-\widetilde{q}\,\psi_-^\alpha\, x^{-\alpha}}\right)\\
&= e^{-\widetilde{p}\,(\psi_+^\alpha \vee \psi_-^\alpha)\, x^{-\alpha}}\ e^{-\widetilde{q}\,(\psi_+^\alpha \vee \psi_-^\alpha)\, x^{-\alpha}}\\
&= e^{-\,\psi^\alpha\, x^{-\alpha}}.
\end{aligned}$$

□

Several other limit results can be obtained as consequences of Theorem 3.2.3, e.g. convergence to extremal processes, convergence of the point process of exceedances to compound Poisson processes etc. For details we refer to Davis and Resnick (1985), Section 3. Limit results for covariance and correlation functions of linear processes with regularly varying noise can also be found in Davis and Resnick (1986).

3.3 Random Walks with Independent Jump Sizes

In Mikosch and Račkauskas (2010) the limit distribution of the maximum increment of a random walk with iid regularly varying Banach-valued jump sizes was studied. We recall some of the results as a benchmark for the case of dependent jump sizes, but we restrict ourselves to real-valued random variables X_t.

For an iid sequence (X_t) of random variables we define the corresponding *random walk*

$$S_0 = 0 \quad \text{and} \quad S_n = X_1 + \cdots + X_n \quad \text{for} \quad n \ge 1,$$

and $(\overline{X}_n)$ denotes the corresponding sequence of the sample means.

Definition 3.3.1 (Maximum Increment). The *maximum increment* of the random walk (S_n) is given by the quantity

$$\widetilde{M}_n^{(\gamma)} := \max_{1\le l\le n} (f(l))^{-1} \max_{0\le k\le n-l} |S_{k+l} - S_k|\,, \quad n \ge 1\,, \tag{3.3.1}$$

and by the corresponding centered version

$$\widetilde{T}_n^{(\gamma)} := \max_{1\le l\le n} (f(l(1-l/n)))^{-1} \max_{0\le k\le n-l} |S_{k+l} - S_k - l\overline{X}_n|\,, \quad n \ge 1, \tag{3.3.2}$$

we denote the *centered maximum increment* of (S_n).

The functions f are chosen from the class $\mathcal{F}_\gamma$, $\gamma \ge 0$, given by

$$\begin{aligned}\mathcal{F}_\gamma := \big\{ & f : f \text{ is a non-decreasing function on } [0,\infty),\ f(1)=1, f(l)\ge l^\gamma \text{ for } l\ge 1 \text{ and} \\ & \text{for any increasing sequence } (d_n) \text{ of positive numbers such that } d_n^2/n \to 0, \\ & \text{the following relation holds } \lim_{n\to\infty} \inf_{1\le l\le d_n} f(l(1-l/n))/f(l) = 1\big\}.\end{aligned}$$

The class $\mathcal{F}_\gamma$, $\gamma > 0$, contains the functions $f(x) = x^{\gamma'}$ for $\gamma' \ge \gamma$, and $f(x) = x^\gamma \log^\beta(1+x)$ for $\beta > 0$.

The main result of Mikosch and Račkauskas (2010) in the case of real-valued random variables is the following.

Theorem 3.3.2 (Mikosch and Račkauskas (2010), Theorem 2.2). *Let (X_t) be a sequence of iid random variables which are regularly varying with index $\alpha > 0$. In addition, assume $\mathbb{E}\,X = 0$ if $\mathbb{E}\,|X| < \infty$. Then, for $f \in \mathcal{F}_\gamma$, $\gamma > \max(0, 0.5 - \alpha^{-1})$, with the normalizing sequence (a_n) chosen such that $nP(|X| > a_n) \to 1$,*

$$\begin{aligned}\lim_{x\to\infty} P(a_n^{-1}\widetilde{M}_n^{(\gamma)} \le x) &= \Phi_\alpha(x)\,, \quad x > 0\,, \\ \lim_{x\to\infty} P(a_n^{-1}\widetilde{T}_n^{(\gamma)} \le x) &= \Phi_\alpha(x)\,, \quad x > 0\,,\end{aligned}$$

where $\Phi_\alpha(x) = \exp\{-x^{-\alpha}\}$, $x > 0$, denotes the Fréchet *distribution function.*

Remark 3.3.3. It is known from classical extreme value theory (e.g. Section 2.1.2 or Embrechts et al. (1997), Chapter 3) that regular variation of $|X|$ is necessary and sufficient for the convergence in distribution of the sequence of normalized partial maxima $(a_n^{-1} \max_{i=1,\dots,n} |X_i|)$ towards a Φ_α-distributed random variable. Hence, for $l \ge 2$ the normalized random walk increments $a_n^{-1}(f(l))^{-1} \max_{0\le k\le n-l} |S_{k+l} - S_k|$ do not contribute to the limits of $a_n^{-1}\widetilde{M}_n^{(\gamma)}$ in Theorem 3.3.2.

Remark 3.3.4. The results in Mikosch and Račkauskas (2010) are sharp in the following sense. If $f(l) = l^\gamma$ for some $\gamma < 0.5 - \alpha^{-1}$ and some $\alpha > 2$, then an application of the invariance principle in Hölder space yields

$$n^{-0.5+\gamma}\widetilde{M}_n^{(\gamma)} \xrightarrow{d} \sup_{s,t\in[0,1],s\neq t} \frac{|W(t)-W(s)|}{|t-s|^\gamma},$$

where W is a Brownian motion on $[0,1]$; see Račkauskas and Suquet (2006).
One of the original motivating ideas for considering limit theory for statistics of the type $\widetilde{M}_n^{(\gamma)}$ or $\widetilde{T}_n^{(\gamma)}$ was to use these statistics for detecting epidemic changes in a sample. We refer to the monograph Csörgö and Horváth (1997) as a general reference to change point problems and the recent papers Račkauskas and Suquet (2004) and Račkauskas and Suquet (2006) for advanced limit theory in the context of epidemic changes.

Besides $\widetilde{M}_n^{(\gamma)}$ and $\widetilde{T}_n^{(\gamma)}$, several one-sided characteristics have also been discussed in Mikosch and Račkauskas (2010). Define

$$\begin{aligned} M_n^{(\gamma)} &:= \max_{1\le l\le n}(f(l))^{-1}\max_{0\le k\le n-l}(S_{k+l}-S_k) \quad \text{and} \\ r_n^{(\gamma)} &:= \min_{1\le l\le n}(f(l))^{-1}\min_{0\le k\le n-l}(S_{k+l}-S_k). \end{aligned}$$

Theorem 3.3.5 (Mikosch and Račkauskas (2010), Theorem 2.10). *Let (X_t) be a sequence of iid real-valued random variables with common distribution function F which is regularly varying with index $\alpha > 0$ in the sense of (3.1.3) and assume $\mathbb{E}X = 0$ if $\mathbb{E}|X| < \infty$. Then, for $b_n = \widetilde{p}^{\,1/\alpha}a_n$, where (a_n) is chosen such that $nP(|X| > a_n) \to 1$, $f \in \mathcal{F}_\gamma$ and $\gamma > \max(0, 0.5 - \alpha^{-1})$*

$$\lim_{n\to\infty} P\left(b_n^{-1}r_n^{(\gamma)} \le -x,\ b_n^{-1}M_n^{(\gamma)} \le y\right) = \Phi_\alpha(y)\left(1-\Phi_\alpha^{\widetilde{p}/\widetilde{q}}(x)\right),$$

where $x, y > 0$.

One can also consider the centered characteristic

$$\widehat{T}_n^{(\gamma)} := \max_{1\le l\le n}(l(1-l/n))^{-\gamma}\max_{0\le k\le n-l}(S_{k+l}-S_k-l\overline{X}_n)$$

which is the one-sided version of $\widetilde{T}_n^{(\gamma)}$.

Theorem 3.3.6 (Mikosch and Račkauskas (2010), Theorem 2.12). *Let (X_t) be a sequence of iid real-valued random variables with common distribution function F which is regularly varying with index $\alpha > 0$ in the sense of (3.1.3) and assume $\mathbb{E}X = 0$ if*

$\mathbb{E}\,|X| < \infty$. *Then, for* $b_n = \widetilde{p}^{\,1/\alpha} a_n$, *where* (a_n) *is chosen such that* $nP(|X| > a_n) \to 1$, $\gamma > \max(0, 0.5 - \alpha^{-1})$ *and* $x > 0$

$$\lim_{n\to\infty} P\left(b_n^{-1}\widehat{T}_n^{(\gamma)} \le x\right) = \Phi_\alpha(x).$$

A symmetric version with similar structure than $\widetilde{M}_n^{(\gamma)}$ can be given by

$$\widehat{M}_n^{(\gamma)} := \max_{1\le l\le n}(f(l)))^{-1} \max_{l+1\le k\le n-l}(S_{k+l} - S_{k-l} - 2S_k).$$

As the random variables $S_{k+l} - S_{k-l} - 2S_k$ are symmetric, the condition $\mathbb{E}\,X = 0$ can be dropped.

Theorem 3.3.7 (Mikosch and Račkauskas (2010), Theorem 2.13). *Let* (X_t) *be a sequence of iid real-valued random variables with common distribution function* F *which is regularly varying with index* $\alpha > 0$ *in the sense of* (3.1.3). *Then, for* (a_n) *chosen such that* $nP(|X| > a_n) \to 1$, $f \in \mathcal{F}_\gamma$ *and* $\gamma > \max(0, 0.5 - \alpha^{-1})$

$$\lim_{n\to\infty} P\left(a_n^{-1}\widehat{M}_n^{(\gamma)} \le x\right) = \Phi_\alpha^2(x),$$

where $x > 0$.

3.4 Random Walks with Linearly Dependent Jump Sizes

3.4.1 The Jump Size is a Finite Moving Average

In this section, we derive the limit distribution of the quantities $\widetilde{M}_n^{(\gamma)}$ and $\widetilde{T}_n^{(\gamma)}$ defined in (3.3.1) and (3.3.2), respectively, for a *moving average of finite order* q, i.e.

$$X_t = X_t^{(q)} = \sum_{i=1}^{q} \psi_i Z_{t-i}\,, \quad t \in \mathbb{Z}\,, \tag{3.4.1}$$

for an iid sequence (Z_t). The following result is the analog of Theorem 3.3.2 for moving averages.

Theorem 3.4.1. *Let* (X_t) *be a moving average process of order* $q > 1$ *with an iid noise sequence* (Z_t) *of regularly varying random variables with index* $\alpha > 0$. *If* $\mathbb{E}\,|Z| < \infty$ *we also assume* $\mathbb{E}\,Z = 0$. *Then, for* $f \in \mathcal{F}_\gamma$, $\gamma > \max(0, 0.5 - 1/\alpha)$, *the following relations hold:*

$$\lim_{n\to\infty} P(a_n^{-1}\widetilde{M}_n^{(\gamma)} \le x) = \Phi_\alpha^{m_q^\alpha}(x), \quad x > 0, \tag{3.4.2}$$

$$\lim_{n\to\infty} P(a_n^{-1}\widetilde{T}_n^{(\gamma)} \le x) = \Phi_\alpha^{m_q^\alpha}(x), \quad x > 0, \tag{3.4.3}$$

where Φ_α *again denotes the Fréchet distribution,* (a_n) *is chosen such that* $n\,P(|Z| > a_n) \to 1$ *and*

$$m_q := \max_{1\le l\le q} \max_{1\le k\le q-l+1} \frac{|\psi_{k,k+l-1}|}{f(l)} \quad \textit{with} \quad \psi_{i,j} := \sum_{k=i}^{j} \psi_k, \quad \textit{for} \quad 1 \le i \le j \le q.$$

The proofs of this and the other results of this section will be given in Section 3.4.3.

Remark 3.4.2. For $\gamma \ge 1$, we have $l/f(l) \le 1$, $l \ge 1$, and then direct calculation shows that

$$\frac{S_{k+l} - S_k}{f(l)} = \frac{X_{k+1} + \cdots + X_{k+l}}{f(l)} \le \frac{l \max_{1\le i\le l} X_{k+i}}{f(l)} \le \max_{1\le i\le l} X_{k+i}$$

for all $1 \le l \le n$, $0 \le k \le n-l$ and thus $\widetilde{M}_n^{(\gamma)} = \max_{t=1,\ldots,n} |X_t|$. Hence, it follows from Davis and Resnick (1985) that $\lim_{n\to\infty} P(a_n^{-1}\widetilde{M}_n^{(\gamma)} \le x) = \Phi_\alpha^{\psi_q^\alpha}(x)$, where $\psi_q = \max_{q\ge i\ge 1} |\psi_i|$. This is in agreement with (3.4.2), since, for $\gamma \ge 1$,

$$\max_{1\le k\le q} |\psi_k| = \max_{1\le k\le q} (f(1))^{-1}|\psi_k| \le m_q \le \max_{1\le l\le q}(l/f(l)) \max_{1\le k\le q} |\psi_k| = \max_{1\le k\le q} |\psi_k|\,.$$

A corresponding remark applies to (3.4.3).

Similar to Section 3.3, it is also possible to derive results for quantities based on the one-sided increments of a random walk, for example,

$$\begin{aligned} M_n^{(\gamma)} &:= \max_{1\le l\le n} (f(l))^{-1} \max_{0\le k\le n-l} (S_{k+l} - S_k)\,, \\ r_n^{(\gamma)} &:= \min_{1\le l\le n} (f(l))^{-1} \min_{0\le k\le n-l} (S_{k+l} - S_k)\,, \\ T_n^{(\gamma)} &:= \max_{1\le l\le n} f(l(1-l/n))^{-1} \max_{1\le k\le n-l} (S_{k+l} - S_k - l\overline{X}_n)\,. \end{aligned}$$

Theorem 3.4.3. *Assume the conditions of Theorem 3.4.1. Choose* (b_n) *such that* $nP(Z > b_n) \to 1$*, where we also assume that* $\widetilde{p} > 0$*. Then, for* $f \in \mathcal{F}_\gamma$*,* $\gamma > \max(0, 0.5 - 1/\alpha)$ *and* $x, y > 0$*, the following limit relations hold*

$$\begin{aligned} \lim_{n\to\infty} P\left(b_n^{-1} r_n^{(\gamma)} \le -x\,,\ b_n^{-1} M_n^{(\gamma)} \le y\right) &= \Phi_\alpha^{(m_q^+)^\alpha}(y)\left(1 - \Phi_\alpha^{(\widetilde{q}/\widetilde{p})(m_q^-)^\alpha}(x)\right), \\ \lim_{n\to\infty} P(b_n^{-1} T_n^{(\gamma)} \le x) &= \Phi_\alpha^{(m_q^+)^\alpha}(x)\,, \end{aligned}$$

where

$$\begin{aligned} m_q^+ &= \max_{1\le l\le q} \max_{1\le k\le q-l+1} \frac{(\psi_{k,k+l-1})_+}{f(l)}\,, \\ m_q^- &= \max_{1\le l\le q} \max_{1\le k\le q-l+1} \frac{(\psi_{k,k+l-1})_-}{f(l)}\,. \end{aligned}$$

3.4.2 The Jump Sizes Constitute a Linear Process

In this section we analyze the limit distribution of the maximum increment of a random walk whose jump sizes constitute an infinite order moving average, i.e. $q = \infty$ in (3.4.1). Then the limit distributions of $\widetilde{M}_n^{(\gamma)}$ and $\widetilde{T}_n^{(\gamma)}$ are not straightforward consequences of Theorem 3.4.1 by simply letting $q \to \infty$.

Throughout we assume that the infinite series X_t in (3.4.1) is finite a.s.; Lemma 3.2.2 provides sufficient conditions. In what follows, we assume additional conditions on the coefficients (ψ_j):

$$\sum_{j=3}^{\infty} |\psi_j| \ell(j) < \infty, \tag{3.4.4}$$

where

$$\ell(j) = \begin{cases} (j \log\log j)^{1/2} & \text{if } \operatorname{var}(Z) < \infty \\ j^{1/p} & \text{for some } p < \alpha \text{ if } \alpha < 2 \text{ or } \alpha = 2 \text{ and } \operatorname{var}(Z) = \infty. \end{cases}$$

In view of Lemma 3.2.2, (3.4.4) implies the a.s. convergence of the series X_t for $\alpha > 1$ if in addition $\mathbb{E}\, Z = 0$.

Theorem 3.4.4. *Let (X_n) be a linear process (3.1.2) with iid regularly varying noise sequence (Z_t) with index $\alpha > 0$. If $\mathbb{E}\,|Z| < \infty$ we also assume $\mathbb{E}\, Z = 0$. Furthermore, assume (3.4.4), and if $\alpha \le 1$ also $\sum_{j=0}^{\infty} |\psi_j|^{\alpha-\delta} < \infty$ for some $\delta < \alpha$. Then for $f \in \mathcal{F}_\gamma$, $\gamma > \max(0, 0.5 - 1/\alpha)$ we have*

$$\lim_{n\to\infty} P(a_n^{-1} \widetilde{M}_n^{(\gamma)} \le x) = \Phi_\alpha^{m_\infty^\alpha}(x), \quad x > 0, \tag{3.4.5}$$

$$\lim_{n\to\infty} P(a_n^{-1} \widetilde{T}_n^{(\gamma)} \le x) = \Phi_\alpha^{m_\infty^\alpha}(x), \quad x > 0\,, \tag{3.4.6}$$

where

$$m_\infty := \lim_{q\to\infty} m_q = \max_{l\ge 1} \max_{k\ge 1} \frac{|\psi_{k,k+l-1}|}{f(l)}.$$

Remark 3.4.5. From the proofs of Theorems 3.4.1, 3.4.3 and 3.4.4 it is straightforward that Theorem 3.4.3 extends to the infinite order moving average case as well provided the conditions of Theorem 3.4.4 for (ψ_j) are satisfied and the constants m_q^+ and m_q^- are replaced by

$$m_\infty^+ = \lim_{q\to\infty} m_q^+ = \max_{l\ge 1} \max_{k\ge 1} \frac{(\psi_{k,k+l-1})_+}{f(l)} \quad \text{and} \quad m_\infty^- = \lim_{q\to\infty} m_q^- = \max_{l\ge 1} \max_{k\ge 1} \frac{(\psi_{k,k+l-1})_-}{f(l)},$$

respectively.

Remark 3.4.6. In the case $\gamma \geq 1$, a remark corresponding to Remark 3.4.2 can be given. Again we have $l/f(l) \leq 1$, $l \geq 1$, and thus $\widetilde{M}_n^{(\gamma)} = \max_{t=1,\ldots,n} |X_t|$. Now it follows from Corollary 3.2.6 that $\lim_{n\to\infty} P(a_n^{-1}\widetilde{M}_n^{(\gamma)} \leq x) = \Phi_\alpha^{\psi^\alpha}(x)$, where $\psi = \max_{i\geq 1} |\psi_i|$. On the other hand,

$$\psi = \max_{k\geq 1}(f(1))^{-1}|\psi_k| \leq m_\infty \leq \max_{l\geq 1}(l/f(l)) \max_{k\geq 1} |\psi_k| = \max_{k\geq 1} |\psi_k| = \psi$$

and hence, the result of (3.4.5) coincides with the classical results of Davis and Resnick (1985).

We mention in passing that Juodis et al. (2009) and Račkauskas and Suquet (2010) proved functional central limit theorems in Hölder space for partial sum processes of linear processes. Results of this type yield limits for $\widetilde{M}_n^{(\gamma)}$ and $\widetilde{T}_n^{(\gamma)}$ when $\alpha > 2$ and $\gamma < 0.5 - \alpha^{-1}$. The mentioned results show that the statements of Theorem 3.4.4 are sharp in the sense that different limit distributions appear when the normalizing functions $f(l)$ increase too slowly.

3.4.3 Proofs

Proof of Theorem 3.4.1

The proof uses arguments similar to the ones in the proof of Theorem 2.2 in Mikosch and Račkauskas (2010). Lemma 3.4.7 in this section replaces Lemma 2.4 in Mikosch and Račkauskas (2010).

Lemma 3.4.7. *Assume the conditions of Theorem 3.4.1 hold for (Z_t) and (X_t). Then the following statements hold.*

(1) For any $f \in \mathcal{F}_\gamma$, $\gamma \geq 0$ and $h \geq 1$,

$$\lim_{n\to\infty} P\Big(a_n^{-1} \max_{1\leq l\leq h} (f(l))^{-1} \max_{0\leq k\leq n-l} |S_{k+l} - S_k| \leq x\Big) = \Phi_\alpha^{m_{q,h}^\alpha}(x),$$

where

$$m_{q,h} := \max_{1\leq l\leq h\wedge q} \max_{1\leq k\leq q-l+1} \frac{|\psi_{k,k+l-1}|}{f(l)}.$$

(2) For any $\delta > 0$ and $f \in \mathcal{F}_\gamma$, $\gamma > \max\{0, 0.5 - \alpha^{-1}\}$,

$$\lim_{h\to\infty} \limsup_{n\to\infty} P\left(\max_{h\leq l\leq n} (f(l))^{-1} \max_{0\leq k\leq n-l} |S_{k+l} - S_k| > \delta a_n\right) = 0.$$

Remark 3.4.8. Notice that for $h \geq q$, $m_{q,h} = m_q$ as defined in Theorem 3.4.1.

Proof. **(1)** The proof uses arguments from the proof of Lemma 2.4(1) in Mikosch and Račkauskas (2010) and a continuous mapping argument for point processes, going back to Davis and Resnick (1985). Write $M_p(E)$ for the set of point measures on a Borel state space E of some Euclidean space, and equip $M_p(E)$ with the vague topology. Introduce the point processes

$$\widetilde{N}_n^{(h)} := \sum_{t=1}^{n} \varepsilon_{a_n^{-1}(Z_{t-q}, Z_{t-q+1}, \ldots, Z_{t+h-2})} \quad \text{and} \quad N_n^{(h)} := \sum_{t=1}^{n} \varepsilon_{a_n^{-1}(X_t, X_t+X_{t+1}, \ldots, X_t+\cdots+X_{t+h-1})} ,$$

on $M_p(\overline{\mathbb{R}}_0^{q+h-1})$ and $M_p(\overline{\mathbb{R}}_0^{h})$, respectively. It follows from Theorem 2.2.11 that

$$\widetilde{N}_n^{(h)} \xrightarrow{d} \widetilde{N}^{(h)} := \sum_{k=1}^{q+h-1} \sum_{i=1}^{\infty} \varepsilon_{J_i \mathbf{e}_k} , \quad n \to \infty ,$$

where (J_i) are the points of a Poisson random measure on $\overline{\mathbb{R}}_0$ with intensity

$$\alpha |x|^{\alpha-1} [\widetilde{p}\, \mathbb{1}_{(0,\infty]}(x) + \widetilde{q}\, \mathbb{1}_{[-\infty,0)}(x)] dx$$

and $\mathbf{e}_k$ is the kth unit vector in $\mathbb{R}^{q+h-1}$. According to the proof of Theorem 2.4 in Davis and Resnick (1985), the continuous function

$$\begin{aligned}(Z_{t-q}, Z_{t-q+1}, \ldots, Z_{t+h-2}) \mapsto & \left(\sum_{k=t}^{t} \sum_{i=1}^{q} \psi_i Z_{k-i}, \sum_{k=t}^{t+1} \sum_{i=1}^{q} \psi_i Z_{k-i}, \ldots, \sum_{k=t}^{t+h-1} \sum_{i=1}^{q} \psi_i Z_{k-i} \right) \\ & = (X_t, X_t + X_{t+1}, \ldots, X_t + \cdots + X_{t+h-1}),\end{aligned}$$

induces a continuous mapping on the limit relation $\widetilde{N}_n^{(h)} \xrightarrow{d} \widetilde{N}^{(h)}$, resulting in the following convergence in $M_p(\overline{\mathbb{R}}_0^{h})$

$$N_n^{(h)} \xrightarrow{d} N^{(h)} = \sum_{k=2-h}^{q} \sum_{i=1}^{\infty} \varepsilon_{J_i(\psi_{1\vee k, q\wedge k}, \psi_{1\vee k, q\wedge(k+1)}, \ldots, \psi_{1\vee k, q\wedge(k+h-1)})} , \tag{3.4.7}$$

where we set $\psi_{i,j} = 0$ if $j < 1$. As in Mikosch and Račkauskas (2010) we write for $l \geq 1$, $\widetilde{M}_{nl} = \max\limits_{0\leq k\leq n} |S_{k+l} - S_k|$ and introduce the sets

$$B_f(y) := \left\{ (x_1, \ldots, x_h) \in \mathbb{R}_0^h : |x_i| \leq y f(i) \text{ for } i = 1, \ldots, h \right\} , \quad y > 0 .$$

Write $J = \sup_{i\geq 1} |J_i|$. It follows from (3.4.7) and the monotonicity of f that

$$P(a_n^{-1} \max_{1\leq l\leq h} (f(l))^{-1} \widetilde{M}_{nl} \leq y) =$$

$$
\begin{aligned}
&= P(a_n^{-1}\widetilde{M}_{n1} \le yf(1), \ldots, a_n^{-1}\widetilde{M}_{nh} \le yf(h)) \\
&= P(N_n^{(h)}((B_f(y))^c) = 0) \\
&\to P(N^{(h)}((B_f(y))^c) = 0) \\
&= P\Big(J \max_{1\le k\le q} |\psi_{1\vee k, q\wedge k}| \le yf(1), J \max_{0\le k\le q} |\psi_{1\vee k, q\wedge(k+1)}| \le yf(2), \ldots, \\
&\qquad\qquad J \max_{2-h\le k\le q} |\psi_{1\vee k, q\wedge(k+h-1)}| \le yf(h)\Big) \\
&= P(J \le y/m_{q,h}) \\
&= \Phi_\alpha(y/m_{q,h}).
\end{aligned}
$$

This concludes the proof of the first statement.

(2) Let $(\widetilde{S}_n)_{n\ge 0}$ denote the random walk generated by the iid sequence (Z_t). Observe that for $1 \le l \le n$ and some constant $c > 0$ depending on (ψ_j), using the stationarity of the iid sequence (Z_t),

$$
\begin{aligned}
\max_{h\le l\le n} \max_{k\le n-l} |S_{k+l} - S_k| &\le c \max_{h\le l\le n} \max_{k\le n-l} \max_{1\le i\le q} \Big| \sum_{j=k+1}^{k+l} Z_{j-i} \Big| \\
&\stackrel{d}{=} c \max_{h\le l\le n} \max_{k\le n-l} \max_{1\le i\le q} |\widetilde{S}_{k+q-i+l} - \widetilde{S}_{k+q-i}| \\
&\le c \max_{h\le l\le n} \max_{0\le k\le n+q-l} |\widetilde{S}_{k+l} - \widetilde{S}_k| \,.
\end{aligned}
$$

An application of Lemma 2.4(2) in Mikosch and Račkauskas (2010) yields that for every $\delta > 0$,

$$
\begin{aligned}
\lim_{h\to\infty} \limsup_{n\to\infty} P\Big(\max_{h\le l\le n} \max_{0\le k\le n-l} |S_{k-l} - S_k| > \delta a_n \Big) &\le \\
&\le \lim_{h\to\infty} \limsup_{n\to\infty} P\Big(\max_{h\le l\le n+q} \max_{0\le k\le n+q-l} |\widetilde{S}_{k+l} - \widetilde{S}_k| > a_n\delta \Big) \\
&= 0.
\end{aligned}
$$

This concludes the proof of the lemma.

□

Relation (3.4.2) is now a straightforward consequence of Lemma 3.4.7.

For the proof of (3.4.3), we can use the following lemma which is an application of the Remarks 2.5 and 2.6 of Mikosch and Račkauskas (2010) to the setting of finite moving averages. It shows the connection between the limit distributions of $\widetilde{M}_n^{(\gamma)}$ and $\widetilde{T}_n^{(\gamma)}$ that we need for the proof of (3.4.3).

Lemma 3.4.9. *Let (X_n) be the Moving Average process of order $q < \infty$ and let*

$$\zeta_n^{(\gamma)} := \max_{1\le l\le n} f(l)^{-1} \max_{0\le k\le n-l} |S_{k+l} - S_k - l\overline{X}_n|.$$

(1) If $\gamma \ge 1$, we obtain

$$\widetilde{M}_n^{(\gamma)} = \max_{1\le k\le n} |X_k| \quad \text{and} \quad \zeta_n^{(\gamma)} = \max_{1\le k\le n} |X_k - \overline{X}_n|$$

and thus the relations (3.4.2) *and* (3.4.3) *are trivially satisfied.*

(2) If the assumptions of Theorem 3.4.1 hold, then the sequences $a_n^{-1}\widetilde{M}_n^{(\gamma)}$ and $a_n^{-1}\zeta_n^{(\gamma)}$ have the same limit distribution.

Proof. **(1)** Using the assumption $\gamma \ge 1$ and the property $f(l) \ge l^\gamma$ we conclude

$$\begin{aligned}
\max_{1\le k\le n} |X_k - \overline{X}_n| &\le \max_{1\le l\le n} f(l)^{-1} \max_{0\le k\le n-l} |S_{k+l} - S_k - l\overline{X}_n| \\
&= \zeta_n^{(\gamma)} \\
&\le \max_{1\le l\le n} f(l)^{-1} \max_{0\le k\le n-l} \sum_{i=k+1}^{k+l} |X_i - \overline{X}_n| \\
&\le \max_{1\le l\le n} f(l)^{-1} l \max_{1\le k\le n} |X_k - \overline{X}_n| \\
&\le \max_{1\le k\le n} |X_k - \overline{X}_n|
\end{aligned}$$

and likewise for $\widetilde{M}_n^{(\gamma)}$, see also Remark 3.4.2. Then (3.4.2) follows by classical extreme value theory for Moving Averages, see Resnick (1987), Chapter 4.5, and Davis and Resnick (1985), and (3.4.3) follows in view of the proof of Lemma 3.4.9 *(2)*.

(2) We make use of the relation

$$\left|\overline{X}_n\right| = \Big|\sum_{i=1}^{q} \psi_i \frac{1}{n}\sum_{j=1}^{n} Z_{j-i}\Big| \le \left(\sum_{i=1}^{q} |\psi_i|\right) \max_{1\le i\le q} \Big|\frac{1}{n}\sum_{j=1}^{n} Z_{j-i}\Big|.$$

Since $1/n\sum_{j=1}^{n} Z_{j-i}$ has the same distribution as $\overline{Z}_n$ for all $i = 1, \dots, q$, we can follow the argumentation of Mikosch and Račkauskas (2010), Remark 2.6, which shows that $a_n^{-1}|\overline{Z}_n| \xrightarrow{P} 0$ in the case $\gamma \ge 1$ and $a_n^{-1} \max_{1\le l\le n} l^{1-\gamma}|\overline{Z}_n| \xrightarrow{P} 0$ for $\gamma \in (0,1)$.

□

This concludes the proof of (3.4.3), since the same arguments as in the iid case apply, using that (X_t) is a finite moving average process.

Proof of Theorem 3.4.3

The proof is similar to the one of Theorem 3.4.1. We focus on the proof of the joint limit of $b_n^{-1}(r_n^{(\gamma)}, M_n^{(\gamma)})$.

Lemma 3.1 in Mikosch and Račkauskas (2010) is still valid if we substitute a_n by b_n. The only difference is that we get a slightly different limiting Poisson random measure with mean measure μ given by $\mu(dx) = \alpha|x|^{-\alpha}[\mathbb{1}_{(0,\infty)}(x) + (\widetilde{q}/\widetilde{p})\mathbb{1}_{[-\infty,0)}(x)]dx$. Now we follow the proof of Lemma 3.4.7, replace (a_n) by (b_n) everywhere and use the same notation. Write

$$B_f(x,y) := \big\{(z_1,\dots,z_h) \in \mathbb{R}^h : -xf(i) \le z_i \le yf(i) \text{ for } i = 1,\dots,h\big\}\,, \quad x,y>0\,.$$

We obtain

$$\begin{aligned}
&P(N_n^{(h)}((B_f(x,y))^c) = 0) = \\
&= P\Big(b_n^{-1}\max_{1\le l\le h}(f(l))^{-1}\max_{0\le k\le n-l}(S_{k+l}-S_k) \le y, b_n^{-1}\min_{1\le l\le h}(f(l))^{-1}\min_{0\le k\le n-l}(S_{k+l}-S_k) \ge -x\Big)\\
&\to P(N^{(h)}((B_f(x,y))^c) = 0)\\
&= e^{-\mu\left((-\infty,-x/m_{q,h}^-)\cup(y/m_{q,h}^+,\infty)\right)}\\
&= \Phi_\alpha^{\widetilde{q}/\widetilde{p}}(x/m_{q,h}^-)\,\Phi_\alpha(y/m_{q,h}^+)\,,
\end{aligned}$$

where

$$m_{q,h}^+ = \max_{1\le l\le h\wedge q}\ \max_{1\le k\le q-l+1}\frac{(\psi_{k,k+l-1})_+}{f(l)}$$

and

$$m_{q,h}^- = \max_{1\le l\le h\wedge q}\ \max_{1\le k\le q-l+1}\frac{(\psi_{k,k+l-1})_-}{f(l)}\,.$$

Therefore for $x,y>0$,

$$\begin{aligned}
&P\Big(b_n^{-1}\max_{1\le l\le h}(f(l))^{-1}\max_{0\le k\le n-l}(S_{k+l}-S_k) \le y, b_n^{-1}\min_{1\le l\le h}(f(l))^{-1}\min_{0\le k\le n-l}(S_{k+l}-S_k) \le -x\Big)\\
&= P\Big(b_n^{-1}\max_{1\le l\le h}(f(l))^{-1}\max_{0\le k\le n-l}(S_{k+l}-S_k) \le y\Big) -\\
&P\Big(b_n^{-1}\max_{1\le l\le h}(f(l))^{-1}\max_{0\le k\le n-l}(S_{k+l}-S_k) \le y, b_n^{-1}\min_{1\le l\le h}(f(l))^{-1}\min_{0\le k\le n-l}(S_{k+l}-S_k) > -x\Big)\\
&\to \big(1-\Phi_\alpha^{\widetilde{q}/\widetilde{p}}(x/m_{q,h}^-)\big)\Phi_\alpha(y/m_{q,h}^+).
\end{aligned}$$

Finally, in view of Lemma 3.4.7(2), and observing that $m_{q,h}^+ = m_q^+$ and $m_{q,h}^- = m_q^-$ for large h, we derived the limit distribution of the sequence $(b_n^{-1}(r_n^{(\gamma)}, M_n^{(\gamma)}))$.

Proof of Theorem 3.4.4

For $q > 1$ write

$$X_t = \sum_{j=1}^{q} \psi_j Z_{t-j} + \sum_{j=q+1}^{\infty} \psi_j Z_{t-j}.$$

In the analysis of the quantities $\widetilde{M}_n^{(\gamma)}$ and $\widetilde{T}_n^{(\gamma)}$, the first term on the right-hand side can now be controlled by Theorem 3.4.1 when $q \to \infty$ and the second term is handled by the following lemma.

Lemma 3.4.10. *Under the assumptions of Theorem 3.4.4, for any $\delta > 0$ and $f \in \mathcal{F}_\gamma$, $\gamma > \max\{0, 0.5 - \alpha^{-1}\}$,*

$$\lim_{q\to\infty} \limsup_{n\to\infty} P\Big(\max_{1\le l\le n} (f(l))^{-1} \max_{1\le k\le n} \Big| \sum_{t=k+1}^{k+l} \sum_{j=q+1}^{\infty} \psi_j Z_{t-j} \Big| > \delta \Big) = 0.$$

Proof. We have

$$\begin{aligned} &\max_{1\le l\le n} (f(l))^{-1} \max_{1\le k\le n} \Big| \sum_{t=k+1}^{k+l} \sum_{j=q+1}^{\infty} \psi_j Z_{t-j} \Big| \\ &\quad \le \max_{1\le l\le n} (f(l))^{-1} \max_{1\le k\le n} \Big(\sum_{j=q+1}^{n} |\psi_j| \Big| \sum_{t=k+1}^{k+l} Z_{t-j} \Big| + \sum_{j=n+1}^{\infty} |\psi_j| \Big| \sum_{t=k+1}^{k+l} Z_{t-j} \Big| \Big). \end{aligned} \tag{3.4.8}$$

The first maximum expression on the right-hand side of (3.4.8) is bounded by

$$R_n^{(q)} \;=\; \sum_{j=q+1}^{\infty} |\psi_j| \quad \max_{1\le l\le n} (f(l))^{-1} \max_{-n\le k\le n} \Big| \sum_{t=k+1}^{k+l} Z_t \Big| .$$

In view of the results in Mikosch and Račkauskas (2010) for the maximum increment of the iid sequence (Z_t) and since $\sum_{j>q} |\psi_j| \to 0$ as $q \to \infty$ we have

$$\lim_{q\to\infty} \limsup_{n\to\infty} P(R_n^{(q)} > \delta) = 0\,, \quad \delta > 0\,.$$

Next we bound the second maximum term on the right-hand side of (3.4.8). For $\alpha \ge 2$ and if $\mathrm{var}(Z) < \infty$ we can bound this term by

$$\begin{aligned} &\max_{1\le l\le n} (f(l))^{-1} \max_{1\le k\le n} \sum_{j=n+1}^{\infty} |\psi_j| \Big| \sum_{t=k+1}^{k+l} Z_{t-j} \Big| \le \\ &\qquad \le \max_{1\le l\le n} (f(l))^{-1} \max_{1\le k\le n} \sum_{j=n+1}^{\infty} |\psi_j| \Big(\Big| \sum_{t=k+1-j}^{0} Z_t \Big| + \Big| \sum_{t=k+l-j}^{1} Z_t \Big| \Big) \end{aligned}$$

$$\leq c \sum_{j=n+1}^{\infty} |\psi_j| \sqrt{j \log\log j} \sup_{r\geq 3} \frac{1}{\sqrt{r \log\log r}} \Big(\Big| \sum_{t=-r}^{0} Z_t \Big| + \Big| \sum_{t=1}^{r} Z_t \Big| \Big)$$
$$= Q_n \, .$$

By virtue of the law of the iterated logarithm the right-hand supremum is bounded a.s. Therefore and since (3.4.4) holds, $\lim_{n\to\infty} P(Q_n > \delta) = 0$ for $\delta > 0$. This proves the lemma for $\alpha \geq 2$ and $\mathrm{var}(Z) < \infty$.

In the cases $\alpha < 2$ or $\alpha = 2$, $\mathrm{var}(Z) = \infty$, the argument with the law of the iterated logarithm is replaced by a Marcinkiewicz-Zygmund strong law of large numbers (see e.g. Theorem 6.9 in Petrov (1995)) with normalization $n^{1/p}$ for some $p < \alpha$.

□

The proof of Theorem 3.4.4 now follows by some elementary calculations. Indeed, since

$$\Big| \max_{1\leq l\leq n} (f(l))^{-1} \max_{1\leq k\leq n-l} |S_{k+l} - S_k| - \max_{1\leq l\leq n} (f(l))^{-1} \max_{1\leq k\leq n-l} \Big| \sum_{t=k+1}^{k+l} \sum_{j=1}^{q} \psi_j Z_{t-j} \Big| \Big| \leq$$
$$\leq \max_{1\leq l\leq n} (f(l))^{-1} \max_{1\leq k\leq n-l} \Big| \sum_{t=k+1}^{k+l} \sum_{j=q+1}^{\infty} \psi_j Z_{t-j} \Big| \, ,$$

we can combine Lemma 3.4.10 and Theorem 3.4.1, by first letting $n \to \infty$ and then $q \to \infty$, to obtain

$$\lim_{n\to\infty} P(a_n^{-1} \widetilde{M}_n^{(\gamma)} \leq x) \;\; = \;\; \lim_{q\to\infty} \Phi_\alpha(x/m_q) = \Phi_\alpha(x/m_\infty) \, .$$

This concludes the proof of the theorem.

3.5 Random Walks with General Dependent Jump Sizes

In this section we study the maximum increment of a random walk with jump sizes which constitute a general strictly stationary regularly varying sequence (X_t) with index $\alpha > 0$. In contrast to the results in Section 3.4 we do not assume any particular dependence structure of (X_t). As a compensation, we will need some mixing and anti-clustering conditions. These conditions are in general hard to verify for a regularly varying linear process with iid noise. In particular, for such a sequence mixing conditions are in general difficult to check and sometimes not true.

3.5.1 General Theory

In what follows, we write $\mathcal{M}$ for the collection of Radon counting measures on $\overline{\mathbb{R}}^d \setminus \{\mathbf{0}\}$, $\widetilde{\mathcal{M}}$ is the subset of those measures μ in $\mathcal{M}$ for which $\mu(\{x : |x| > 1\}) = 0$ and $\mu(\{x : |x| = 1\}) > 0$. Moreover, $\mathcal{B}(\widetilde{\mathcal{M}})$ is the Borel σ-field of $\widetilde{\mathcal{M}}$. We recall a result from Davis and Mikosch (1998), Theorem 2.8, which is a multivariate version of Theorem 2.7 in Davis and Hsing (1995).

Theorem 3.5.1. *Assume that the sequence (X_t) of $\mathbb{R}^d$-valued random variables is regularly varying with index $\alpha > 0$ and satisfies the mixing condition $\mathcal{A}(b_n)$ of Davis and Hsing (1995), where (b_n) is chosen such that $n\,P(|X_1| > b_n) \to 1$, and the anti-clustering condition*

$$\lim_{h\to\infty} \limsup_{n\to\infty} P\Big(\vee_{h\le|t|\le r_n} |X_t| > \delta b_n \Big| |X_0| > \delta b_n \Big) = 0\,, \quad \delta > 0\,, \tag{3.5.1}$$

holds for a certain sequence $r_n \to \infty$ such that $r_n = o(n)$; see Remark 3.5.2. Then

$$N_n = \sum_{t=1}^{n} \varepsilon_{b_n^{-1} X_t} \xrightarrow{d} N = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i Q_{ij}}\,,$$

where (P_i) are the points of a Poisson process on $(0,\infty)$ with intensity measure given by $\nu(x,\infty) = \theta_{|X|} x^{-\alpha}$ for $x > 0$, $\theta_{|X|} \in [0,1]$ is the extremal index of the sequence $(|X_t|)$ and $\sum_{j=1}^{\infty} \varepsilon_{Q_{ij}}$, $i \ge 1$, constitute an iid sequence of point processes with common distribution Q on $(\widetilde{\mathcal{M}}, \mathcal{B}(\widetilde{\mathcal{M}}))$, and these point processes are independent of (P_i). The distribution Q is given in Theorem 2.8 in Davis and Mikosch (1998).

Remark 3.5.2. Condition $\mathcal{A}(b_n)$ is a rather general mixing condition which is suited for the purposes of extreme value theory for dependent sequences. This condition follows for example from strong mixing of (X_t) with geometric rate. The definition of $\mathcal{A}(b_n)$ involves the block length $r_n \to \infty$ such that $r_n = o(n)$. Under strong mixing with geometric rate, one can choose $r_n = n^\gamma$ for any $\gamma \in (0,1)$. The block length r_n also appears in the anti-clustering condition (3.5.1).
The extremal index of a strictly stationary real-valued sequence is a measure of the extremal clustering in the sequence; see Leadbetter et al. (1983) and Embrechts et al. (1997), Section 8.1.

Next we apply Theorem 3.5.1 to the lagged vector sequence $X_t^{(h)} = (X_t, \ldots, X_{t+h-1})$, $t \in \mathbb{Z}$, of a real-valued strictly stationary sequence (X_t) for increasing $h \ge 1$. Instead of the normalization (b_n) which would depend on the dimension h we choose a sequence (a_n) such

that $n\,P(|X_1| > a_n) \to 1$. By regular variation, we have $b_n/a_n \to c_h$ for certain constants $c_h > 0$.

Corollary 3.5.3. *Assume that (X_t) is regularly varying with index $\alpha > 0$, satisfies the mixing condition $\mathcal{A}(a_n)$ of Davis and Hsing (1995), where the sequence (a_n) is chosen such that $n\,P(|X_1| > a_n) \to 1$, the anti-clustering condition* (3.5.1) *(with (b_n) replaced by (a_n)) and the extremal index $\theta_{|X|}$ is positive. Then*

$$N_n^{(h)} = \sum_{t=1}^{n} \varepsilon_{a_n^{-1} X_t^{(h)}} \xrightarrow{d} N^{(h)} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i^{(h)} Q_{ij}^{(h)}}\,,$$

where the limiting quantities were explained in Theorem 3.5.1. The Poisson points $(P_i^{(h)})$ now have an intensity measure given by $\nu_h(x,\infty) = (\theta_{|X^{(h)}|}/c_h)x^{-\alpha}$, $x > 0$.

Now we are ready to formulate a general result.

Theorem 3.5.4. *Assume that (X_t) is regularly varying with index $\alpha > 0$, satisfies the mixing condition $\mathcal{A}(a_n)$ of Davis and Hsing (1995), where (a_n) is chosen such that $n\,P(|X_1| > a_n) \to 1$, the anti-clustering condition* (3.5.1) *and $\theta_{|X|} > 0$. If the condition*

$$\lim_{h\to\infty} \limsup_{n\to\infty} P(\max_{h\le l\le n} (f(l))^{-1} \max_{0\le k\le n-l} |S_{k+l} - S_k| > \delta a_n) = 0\,, \quad \delta > 0\,, \tag{3.5.2}$$

holds for $f \in \mathcal{F}_\gamma$, some $\gamma > 0$, then

$$\lim_{n\to\infty} P(a_n^{-1} \widetilde{M}_n^{(\gamma)} \le x) = \lim_{h\to\infty} P\left(\sup_{i\ge 1} P_i^{(h)} V_i^{(h)} \le x\right)\,, \quad x > 0\,, \tag{3.5.3}$$

where the iid sequence $(V_i^{(h)})$ is defined in (3.5.5) *and independent of the Poisson points $(P_i^{(h)})$ from Corollary 3.5.3. Moreover, the limit* (3.5.3) *can be written in the form $\Phi_\alpha^\xi(x)$, $x > 0$, for some constant $\xi > 0$.*

Proof. A continuous mapping argument (the map acting on the points $X_t^{(h)}$ is continuous and maps zero into zero) yields that for any non-decreasing sequence of positive numbers $(f(l))$ and $h \ge 1$,

$$\sum_{t=1}^{n} \varepsilon_{a_n^{-1}\left(X_t/f(1),(X_t+X_{t+1})/f(2),\ldots,(X_t+\cdots+X_{t+h-1})/f(h)\right)} \xrightarrow{d} \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \varepsilon_{P_i^{(h)}\left(Q_{ij}^{(h,1)}/f(1),(Q_{ij}^{(h,1)}+Q_{ij}^{(h,2)})/f(2),\ldots,(Q_{ij}^{(h,1)}+\cdots+Q_{ij}^{(h,h)})/f(h)\right)}\,.$$

Here $Q_{ij}^{(h,k)}$ denotes the kth component of the vector $Q_{ij}^{(h)}$. Hence we may conclude that for every $f \in \mathcal{F}_\gamma$, $\gamma \geq 0$, and $h \geq 1$,

$$\lim_{n\to\infty} P(a_n^{-1} \max_{1\leq l\leq h} (f(l))^{-1} \max_{1\leq k\leq n} |S_{k+l} - S_k| \leq x) = \\ = P(\sup_{i\geq 1} P_i^{(h)} \max_{1\leq l\leq h} (f(l))^{-1} \sup_{j\geq 1} |Q_{ij}^{(h,1)} + \cdots + Q_{ij}^{(h,l)}| \leq x)\,, \quad x > 0\,. \tag{3.5.4}$$

The random variables

$$V_i^{(h)} = \max_{1\leq l\leq h} (f(l))^{-1} \sup_{j\geq 1} |Q_{ij}^{(h,1)} + \cdots + Q_{ij}^{(h,l)}|, \quad i = 1, 2, \ldots, \tag{3.5.5}$$

constitute an iid sequence independent of the Poisson points $(P_i^{(h)})$. It is well known (see e.g. de Haan and Ferreira (2006), Corollary 9.4.5) that $Y_h = \sup_{i\geq 1} P_i^{(h)} V_i^{(h)}$ has a Fréchet distribution $\Phi_\alpha^{\xi_h}$ with shape parameter $\alpha > 0$ and scale factor $\xi_h = \mathbb{E}\,[(V^{(h)})^\alpha]\theta_{|X^{(h)}|}/c_h$.

In view of (3.5.4) we have for every $h \geq 1$,

$$Y_{nh} = a_n^{-1} \max_{1\leq l\leq h} (f(l))^{-1} \max_{1\leq k\leq n} |S_{k+l} - S_k| \xrightarrow{d} Y_h = \sup_{i\geq 1} P_i^{(h)} V_i^{(h)}$$

and the limit has a Fréchet distribution $\Phi_\alpha^{\xi_h}$. By virtue of (3.5.2) and Theorem 2 in Dehling et al. (2009) we conclude that $Y_h \xrightarrow{d} Y$ for a random variable Y and $a_n^{-1}\widetilde{M}_n^{(\gamma)} \xrightarrow{d} Y$. Since $\lim_{h\to\infty} \Phi_\alpha^{\xi_h}$ exists, the limit $\xi = \lim_{h\to\infty} \xi_h$ exists as well and Φ_α^ξ is the distribution of Y. Moreover, ξ is positive since $a_n^{-1} \max_{t=1,\ldots,n} |X_t| \leq a_n^{-1} M_n^{(\gamma)}$ and the sequence on the left-hand side has the limiting distribution $\Phi_\alpha^{\theta_{|X|}}$ with $\theta_{|X|}$ positive by assumption; see Davis and Hsing (1995).

□

3.5.2 The Increment Process is a Process with Multiplicative Noise

In this section we assume that the strictly stationary real-valued process (X_t) has the form

$$X_t = \sigma_t\, Z_t\,, \quad t \in \mathbb{Z}\,, \tag{3.5.6}$$

where (Z_t) is an iid sequence and (σ_t) is a strictly stationary sequence of non-negative random variables such that σ_t and Z_t are independent for every t. We mention two popular specifications of the volatility process (σ_t).

3.5.2.1 The GARCH Process

The perhaps best known process of type (3.5.6) is a GARCH process introduced in Bollerslev (1986).

Definition 3.5.5 (GARCH Process). Let $\alpha_i, \beta_j \in [0, \infty)$ and $p, q \in \mathbb{N}$ such that $\alpha_p \beta_q \neq 0$. Then the process (X_t) given by (3.5.6) with a volatility process (σ_t) satisfying the recursion

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 , \quad t \in \mathbb{Z} . \tag{3.5.7}$$

is called a *GARCH(p,q)* process.

In Basrak et al. (2002) one finds general conditions for the existence of a strictly stationary version of the process $X_t = \sigma_t Z_t$, i.e. one has to ensure that the recursion (3.5.7) has a strictly stationary causal solution. These conditions require that $\alpha_0 > 0$, $\mathbb{E} \log^+ |Z_1| < \infty$ and negativity of the Lyapunov exponent of a certain matrix whose entries depend on the Z_t^2's and the parameters α_i and β_j. See Basrak et al. (2002), Theorem 3.1(A), for details and for some simple sufficient conditions for the Lyapunov exponent to be negative. In the same result a sufficient condition for regular variation of (X_t) is given: Z_1 has a positive density on $\mathbb{R}$ and there exists $0 < h_0 \leq \infty$ such that $\mathbb{E} |Z_1|^h < \infty$ for $h < h_0$ and $\mathbb{E} |Z_1|^{h_0} = \infty$. (If $h_0 = \infty$ the latter relation has the interpretation that $\mathbb{E} |Z_1|^h \to \infty$ as $h \to \infty$.) It follows from Theorem 3.1(B) in Basrak et al. (2002) that the sequence $((X_t^2, \sigma_t^2))$ is regularly varying with some index $\alpha/2 > 0$. In Basrak et al. (2002) the result was proved under the additional condition that $\alpha/2$ is not an even integer. Following the recent approach in Boman and Lindskog (2009), this additional condition can be dropped. Then one may follow the proof of Corollary 3.5 in Basrak et al. (2002) to see that, under the conditions above, $((X_t, \sigma_t))_{t \in \mathbb{Z}}$ is a regularly varying sequence with index α. We mention that in the GARCH(1, 1) case with $\mathbb{E} Z_1 = 0$ and $\mathbb{E} Z_1^2 = 1$, α is the unique positive solution to the equation $\mathbb{E} (\alpha_1 Z_1^2 + \beta_1)^{\alpha/2} = 1$. The regular variation of the GARCH process is a consequence of Kesten (1973) on the tail of solutions to stochastic recurrence equations; see also Goldie (1991). Finally, Corollary 3.5 in Basrak et al. (2002) also gives strong mixing of $((X_t, \sigma_t))_{t \in \mathbb{Z}}$ with geometric rate provided Z_t is symmetric and has a density on $\mathbb{R}$. The latter condition is satisfied in most applications where Z_t is assumed standard normal or t-distributed with variance 1. It was mentioned in Remark 3.5.2 that regular variation and strong mixing with geometric rate for the sequence (X_t) imply the mixing condition $\mathcal{A}(a_n)$. Moreover, the anti-clustering condition (3.5.1) is also satisfied as proved for Theorem 2.10 in Basrak et al. (2002). We summarize as follows.

Theorem 3.5.6. *Let (X_t) be a strictly stationary* GARCH(p, q) *process with an iid symmetric noise sequence (Z_t) such that $\mathbb{E} Z_1^2 = 1$. Further assume that Z_1 has a positive density on $\mathbb{R}$ and there exists $h_0 \leq \infty$ such that $\mathbb{E} |Z_1|^h < \infty$ for $h < h_0$ and $\mathbb{E} |Z_1|^{h_0} = \infty$.*

Then (X_t) is regularly varying with some index $\alpha > 0$ and the statement of Theorem 3.5.4 holds for $\gamma > \max(0, 0.5 - \alpha^{-1})$.

Proof. In view of the discussion preceding the theorem, all conditions of Theorem 3.5.4 but (3.5.2) have been verified. Thus we focus on the latter condition. We also notice that the condition $f \in \mathcal{F}_\gamma$ for some $\gamma > \max(0, 0.5 - \alpha^{-1})$ has not been used so far. It will be crucial in the sequel. For notational simplicity we restrict ourselves to the functions $f(l) = l^\gamma$.

We introduce the truncated random variables $X'_t = \sigma_t \mathbb{1}_{\{\sigma_t \le h^\gamma a_n\}} Z_t$ (we suppress the dependence on n and h in the notation) and the corresponding partial sums

$$S'_k = \sum_{t=1}^{k} X'_t\,, \quad k \ge 1\,, \qquad S'_0 = 0\,.$$

Then for any $\delta > 0$,

$$\begin{aligned} P(\max_{h\le l\le n} l^{-\gamma} \max_{0\le k\le n-l} |S_{k+l} - S_k| > \delta a_n) \le \\ \le P(\max_{k\le n} \sigma_k > h^\gamma a_n) + P(\max_{h\le l\le n} l^{-\gamma} \max_{0\le k\le n-l} |S'_{k+l} - S'_k| > \delta a_n) \\ \le P(\max_{k\le n} \sigma_k > h^\gamma a_n) + 2 \sum_{j=1}^{\log_2(n/h)} 2^j T_j\,, \end{aligned} \tag{3.5.8}$$

where $\log_2$ denotes logarithm with base 2 and

$$T_j = P(\max_{1\le k\le 2n2^{-j}} |S'_k| > \delta (n2^{-j})^\gamma a_n)\,. \tag{3.5.9}$$

In the last step we used Lemma 3.3 in Mikosch and Račkauskas (2010). There the proof was given for an iid sequence (X_t) but the proof remains the same for a strictly stationary sequence.

It is well known that $a_n^{-1} \max_{k\le n} \sigma_k$ converges in distribution to a Fréchet $\Phi_\alpha^{\theta_\sigma}$ distributed random variable, where $\theta_\sigma > 0$ is the extremal index of the sequence (σ_t). This fact follows e.g. from the point process convergence results in Basrak et al. (2002). Hence

$$\lim_{h\to\infty} \lim_{n\to\infty} P(a_n^{-1} \max_{k\le n} \sigma_k > h^\gamma) = 0\,. \tag{3.5.10}$$

Since Z_t is assumed symmetric an application of Lévy's maximal inequality (e.g. Petrov (1995), Theorem 2.2) conditional on $(\sigma_t | Z_t|)$ yields

$$T_j \le 2P(|S'_{2N}| > \delta N^\gamma a_n)\,, \quad \text{where } N = [n2^{-j}]. \tag{3.5.11}$$

Assume $\alpha < 2$. Then by Markov's inequality and Karamata's theorem (see Bingham et al. (1987)),

$$\begin{aligned} T_j &\le c\, N^{1-2\gamma} a_n^{-2} \mathbb{E}\,[Z_1^2 \sigma_1^2 \mathbb{1}_{\{\sigma_1 \le h^\gamma a_n\}}] \\ &\sim c\, h^{2\gamma} N^{1-2\gamma} P(\sigma_1 > h^\gamma a_n) \\ &\le c h^{\gamma(2-\alpha)} n^{-1} N^{1-2\gamma}. \end{aligned}$$

Hence

$$\sum_{j=1}^{\log_2(n/h)} 2^j T_j \le c\, n^{-2\gamma} h^{\gamma(2-\alpha)} \sum_{j=1}^{\log_2(n/h)} 2^{2\gamma j} \le c\, h^{-\gamma\alpha}\,. \tag{3.5.12}$$

In sum, we conclude from (3.5.8)–(3.5.12) that (3.5.2) holds for every $\delta > 0$, $f(l) = l^\gamma$ and $\gamma > 0$. Now assume $\alpha \ge 2$. It follows from the assumptions that $\mathbb{E}\,|Z_1|^p < \infty$ for some $p > \alpha$. Then Markov's and Burkholder's inequalities and Karamata's theorem yield

$$\begin{aligned} P(|S'_{2N}| > \delta N^\gamma a_n) &\le c\,(N^\gamma a_n)^{-p} \mathbb{E}\,|S'_{2N}|^p \\ &\le c\,(N^\gamma a_n)^{-p} N^{p/2} \mathbb{E}\,\sigma_1^p \mathbb{1}_{\{\sigma_1 \le h^\gamma a_n\}} \\ &\sim c N^{p(0.5-\gamma)} h^{\gamma p} P(\sigma_1 > h^\gamma a_n) \\ &\sim c N^{p(0.5-\gamma)} h^{\gamma(p-\alpha)} n^{-1}\,. \end{aligned} \tag{3.5.13}$$

Since $\gamma > 0.5 - \alpha^{-1}$,

$$\sum_{j=1}^{\log_2(n/h)} 2^j T_j \le c h^{\gamma(p-\alpha)} n^{-1+p(0.5-\gamma)} \sum_{j=1}^{\log_2(n/h)} 2^{j(1-p(0.5-\gamma))} \le c\, h^{-\alpha\gamma-1+0.5p}$$

Now choose $p > \alpha$ so close to α that $-\alpha\gamma - 1 + 0.5p < 0$. Then (3.5.2) follows.

□

3.5.2.2 The stochastic volatility model

Another model of the type (3.5.6) has attracted some attention in the financial time series literature: the stochastic volatility model.

Definition 3.5.7 (Stochastic Volatility Model). Let (σ_t) be a strictly stationary sequence independent of the iid sequence (Z_t). Then the process (X_t) given by (3.5.6) is called a *stochastic volatility process*.

Remark 3.5.8. There is a number of different stochastic volatility models in the literature. They differ in particular by the chosen volatility process σ_t. See also Section 4.5 for a multivariate stochastic volatility model where the volatility process Σ_t is modeled by a positive semi-definite supOU process.

If Z is regularly varying with index $\alpha > 0$ and $\mathbb{E}\,\sigma^p < \infty$ for some $p > \alpha$ then (X_t) is regularly varying with index α. In this case, the limiting measures of the regularly varying sequence, i.e. the measures appearing as vague limits of $nP(a_n^{-1}(X_1, \ldots, X_h) \in \cdot)$, are concentrated on the axes; see Davis and Mikosch (2001). For this reason, the stochastic volatility model (X_t) has very much the same extremal behavior as an iid sequence with the marginal distribution as X_1. In particular, the extremal index of (X_t) is one and the following point process convergence result holds: if the tail balance condition (3.1.3) holds and $(\log \sigma_t)$ is a linear Gaussian process, then

$$\sum_{t=1}^{n} \varepsilon_{a_n^{-1}(X_t,\ldots,X_{t+h-1})} \xrightarrow{d} \sum_{k=1}^{h} \sum_{i=1}^{\infty} \varepsilon_{J_i \mathbf{e}_k}\,, \tag{3.5.14}$$

where $\mathbf{e}_k$ is the kth unit vector in $\mathbb{R}^h$, (J_i) are the points of a Poisson process on $\overline{\mathbb{R}}_0$ with intensity $\alpha\,|x|^{-\alpha-1}[\widetilde{p}\,\mathbb{1}_{(0,\infty)}(x) + \widetilde{q}\,\mathbb{1}_{(-\infty,0)}(x)]\,dx$ and (a_n) is chosen such that $n\,P(|X_1| > a_n) \to 1$. In the case of an iid regularly varying sequence (X_t) with tail balance condition (3.1.3) (with Z replaced by X) relation (3.5.14) holds as well.

Now we are ready to formulate the following analog of Theorem 3.5.6.

Theorem 3.5.9. *Let (X_t) be a stochastic volatility process. Assume that Z_1 is symmetric and has a regularly varying distribution with index α and $(\log \sigma_t)$ is a Gaussian linear process. Then for $f \in \mathcal{F}_\gamma$, $\gamma > \max(0, 0.5 - \alpha^{-1})$,*

$$\lim_{n\to\infty} P(a_n^{-1} \max_{1\le l\le n} (f(l))^{-1} \max_{1\le k\le n} |S_{k+l} - S_k| \le x) = \Phi_\alpha(x)\,, \quad x > 0\,.$$

Proof. The point process convergence (3.5.14) was the starting point in Mikosch and Račkauskas (2010) for proving that for any $f \in \mathcal{F}_\gamma$ and $\gamma > 0$,

$$\lim_{n\to\infty} P(a_n^{-1} \max_{1\le l\le h} (f(l))^{-1} \max_{1\le k\le n} |S_{k+l} - S_k| \le x) = \Phi_\alpha(x)\,, \quad x > 0\,.$$

The same arguments apply in the stochastic volatility case, i.e. the relation above remains valid. Now consider the truncated random variables $X_t' = \sigma_t Z_t \mathbb{1}_{\{|Z_t| \le h^\gamma a_n\}}$ and denote their partial sums by S_k', $k \ge 0$. Then the following analog of (3.5.8) holds:

$$\begin{aligned}
P(\max_{h\le l\le n} l^{-\gamma} \max_{0\le k\le n-l} |S_{k+l} - S_k| > \delta a_n) &\le \\
&\le P(\max_{k\le n} |Z_k| > h^\gamma a_n) + P(\max_{h\le l\le n} l^{-\gamma} \max_{0\le k\le n-l} |S_{k+l}' - S_k'| > \delta a_n) \\
&\le P(\max_{k\le n} |Z_k| > h^\gamma a_n) + 2 \sum_{j=1}^{\log_2(n/h)} 2^j T_j\,,
\end{aligned}$$

where T_j is defined in (3.5.9). Since $a_n^{-1} \max_{k=1,\ldots,n} |Z_t|$ converges in distribution to a Fréchet distributed random variable,

$$\lim_{h\to\infty} \limsup_{n\to\infty} P(\max_{k\le n} |Z_k| \ge h^\gamma a_n) = 0\,.$$

Now assume that Z_1 is also symmetric. Then, by Lévy's maximal inequality conditional on (σ_t), relation (3.5.11) follows. We consider the case $\alpha < 2$ first. An application of Markov's inequality and the conditional independence of the X_t's yield

$$\begin{aligned} T_j &\le c\, N^{1-2\gamma} a_n^{-2} \mathbb{E}\,[\sigma_1^2 Z_1^2 \mathbb{1}_{\{|Z_1|\le h^\gamma a_n\}}] \\ &\sim c\, h^{2\gamma} N^{1-2\gamma} P(|Z_1| > h^\gamma a_n) \\ &\le c h^{\gamma(2-\alpha)} n^{-1} N^{1-2\gamma}, \end{aligned}$$

Now one can follow the lines of the proof of Theorem 3.5.6. Next we consider the case $\alpha \ge 2$. Then for $p > \alpha$ the Markov and Marzinkiewicz-Zygmund inequalities conditional on (σ_t) and the Minkowski inequality yield

$$\begin{aligned} \mathbb{E}\,[P(|S'_{2N}| > \delta N^\gamma a_n \mid (\sigma_t))] &\le c\, a_n^{-p} N^{-p(\gamma-0.5)} \mathbb{E}\,|Z_1|^p \mathbb{1}_{\{|Z_1|\le h^\gamma a_n\}} \mathbb{E}\left(N^{-1}\sum_{t=1}^{2N} \sigma_t^2\right)^{p/2} \\ &\le c\, a_n^{-p} N^{-p(\gamma-0.5)} \mathbb{E}\,|Z_1|^p \mathbb{1}_{\{|Z_1|\le h^\gamma a_n\}} \mathbb{E}\,\sigma_1^p\,. \end{aligned}$$

Now we may follow the lines of the proof of Theorem 3.5.6 for $\alpha \ge 2$.

□

Chapter 4

Finite Dimensional Regular Variation of Multivariate Lévy-Driven Mixed Moving Average Processes and SupOU Stochastic Volatility Models[1]

The class of multivariate Lévy-driven mixed moving average (MMA) processes of the type $X_t = \int\int f(A, t-s)\Lambda(dA, ds)$ covers a wide range of well-known and extensively used processes such as Ornstein-Uhlenbeck processes, superpositions of Ornstein-Uhlenbeck (supOU) processes, (fractionally integrated) CARMA processes and increments of fractional Lévy processes. In this chapter, we introduce multivariate MMA processes and give conditions for their existence and finite dimensional regular variation of the stationary distributions. Furthermore, we study the tail behavior of multivariate supOU processes and of a stochastic volatility model where a positive semidefinite supOU process models the stochastic volatility.

4.1 Introduction

In many areas of application Lévy-driven processes are used for modeling time series. One elementary example of the processes used is the Lévy-driven Ornstein-Uhlenbeck (OU) type process

$$X_t = \int_{-\infty}^{t} e^{-a(t-s)} dL_s,$$

[1]The content of this chapter is based on Moser and Stelzer (2011).

where L is a Lévy process (see Section 2.3 and Sato (2002) for detailed introductions). These processes are used, for instance, to model the variance (i.e., the volatility in the terminology of mathematical finance) in the OU type stochastic volatility model of Barndorff-Nielsen and Shephard (2001) which has been extended to the multivariate setting by Pigorsch and Stelzer (2009). Even though this model has many nice properties (e.g. stochastic volatility with jumps and clustering, heavy tails etc.) it does not account for the long memory effects that can often be found in real data. This problem can be bypassed by the superposition of OU type processes which leads to supOU processes of the type

$$X_t = \int_{\mathbb{R}} \int_{-\infty}^{t} e^{a(t-s)} \Lambda(da, ds),$$

where Λ is a so-called Lévy basis. These processes have been introduced by Barndorff-Nielsen (2001), extended to a multivariate setting by Barndorff-Nielsen and Stelzer (2011a) and they are used in the multivariate supOU type stochastic volatility model of Barndorff-Nielsen and Stelzer (2011b). Bayesian estimation of univariate supOU stochastic volatility models is e.g. carried out in Griffin and Steel (2010).
The aim of this chapter is to analyze the tail behavior of the multivariate mixed moving average (MMA) processes

$$X_t = \int_{M_d^-} \int_{\mathbb{R}} f(A, t-s) \Lambda(dA, ds)$$

that allow for a general kernel function $f : M_d^- \times \mathbb{R} \mapsto M_{n,d}$ (Λ is an $\mathbb{R}^d$-valued Lévy basis in this setting). They reach back to Surgailis et al. (1993) and they cover both, OU and supOU processes, as well as CARMA processes, fractionally integrated CARMA processes (cf. Brockwell (2004), Marquardt (2007)) and increments of fractional Lévy processes (cf. Marquardt (2006), Bender et al. (2011) and references therein). The tail behavior of univariate MMA processes has already been studied by Fasen (2005) and Jacobsen et al. (2009) and we extend the results to a multivariate setting and analyze also the special case of supOU processes and the related stochastic volatility model given by

$$\begin{aligned} dX_t &= a_t dt + \Sigma_t^{1/2} dW_t + \Psi(dL_t) \\ X_0 &= 0, \end{aligned}$$

where a is an $\mathbb{R}^d$-valued predictable process, W is the standard d-dimensional Brownian motion, L is the Lévy process associated with Λ, $\Psi : \mathbb{S}_d \mapsto \mathbb{R}^d$ is a linear operator and

the stochastic volatility process $(\Sigma_t)_{t\in\mathbb{R}}$ is a matrix-valued positive semidefinite supOU process. The multivariate extension is non-trivial, since the definition of regular variation is considerably more involved in the multivariate setting and we have to take the peculiarities created by the use of matrices into account.
In finance understanding the tail behavior is of great importance for risk assessment and risk management. Moreover, the results of this chapter allow one to understand how one can model the so-called "correlation breakdown" effect (viz. in times of extreme crisis basically all correlations get close to one) which is regarded by econometrics to be typically present in observed financial data.
The chapter is structured as follows. We start by giving some general notation in Section 4.2.1. In Section 4.2.2 we will give a short excursion to multivariate regular variation that we need when we analyze the tail behavior of the processes given. An introduction to Lévy bases and conditions for the existence of integrals with respect to Lévy bases will be given in Section 4.2.3. Based on these preliminaries, we can then define and analyze multivariate mixed moving average processes in Section 4.3. We give sufficient conditions for the mixed moving average processes to be regularly varying given that the driving Lévy basis is regularly varying. Furthermore, we examine the restrictiveness of the conditions by establishing closely related necessary conditions. In Section 4.4 we apply these results to multivariate supOU processes and give some more accessible conditions for this special case. Moreover, we consider a stochastic volatility model that is based on positive semidefinite supOU processes and analyze its tail behavior in Section 4.5, which is very important for risk assessment. Finally, we discuss the model's relevance to finance in Section 4.5.2.

4.2 Preliminaries

4.2.1 Notation

Given the real numbers $\mathbb{R}$ we use the notation $\mathbb{R}^+$ for the positive real numbers and $\mathbb{R}^-$ for the negative real numbers, both without 0. $\mathbb{N}$ is the set of positive integers. The Borel sets are denoted by $\mathcal{B}$, where $\mathcal{B}_b$ are the bounded Borel sets and $\mathcal{B}_\mu := \{B \in \mathcal{B} : \mu(\partial B) = 0\}$ describes all Borel sets with no μ-mass at the boundary ∂B. The closure of a set B is given by $\overline{B}$. S is the unit sphere, λ is the Lebesgue measure on $\mathbb{R}$ and $N(0, I_d)$ is the standard normal distribution in $\mathbb{R}^d$.

For matrices, $M_{n,d}$ is the set of all $n \times d$ matrices and M_d the set of all $d \times d$ matrices. M_d^- is the set of all $d \times d$ matrices with eigenvalues having strictly negative real part. I_d is the $d \times d$ identity matrix, $\mathbb{S}_d$ denotes the symmetric $d \times d$ matrices and $\mathbb{S}_d^+$ the positive

semidefinite $d \times d$ matrices. We write A^T for the transposed of a matrix A and $\|A\|$ for its matrix norm. Since all norms are equivalent, the type of norm is not important for the results, but if we make no further specifications, we use the operator norm induced by the Euclidean norm. $j(A) := \min_{\|x\|=1} \|Ax\|$ is the modulus of injectivity of A. vec(A) is the well-known operation that creates a vector by stacking the columns of an $n \times n$ matrix A below each other to obtain an $\mathbb{R}^{n^2}$-valued vector and $\otimes$ is the tensor product of two matrices.
Vague convergence is denoted by $\xrightarrow{v}$. It is defined on the one-point uncompactification $\overline{\mathbb{R}}^d \backslash \{0\}$, which assures that the sets $B \subseteq V_r := \{x : \|x\| > r\}$, $r > 0$, that are bounded away from the origin can be referred to as the relatively compact sets in the vague topology. In this topology, the compact sets shall be denoted by $\mathcal{K}$ and the open sets by $\mathcal{G}$.

4.2.2 Multivariate Regular Variation

For the analysis of the tail behavior of multivariate stochastic processes, we use the well established concept of regular variation. However, there is not only one single definition of multivariate regular variation, but many different equivalent ones (cf. Section 2.1.3). For detailed and very good introductions into the different approaches to multivariate regular variation, we refer the reader also to Resnick (2007) and Lindskog (2004). Throughout this chapter, we use the following well-known definition of multivariate regular variation (cf. Resnick (1986) and Hult and Lindskog (2006a)).

Definition 4.2.1 (Multivariate Regular Variation). A random vector X with values in $\mathbb{R}^d$ is called *regularly varying* with index $\alpha > 0$ if there exists a slowly varying function $l : \mathbb{R} \mapsto \mathbb{R}$ and a nonzero Radon measure μ defined on $\mathcal{B}(\overline{\mathbb{R}}^d \backslash \{0\})$ with $\mu(\overline{\mathbb{R}}^d \backslash \mathbb{R}^d) = 0$ such that, as $u \to \infty$,

$$u^\alpha l(u) P(u^{-1} X \in \cdot) \xrightarrow{v} \mu(\cdot)$$

on $\mathcal{B}(\overline{\mathbb{R}}^d \backslash \{0\})$. We write $X \in RV(\alpha, l, \mu)$.
Similarly, we call a Radon measure ν regularly varying if α, l and μ exist as above with

$$u^\alpha l(u) \nu(u\cdot) \xrightarrow{v} \mu(\cdot)$$

for $u \to \infty$ and we write $\nu \in RV(\alpha, l, \mu)$.
A stochastic process $(X_t)_{t \in \mathbb{R}} \in \mathbb{R}^d$ is called regularly varying with index α if all its finite dimensional distributions are regularly varying with index α.

Remark 4.2.2. Note that in this chapter we consider regular variation of processes only in terms of (multivariate) regular variation of their finite dimensional distributions. This

is not to be confused with the infinite dimensional regular variation theory of Section 2.1.4 that we will analyze with respect to MMA processes in Chapter 5. Both concepts are linked by the necessary and sufficient conditions of Theorem 2.1.22.

The measure μ is homogeneous, i.e. it necessarily satisfies the condition

$$\mu(tB) = t^{-\alpha}\mu(B)$$

for all $B \in \mathcal{B}(\overline{\mathbb{R}}^d \backslash \{0\})$ and $t > 0$. We make use of this property throughout this chapter. In this chapter, we will deal with infinitely divisible random variables and processes. For those, the following very useful connection between regular variation of the random variable and its Lévy measure exists.

Theorem 4.2.3 (Hult and Lindskog (2006a), Proposition 3.1). *Let $X \in \mathbb{R}^d$ be an infinitely divisible random vector with Lévy measure ν. Then $X \in RV(\alpha, l, \mu)$ if and only if $\nu \in RV(\alpha, l, \mu)$.*

Furthermore, we will also need regular variation of matrix-valued random variables and processes. If we take into account the well-known vec operation that creates a vector by stacking the columns of a matrix below each other, we can simply apply the above definition. This allows us to use all known results for the $\mathbb{R}^d$-valued case also in the matrix-valued case.

4.2.3 Lévy Bases and Integration

In this section we recall $\mathbb{R}^d$-valued Lévy bases which are generalizations of Lévy processes, and the related integration theory. For a general introduction to Lévy processes and infinitely divisible distributions see Sato (2002). Lévy bases are also called *infinitely divisible independently scattered random measures* (i.d.i.s.r.m.) in the literature. For more details on Lévy bases see Rajput and Rosiński (1989) and Pedersen (2003).

Definition 4.2.4 (Lévy Basis). An $\mathbb{R}^d$-valued random measure $\Lambda = (\Lambda(B))$ with $B \in \mathcal{B}_b(M_d^- \times \mathbb{R})$ is called a *Lévy basis* if:

- the distribution of $\Lambda(B)$ is infinitely divisible for all $B \in \mathcal{B}_b(M_d^- \times \mathbb{R})$.
- for any $n \in \mathbb{N}$ the random variables $\Lambda(B_1), \ldots, \Lambda(B_n)$ are independent for pairwise disjoint sets $B_1, \ldots, B_n \in \mathcal{B}_b(M_d^- \times \mathbb{R})$.

- for any pairwise disjoint sets $(B_i)_{i\in\mathbb{N}} \in \mathcal{B}_b(M_d^- \times \mathbb{R})$ with $\bigcup_{n\in\mathbb{N}} B_n \in \mathcal{B}_b(M_d^- \times \mathbb{R})$ we have $\Lambda(\bigcup_{n\in\mathbb{N}} B_n) = \sum_{n\in\mathbb{N}} \Lambda(B_n)$ almost surely.

In this thesis we consider only time-homogeneous and factorisable Lévy bases, i.e. Lévy bases with characteristic function

$$\mathbb{E}\left(e^{iu^T\Lambda(B)}\right) = e^{\varphi(u)\Pi(B)} \tag{4.2.1}$$

for all $u \in \mathbb{R}^d$ and $B \in \mathcal{B}_b(M_d^- \times \mathbb{R})$, where $\Pi = \lambda \times \pi$ is the product of a probability measure π on $M_d^-(\mathbb{R})$ and the Lebesgue measure λ on $\mathbb{R}$ and

$$\varphi(u) = iu^T\gamma - \frac{1}{2}u^T\Sigma u + \int_{\mathbb{R}^d} \left(e^{iu^Tx} - 1 - iu^Tx\mathbb{1}_{[-1,1]}(\|x\|)\right)\nu(dx)$$

is the cumulant transform of an infinitely divisible distribution with characteristic triplet (γ, Σ, ν). By L we denote the *underlying Lévy process* associated with (γ, Σ, ν) and given by $L_t = \Lambda(M_d^- \times (0, t])$ and $L_{-t} = \Lambda(M_d^- \times [-t, 0))$ for $t \in \mathbb{R}^+$. The quadruple $(\gamma, \Sigma, \nu, \pi)$ determines the distribution of the Lévy basis completely and therefore it is called the *generating quadruple*. A definition of $\mathbb{S}_d$-valued Lévy bases follows along the same lines.

Remark 4.2.5. Considering only time homogeneous and factorisable Lévy bases is motivated by possible applications where models with too many parameters are of no real help, and the so far developed theory of special cases, particularly the supOU process, where this assumption is also made. However, it should be noted that this assumption is not overly restrictive, because stationarity of a Lévy-driven MMA requires obviously in general a time-homogeneous Lévy basis, i.e. the Lebesgue measure has to be used on the time axis. In this work it appears very natural only to consider stationary cases. Hence, the only possible generalization would be to allow the infinitely divisible distribution to depend on $A \in M_d^-$. We could have $\varphi(A, u)$ instead of $\varphi(u)$. Then we would also have a characteristic triplet $(\gamma(A), \Sigma(A), \nu(A, dx))$ with ν being a "Lévy kernel" and all the results would have immediate extensions to this case noting that as far as regular variation is concerned one would have to demand that $\Lambda(B)$ has to be regularly varying for all sets B with same index α and slowly varying function l (or "degenerately α-regularly varying", i.e. $u^\alpha l(u)P(u^{-1}\Lambda(B \in \cdot)) \to 0$) and the measure of regular variation would have to be given via a nontrivial kernel $\mu_\nu(A, \cdot)$. Like the univariate literature (see Fasen (2005), Fasen (2009) and Fasen and Klüppelberg (2007)) we refrain from stating the results on this level of generality, since it would not add real insight, but lead to overly technical statements not relevant for applications.

The main focus of this chapter, the mixed moving average processes, are defined by integrating over a function f with respect to a Lévy basis. Regarding the existence of these integrals we recall the following multivariate extension of Theorem 2.7 in Rajput and Rosiński (1989).

Theorem 4.2.6. *Let Λ be an $\mathbb{R}^d$-valued Lévy basis with characteristic function of the form (4.2.1) and let $f : M_d^- \times \mathbb{R} \mapsto M_{n,d}$ be a measurable function. Then f is Λ-integrable as a limit in probability in the sense of Rajput and Rosiński (1989) if and only if*

$$\int_{M_d^-} \int_{\mathbb{R}} \left\| f(A,s)\gamma + \int_{\mathbb{R}^d} f(A,s)x \left(\mathbb{1}_{[0,1]}(\|f(A,s)x\|) - \mathbb{1}_{[0,1]}(\|x\|) \right) \nu(dx) \right\| ds\pi(dA) < \infty, \quad (4.2.2)$$

$$\int_{M_d^-} \int_{\mathbb{R}} \|f(A,s)\Sigma f(A,s)^T\| ds\pi(dA) < \infty \quad \text{and} \quad (4.2.3)$$

$$\int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \left(1 \wedge \|f(A,s)x\|^2\right) \nu(dx) ds\pi(dA) < \infty. \quad (4.2.4)$$

If f is Λ-integrable, the distribution of $X_0 = \int_{M_d^-} \int_{\mathbb{R}^+} f(A,s)\Lambda(dA,ds)$ is infinitely divisible with characteristic triplet $(\gamma_{int}, \Sigma_{int}, \nu_{int})$ given by

$$\gamma_{int} = \int_{M_d^-} \int_{\mathbb{R}} \left(f(A,s)\gamma + \int_{\mathbb{R}^d} f(A,s)x \left(\mathbb{1}_{[0,1]}(\|f(A,s)x\|) - \mathbb{1}_{[0,1]}(\|x\|) \right) \nu(dx) \right) ds\pi(dA),$$

$$\Sigma_{int} = \int_{M_d^-} \int_{\mathbb{R}} f(A,s)\Sigma f(A,s)^T ds\pi(dA) \quad \text{and}$$

$$\nu_{int}(B) = \int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \mathbb{1}_B(f(A,s)x)\nu(dx)ds\pi(dA) \text{ for all Borel sets } B \subseteq \mathbb{R}^d.$$

We give now some more accessible sufficient conditions for the special case of a regular varying driving Lévy measure ν. Therefore, we define the set

$$\mathbb{L}^\delta(\lambda \times \pi) := \left\{ f : M_d^- \times \mathbb{R} \mapsto M_{n,d} \text{ measurable}, \int_{M_d^-} \int_{\mathbb{R}} \|f(A,s)\|^\delta ds\pi(dA) < \infty \right\}.$$

The following theorem is a multivariate analogue of Proposition 3.1 in Fasen (2005), which is non-trivial due to the peculiarities arising from the used matrices.

Theorem 4.2.7. *Let Λ be a Lévy basis with values in $\mathbb{R}^d$ and characteristic quadruple $(\gamma, \Sigma, \nu, \pi)$, let ν be regularly varying with index α and let $f : M_d^- \times \mathbb{R} \mapsto M_{n,d}$. Then f*

is Λ-integrable in the sense of Rajput and Rosiński (1989) and X_0 is well defined and infinitely divisible with the characteristic triplet given in Theorem 4.2.6 if one of the following conditions hold:

(i) L_1 is α-stable with $\alpha \in (0,2)\backslash\{1\}$ and $f \in \mathbb{L}^\alpha \cap \mathbb{L}^1$.

(ii) f is bounded and $f \in \mathbb{L}^\delta$ for some $\delta < \alpha$, $\delta \leq 1$.

(iii) f is bounded, $\mathbb{E}\, L_1 = 0$, $\alpha > 1$ and $f \in \mathbb{L}^\delta$ for some $\delta < \alpha$, $\delta \leq 2$.

Proof. We will prove the result by validating the conditions (4.2.2), (4.2.3) and (4.2.4) given in Theorem 4.2.6 in each of the three settings.

(i). From Theorem 14.3 of Sato (2002) we know that in the α-stable case $\Sigma = 0$, which makes condition (4.2.3) trivial. Furthermore, there is a finite measure θ on the unit sphere S such that

$$\nu(B) = \int_S \int_0^\infty \frac{\mathbb{1}_B(r\xi)}{r^{1+\alpha}} dr\theta(d\xi) \quad \text{for } B \in \mathcal{B}^d.$$

For condition (4.2.2), this yields

$$\begin{aligned}
&\int_{M_d^-} \int_{\mathbb{R}} \left\| f(A,s)\gamma + \int_{\mathbb{R}^d} f(A,s)x \left(\mathbb{1}_{[0,1]}(\|f(A,s)x\|) - \mathbb{1}_{[0,1]}(\|x\|) \right) \nu(dx) \right\| ds\pi(dA) \\
&= \int_{M_d^-} \int_{\mathbb{R}} \left\| f(A,s)\gamma + \int_S \int_0^\infty f(A,s)\xi \left(\mathbb{1}_{[0,1]}(\|f(A,s)r\xi\|) - \mathbb{1}_{[0,1]}(\|r\xi\|) \right) \frac{dr}{r^\alpha} \theta(d\xi)) \right\| ds\pi(dA) \\
&= \int_{M_d^-} \int_{\mathbb{R}} \left\| f(A,s)\gamma + f(A,s) \int_S \xi \int_1^{\|f(A,s)\xi\|^{-1}} r^{-\alpha} dr\theta(d\xi)) \right\| ds\pi(dA) \\
&= \int_{M_d^-} \int_{\mathbb{R}} \left\| f(A,s)\gamma + f(A,s) \int_S \xi \frac{1}{1-\alpha} \left(\|f(A,s)\xi\|^{\alpha-1} - 1 \right) \theta(d\xi)) \right\| ds\pi(dA) \\
&= \int_{M_d^-} \int_{\mathbb{R}} \left\| f(A,s)\gamma + \frac{f(A,s)}{1-\alpha} \int_S \xi \|f(A,s)\xi\|^{\alpha-1} \theta(d\xi)) - \frac{f(A,s)}{1-\alpha} \int_S \xi\theta(d\xi)) \right\| ds\pi(dA) \\
&\leq \int_{M_d^-} \int_{\mathbb{R}} \left(\|f(A,s)\|\, \gamma + \frac{\|f(A,s)\|^\alpha}{1-\alpha} \theta(S) + \frac{\|f(A,s)\|}{1-\alpha} \theta(S) \right) ds\pi(dA) \\
&< \infty,
\end{aligned}$$

where we used $f \in \mathbb{L}^\alpha \cap \mathbb{L}^1$. For condition (4.2.4) we get

$$\int\limits_{M_d^-}\int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}^d} \left(1 \wedge \|f(A,s)x\|^2\right) \nu(dx)ds\pi(dA) =$$

$$= \int\limits_{M_d^-}\int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}^d} \mathbb{1}_{\{\|f(A,s)x\|\geq 1\}} \nu(dx)ds\pi(dA) + \tag{4.2.5}$$

$$+ \int\limits_{M_d^-}\int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}^d} \|f(A,s)x\|^2 \mathbb{1}_{\{\|f(A,s)x\|\leq 1\}} \nu(dx)ds\pi(dA).$$

The first term on the right hand side can be bounded by

$$\int\limits_{M_d^-}\int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}^d} \mathbb{1}_{\{\|f(A,s)x\|\geq 1\}} \nu(dx)ds\pi(dA) =$$

$$= \int\limits_{M_d^-}\int\limits_{\mathbb{R}}\int\limits_{S}\int\limits_{0}^{\infty} \mathbb{1}_{\{\|f(A,s)r\xi\|\geq 1\}} \frac{1}{r^{1+\alpha}} dr\theta(d\xi)ds\pi(dA)$$

$$= \int\limits_{M_d^-}\int\limits_{\mathbb{R}}\int\limits_{S}\int\limits_{\|f(A,s)\xi\|^{-1}}^{\infty} r^{-1-\alpha} dr\theta(d\xi)ds\pi(dA)$$

$$= \frac{1}{\alpha}\int\limits_{M_d^-}\int\limits_{\mathbb{R}}\int\limits_{S} \|f(A,s)\xi\|^\alpha \theta(d\xi)ds\pi(dA)$$

$$\leq \frac{\theta(S)}{\alpha}\int\limits_{M_d^-}\int\limits_{\mathbb{R}} \|f(A,s)\|^\alpha ds\pi(dA)$$

$$< \infty$$

and for the second term on the right hand side we get

$$\int\limits_{M_d^-}\int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}^d} \|f(A,s)x\|^2 \mathbb{1}_{\{\|f(A,s)x\|\leq 1\}} \nu(dx)ds\pi(dA) =$$

$$= \int\limits_{M_d^-}\int\limits_{\mathbb{R}}\int\limits_{S}\int\limits_{0}^{\infty} \|f(A,s)r\xi\|^2 \mathbb{1}_{\{\|f(A,s)r\xi\|\leq 1\}} \frac{1}{r^{1+\alpha}} dr\theta(d\xi)ds\pi(dA)$$

$$= \int\limits_{M_d^-}\int\limits_{\mathbb{R}}\int\limits_{S}\int\limits_{0}^{\|f(A,s)\xi\|^{-1}} \|f(A,s)\xi\|^2 r^{1-\alpha} dr\theta(d\xi)ds\pi(dA)$$

$$
\begin{aligned}
&= \frac{1}{2-\alpha} \int\limits_{M_d^-} \int\limits_{\mathbb{R}} \int\limits_{S} \|f(A,s)\xi\|^\alpha \theta(d\xi) ds\pi(dA) \\
&\leq \frac{\theta(S)}{2-\alpha} \int\limits_{M_d^-} \int\limits_{\mathbb{R}} \|f(A,s)\|^\alpha ds\pi(dA) \\
&< \infty.
\end{aligned}
$$

(ii) and (iii). Condition (4.2.3) can be bounded by

$$
\int\limits_{M_d^-} \int\limits_{\mathbb{R}} \|f(A,s)\Sigma\, f(A,s)^T\| ds\pi(dA) \leq \|\Sigma\| \int\limits_{M_d^-} \int\limits_{\mathbb{R}} \|f(A,s)\|^2 ds\pi(dA) < \infty,
$$

which follows from the boundedness of f together with $f \in \mathbb{L}^\delta$ for some $\delta \leq 2$. For condition (4.2.4) we use (4.2.5) again. For the first term on the right hand side of (4.2.5) we use the inequality

$$
\|f(A,s)\|\,\|x\| \geq \|f(A,s)x\| \geq 1
$$

which implies

$$
\|x\| \geq \frac{1}{\|f(A,s)\|}.
$$

This yields

$$
\begin{aligned}
\int\limits_{M_d^-} \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}^d} \mathbb{1}_{\{\|f(A,s)x\|\geq 1\}} \nu(dx) ds\pi(dA) &\leq \int\limits_{M_d^-} \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}^d} \mathbb{1}_{\left\{\|x\| \geq \frac{1}{\|f(A,s)\|}\right\}} \nu(dx) ds\pi(dA) \\
&= \int\limits_{M_d^-} \int\limits_{\mathbb{R}} \nu\left(\left\{\|x\| \geq \frac{1}{\|f(A,s)\|}\right\}\right) ds\pi(dA).
\end{aligned}
$$

Now we can apply the Potter bounds (Lemma 2.1.5), giving the existence of some t_0 such that for all $t \geq t_0$ a regular varying function (in this case ν) can be bounded. Therefore, we distinguish the cases $1/\|f(A,s)\| > t_0$ and $1/\|f(A,s)\| < t_0$. For the first case we set $\widetilde{C} := \sup\{\|f(A,s)\| : \|f(A,s)\| < 1/t_0\} \leq 1/t_0$. Then we can apply the Potter bounds for $t = 1/\widetilde{C} \geq t_0$ to get

$$
\begin{aligned}
&\int\limits_{M_d^-} \int\limits_{\mathbb{R}} \mathbb{1}_{\{1/\|f(A,s)\|>t_0\}}\, \nu\left(\left\{\|x\| \geq \frac{1}{\|f(A,s)\|}\right\}\right) ds\pi(dA) \leq \\
&\qquad \leq (1+\alpha-\delta) \int\limits_{M_d^-} \int\limits_{\mathbb{R}} \mathbb{1}_{\{1/\|f(A,s)\|>t_0\}}\, \nu\left(\left\{\|x\| \geq \frac{1}{\widetilde{C}}\right\}\right) \left(\frac{\|f(A,s)\|}{\widetilde{C}}\right)^\delta ds\pi(dA)
\end{aligned}
$$

$< \infty$.

In the other case we set $C := \sup \|f(A,s)\| < \infty$ and obtain

$$\begin{aligned}
&\int_{M_d^-} \int_{\mathbb{R}} \mathbb{1}_{\{1/\|f(A,s)\| \le t_0\}} \, \nu\left(\left\{\|x\| \ge \frac{1}{\|f(A,s)\|}\right\}\right) ds\pi(dA) \le \\
&\qquad \le \int_{M_d^-} \int_{\mathbb{R}} \mathbb{1}_{\{1/\|f(A,s)\| \le t_0\}} \, \nu\left(\left\{\|x\| \ge \frac{1}{C}\right\}\right) ds\pi(dA) \\
&\qquad = \nu\left(\left\{\|x\| \ge \frac{1}{C}\right\}\right) \;\; \pi \times \lambda\left(\left\{(A,s) : \|f(A,s)\| \ge \frac{1}{t_0}\right\}\right) \\
&\qquad < \infty,
\end{aligned}$$

since $f \in \mathbb{L}^\delta$. The second term on the right hand side of (4.2.5) can be bounded by

$$\begin{aligned}
&\int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \|f(A,s)x\|^2 \mathbb{1}_{\{\|f(A,s)x\| \le 1\}} \nu(dx) ds\pi(dA) = \\
&\quad = \int_{M_d^-} \int_{\mathbb{R}} \int_{\|x\|<1} \|f(A,s)x\|^2 \mathbb{1}_{\{\|f(A,s)x\| \le 1\}} \nu(dx) \\
&\qquad\qquad + \int_{\|x\| \ge 1} \|f(A,s)x\|^2 \mathbb{1}_{\{\|f(A,s)x\| \le 1\}} \nu(dx) ds\pi(dA) \\
&\quad \le \int_{M_d^-} \int_{\mathbb{R}} \int_{\|x\|<1} \|f(A,s)\|^2 \|x\|^2 \nu(dx) + \int_{\|x\| \ge 1} \|f(A,s)x\|^\delta \nu(dx) ds\pi(dA) \\
&\qquad\qquad + \int_{M_d^-} \int_{\mathbb{R}} \|f(A,s)\|^\delta ds\pi(dA) \int_{\|x\| \ge 1} \|x\|^\delta \nu(dx) \\
&\quad < \infty,
\end{aligned}$$

where we used the fact that for bounded functions f the assumption $f \in \mathbb{L}^\delta$, $\delta < 2$, implies $f \in \mathbb{L}^2$. Moreover, note that $\int_{\|x\| \ge 1} \|x\|^\delta \nu(dx) < \infty$ by Sato (2002), Corollary 25.8, since $0 < \delta < \alpha$ and hence the underlying Lévy process has a finite δth moment (cf. Corollary 2.3.11). Condition (4.2.2) in Theorem 4.2.6 can be reformulated as

$$\begin{aligned}
&\int_{M_d^-} \int_{\mathbb{R}} \left\| f(A,s)\gamma + \int_{\mathbb{R}^d} f(A,s)x \left(\mathbb{1}_{[0,1]}(\|f(A,s)x\|) - \mathbb{1}_{[0,1]}(\|x\|) \right) \nu(dx) \right\| ds\pi(dA) = \\
&= \int_{M_d^-} \int_{\mathbb{R}} \left\| f(A,s)\gamma + \int_{\|x\|>1} f(A,s)x \mathbb{1}_{\{\|f(A,s)x\| \le 1\}} \nu(dx) \right.
\end{aligned}$$

$$- \int_{\|x\|\leq 1} f(A,s)x\mathbb{1}_{\{\|f(A,s)x\|>1\}}\nu(dx)\Big\| ds\pi(dA) =: T.$$

In case (ii) we use $\|f(A,s)\| \leq C$ and thus T can be bounded by

$$T \leq \int_{M_d^-}\int_{\mathbb{R}} \|f(A,s)\|^\delta \left(C^{1-\delta}|\gamma| + \int_{\|x\|>1} \|x\|^\delta \nu(dx) + C^{1-\delta} \int_{\|x\|\in(\frac{1}{C},1]} \|x\|^\delta \nu(dx) \right) ds\pi(dA).$$

In case (iii), we know $\gamma = -\int_{\|x\|>1} x\nu(dx)$. Since $\alpha > 1$ and $\delta < \alpha$, we can arbitrarily choose a $\xi \in (\delta, \alpha)$ with $\xi > 1$. This yields

$$\begin{aligned}
T &= \int_{M_d^-}\int_{\mathbb{R}} \Big\| - \int_{\|x\|>1} f(A,s)x\nu(dx) + \int_{\|x\|>1} f(A,s)x\mathbb{1}_{\{\|f(A,s)x\|\leq 1\}}\nu(dx) \\
&\qquad - \int_{\|x\|\leq 1} f(A,s)x\mathbb{1}_{\{\|f(A,s)x\|>1\}}\nu(dx)\Big\| ds\pi(dA) \\
&= \int_{M_d^-}\int_{\mathbb{R}} \Big\| - \int_{\|x\|>1} f(A,s)x\mathbb{1}_{\{\|f(A,s)x\|>1\}}\nu(dx) \\
&\qquad - \int_{\|x\|\leq 1} f(A,s)x\mathbb{1}_{\{\|f(A,s)x\|>1\}}\nu(dx)\Big\| ds\pi(dA) \\
&\leq \int_{M_d^-}\int_{\mathbb{R}}\int_{\mathbb{R}^d} \|f(A,s)x\|^\xi \mathbb{1}_{\{\|x\|>\frac{1}{C}\}}\nu(dx)ds\pi(dA) \\
&\leq C^{\xi-\delta} \int_{M_d^-}\int_{\mathbb{R}} \|f(A,s)\|^\delta ds\pi(dA) \int_{\|x\|>\frac{1}{C}} \|x\|^\xi \nu(dx) \\
&< \infty.
\end{aligned}$$

□

4.3 Mixed Moving Average Processes

Mixed moving average (short MMA) processes have been first introduced by Surgailis et al. (1993) in the univariate stable case. As we have already mentioned in the previous sections, they are integrals over a given kernel function with respect to a Lévy basis.

Definition 4.3.1 (Mixed Moving Average Process). Let Λ be an $\mathbb{R}^d$-valued Lévy basis on $M_d^- \times \mathbb{R}$ and let $f : M_d^- \times \mathbb{R} \mapsto M_{n,d}$ be a measurable function (*kernel function*). If

the process

$$X_t := \int_{M_d^-} \int_{\mathbb{R}} f(A, t-s)\Lambda(dA, ds)$$

exists in the sense of Theorem 4.2.6 for all $t \in \mathbb{R}$, it is called an *n-dimensional mixed moving average process* (short *MMA process*).

Note that we could also define "generalized MMA" processes by integrating over a slightly more general function $g : M_d^- \times \mathbb{R} \times \mathbb{R} \mapsto M_{n,d}$, which gives us

$$X_t = \int_{M_d^-} \int_{\mathbb{R}} g(A, t, s)\Lambda(dA, ds).$$

However, the extension of all upcoming results is trivial, so we stated the results for the notationally easier case of Definition 4.3.1. Moreover, an MMA process is obviously always stationary and this needs not to be true for generalized MMA processes. Note also that M_d^- can obviously be replaced by M_d or basically any other Borel set. Again we state everything for M_d^-, because this eases notation and is the canonical choice in the supOU case.
Existence of the MMA processes follows directly from Theorem 4.2.6 and Theorem 4.2.7. Especially Theorem 4.2.7 turns out to be very useful in this setting, since it is based on similar conditions compared to the key conditions of the following theorem: regular variation of the driving Lévy measure ν and $f \in \mathbb{L}^\alpha(\lambda \times \pi)$.
The theorem is the multivariate analog of (3.1) in Proposition 3.2 of Fasen (2005), where the same conditions, simplified to the univariate set-up, are used. A similar result also exists for the special case of a univariate filtered Lévy process, where the kernel function f is continuous and of compact support, see Hult and Lindskog (2005), Theorem 22.

Theorem 4.3.2. *Let Λ be an $\mathbb{R}^d$-valued Lévy basis on $M_d^- \times \mathbb{R}$ with generating quadruple $(\gamma, \Sigma, \nu, \pi)$ and let $\nu \in RV(\alpha, l, \mu_\nu)$. If $X_0 = \int_{M_d^-} \int_{\mathbb{R}} f(A, s)\Lambda(dA, ds)$ exists (in the sense of Theorem 4.2.6), $f \in \mathbb{L}^\alpha(\lambda \times \pi)$ and $\mu_\nu(f^{-1}(A, s)(\mathbb{R}^n \backslash \{0\})) = 0$ does not hold for $\pi \times \lambda$ almost-every (A, s), then $X_0 \in RV(\alpha, l, \mu_X)$ with*

$$\mu_X(B) := \int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \mathbb{1}_B\left(f(A, s)x\right) \mu_\nu(dx)ds\pi(dA).$$

Proof. From Theorem 4.2.6 we know that the distribution of X is infinitely divisible. Following Theorem 4.2.3 it is sufficient to prove that its Lévy measure ν_X is regularly

varying. The concrete representation

$$\nu_X = \int\limits_{M_d^-} \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}^d} \mathbb{1}_B(f(A,s)x)\nu(dx)ds\pi(dA)$$

is also known from Theorem 4.2.6. Regular variation of ν then yields the existence of a constant $\alpha > 0$, a slowly varying function l and a Radon measure μ_ν on $\mathcal{B}(\overline{\mathbb{R}}^d\backslash\{0\})$ with $\mu_\nu(\overline{\mathbb{R}}^d\backslash\mathbb{R}^d) = 0$ such that, as $u \to \infty$,

$$u^\alpha l(u)\nu(u\,\cdot) \xrightarrow{v} \mu_\nu(\cdot).$$

Using Theorem 2.1.9 and Fatou's Lemma, we have that for all compact sets $B \in \mathcal{K}$

$$\begin{aligned}
\limsup_{u\to\infty} u^\alpha l(u) \int\limits_{M_d^-} \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}^d} \mathbb{1}_{uB}\left(f(A,s)x\right)\nu(dx)ds\pi(dA) \leq \\
&\leq \int\limits_{M_d^-} \int\limits_{\mathbb{R}} \limsup_{u\to\infty} u^\alpha l(u) \int\limits_{\mathbb{R}^d} \mathbb{1}_{uB}\left(f(A,s)x\right)\nu(dx)ds\pi(dA) \\
&\leq \int\limits_{M_d^-} \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}^d} \mathbb{1}_{B}\left(f(A,s)x\right)\mu_\nu(dx)ds\pi(dA)
\end{aligned}$$

and conversely for all open sets $B \in \mathcal{G}$ that are relatively compact

$$\begin{aligned}
\liminf_{u\to\infty} u^\alpha l(u) \int\limits_{M_d^-} \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}^d} \mathbb{1}_{uB}\left(f(A,s)x\right)\nu(dx)ds\pi(dA) \geq \\
&\geq \int\limits_{M_d^-} \int\limits_{\mathbb{R}} \liminf_{u\to\infty} u^\alpha l(u) \int\limits_{\mathbb{R}^d} \mathbb{1}_{uB}\left(f(A,s)x\right)\nu(dx)ds\pi(dA) \\
&\geq \int\limits_{M_d^-} \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}^d} \mathbb{1}_{B}\left(f(A,s)x\right)\mu_\nu(dx)ds\pi(dA).
\end{aligned}$$

Note here that for any set $B \in \mathcal{K}$ (resp. $\mathcal{G}$) also the preimage $f(A,s)^{-1}(B) \in \mathcal{K}$ (resp. $\mathcal{G}$) for all A, s, since $f(A,s)$ is for fixed A, s a linear mapping. This yields the vague convergence

$$\begin{aligned}
u^\alpha l(u)\nu_X(u\,\cdot) &= u^\alpha l(u) \int\limits_{M_d^-} \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}^d} \mathbb{1}_{u\cdot}\left(f(A,s)x\right)\nu(dx)ds\pi(dA) \\
&\xrightarrow{v} \int\limits_{M_d^-} \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}^d} \mathbb{1}_{(\cdot)}\left(f(A,s)x\right)\mu_\nu(dx)ds\pi(dA) = \mu_X(\cdot).
\end{aligned}$$

It remains to prove that μ_X is again a Radon measure with $\mu_X(\overline{\mathbb{R}}^n \backslash \mathbb{R}^n) = 0$. The second property follows directly from the observation

$$\mathbb{1}_{\left(\overline{\mathbb{R}}^n \backslash \mathbb{R}^n\right)}(f(A,s)x) \leq \mathbb{1}_{\left(\overline{\mathbb{R}}^n \backslash \mathbb{R}^n\right)}(x).$$

For the local finiteness of μ_X, take some compact $B \in \mathcal{B}(\overline{\mathbb{R}}^d \backslash \{0\})$, i.e. there exists some finite $r > 0$ such that $B \subseteq V_r := \{x : \|x\| > r\}$. For all x with $f(A,s)x \in B \subseteq V_r$ we have $r < \|f(A,s)x\| \leq \|f(A,s)\| \, \|x\|$. By using $f \in \mathbb{L}^\alpha(\lambda \times \pi)$ and the local finiteness of μ_ν, we get

$$\begin{aligned}
\mu_X(B) &= \int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \mathbb{1}_B \left(f(A,s)x\right) \mu_\nu(dx) ds \pi(dA) \\
&\leq \int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \mathbb{1}_{(r,\infty)} \left(\|f(A,s)\| \, \|x\|\right) \mu_\nu(dx) ds \pi(dA) \\
&= \int_{M_d^-} \int_{\mathbb{R}} \mu_\nu(\left\{x : \|x\| \geq \|f(A,s)\|^{-1} r\right\}) \, \mathbb{1}_{\mathbb{R}\backslash\{0\}}(\|f(A,s)\|) \, ds \pi(dA) \\
&= \mu_\nu(V_r) \int_{M_d^-} \int_{\mathbb{R}} \|f(A,s)\|^\alpha ds \pi(dA) \\
&< \infty.
\end{aligned}$$

□

The theorem shows that the tail behavior of the driving Lévy measure determines the tail behavior of the MMA process. Since the Lévy measure is related only to the jumps of the underlying Lévy process, we see that the regular variation of the MMA process is caused by the jumps of the underlying Lévy process. Furthermore, we intuitively have that the extremes of the MMA process are caused by a single extremely big jump in the Lévy basis.

Remark 4.3.3. Another important consequence of the theorem is that we know the concrete measure μ_X of regular variation. This is useful to describe the location or mass of the extremes in $\mathbb{R}^n$. It is similar to the spectral measure in an analogue definition of regular variation, see Theorem 2.1.11. See also Example 4.4.4 for some calculations of these measures in the Ornstein-Uhlenbeck case.

As mentioned before, Theorem 4.3.2 uses two crucial conditions. The first one is the regular variation of the driving Lévy measure, meaning that the tail behavior of the input

determines the tail behavior of the resulting MMA process. The second condition $f \in \mathbb{L}^\alpha(\lambda \times \pi)$ is a restriction on the function f. We will now analyze its restrictiveness by looking at necessary conditions. Therefore, we define the set

$$\mathbb{J}^\alpha(\lambda \times \pi) := \left\{ f : M_d^- \times \mathbb{R} \mapsto M_{n,d} \text{ measurable}, \int_{M_d^-} \int_{\mathbb{R}} j(f(A,s))^\alpha ds\pi(dA) < \infty \right\},$$

where $j(A)$ is the modulus of injectivity of A.
The following theorem extends even the univariate work by Fasen (2005) and Hult and Lindskog (2005), where necessary conditions are not considered. Note that the focus is on necessary conditions on f whereas Jacobsen et al. (2009) considered whether regular variation of a moving average implies regular variation of the driving Lévy process in the univariate case.

Theorem 4.3.4. *Let Λ be an $\mathbb{R}^d$-valued Lévy basis on $M_d^- \times \mathbb{R}$ with generating quadruple $(\gamma, \Sigma, \nu, \pi)$ and let $\nu \in RV(\alpha, l, \mu_\nu)$. If $X_0 = \int_{M_d^-} \int_{\mathbb{R}} f(A,s)\Lambda(dA,ds)$ exists and $\mu_\nu(f^{-1}(A,s)(\mathbb{R}^n \backslash \{0\})) = 0$ does not hold for $\pi \times \lambda$ almost-every (A,s), then $f \in \mathbb{J}^\alpha(\lambda \times \pi)$ is a necessary condition for $X_0 \in RV(\alpha, l, \mu_X)$ with*

$$\mu_X(B) := \int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \mathbb{1}_B\left(f(A,s)x\right) \mu_\nu(dx) ds\pi(dA).$$

Proof. We use a simple contradiction. Suppose $f \notin \mathbb{J}^\alpha(\lambda \times \pi)$, i.e.

$$\int_{M_d^-} \int_{\mathbb{R}} j(f(A,s))^\alpha ds\pi(dA) = \infty.$$

Since μ_ν is nonzero there is a positive number $r > 0$ such that $\mu_\nu(V_r) > 0$. Then we use the relation

$$j(f(A,s)) \leq \frac{\|f(A,s)x\|}{\|x\|}$$

for all $x \in \mathbb{R}^d$ and get

$$\begin{aligned}
\mu_X(V_r) &= \int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \mathbb{1}_{V_r}\left(f(A,s)x\right) \mu_\nu(dx) ds\pi(dA) \\
&\geq \int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \mathbb{1}_{(r,\infty)}\left(j(f(A,s))\,\|x\|\right) \mu_\nu(dx) ds\pi(dA)
\end{aligned}$$

$$
\begin{aligned}
&= \int_{M_d^-} \int_{\mathbb{R}} \mu_\nu(\{x : \|x\| \geq j(f(A,s))^{-1} r\}) \, \mathbb{1}_{\mathbb{R}\setminus\{0\}}(j(f(A,s)) \, ds\pi(dA) \\
&= \mu_\nu(V_r) \int_{M_d^-} \int_{\mathbb{R}} j(f(A,s))^\alpha ds\pi(dA) \\
&= \infty
\end{aligned}
$$

and this is a contradiction to μ_X being a Radon measure.

□

Now we have necessary conditions as well as sufficient conditions and both lie close together. Since $j(f(A,s)) \leq \|f(A,s)\|$ we immediately have $\mathbb{L}^\alpha(\lambda \times \pi) \subseteq \mathbb{J}^\alpha(\lambda \times \pi)$. In the univariate case we even have $\mathbb{L}^\alpha(\lambda \times \pi) = \mathbb{J}^\alpha(\lambda \times \pi)$ and thus we get necessary and sufficient conditions.
Having proved the regular variation of the random vector, we can now easily get the regular variation of the process X_t.

Corollary 4.3.5. *Given the conditions of Theorem 4.3.2, the MMA process $(X_t)_{t\in\mathbb{R}}$ is also regularly varying with index α as a process.*

Proof. We have to show that the results also hold for the finite dimensional distributions of X_t. For $m \in \mathbb{N}$ and $\mathbf{t} = (t_1, \ldots, t_m) \in \mathbb{R}^m$ we have

$$
\begin{pmatrix} X_{t_1} \\ \vdots \\ X_{t_m} \end{pmatrix} = \begin{pmatrix} \int_{M_d^-} \int_{\mathbb{R}} f(A, t_1 - s)\Lambda(dA, ds) \\ \vdots \\ \int_{M_d^-} \int_{\mathbb{R}} f(A, t_m - s)\Lambda(dA, ds) \end{pmatrix} = \int_{M_d^-} \int_{\mathbb{R}} \begin{pmatrix} f(A, t_1 - s) \\ \vdots \\ f(A, t_m - s) \end{pmatrix} \Lambda(dA, ds)
$$

$$
= \int_{M_d^-} \int_{\mathbb{R}} g(A, \mathbf{t}, s) \, \Lambda(dA, ds)
$$

with the function $g : M_d^- \times \mathbb{R}^m \times \mathbb{R} \mapsto M_{nm,d}$ defined by

$$
g(A, \mathbf{t}, s) := \begin{pmatrix} f(A, t_1 - s) \\ \vdots \\ f(A, t_m - s) \end{pmatrix}.
$$

Next we show that $f \in \mathbb{L}^\beta(\lambda \times \pi)$ implies $g \in \mathbb{L}^\beta(\lambda \times \pi)$ for all $\beta > 0$. Therefore, we choose the matrix norm

$$
\|A\| := \max_{i,j}\{|a_{ij}|\}.
$$

We get

$$\begin{aligned}
&\int_{M_d^-}\int_{\mathbb{R}} \|g(A,\mathbf{t},s)\|^\beta \, ds\,\pi(dA) = \\
&\qquad = \int_{M_d^-}\int_{\mathbb{R}} \left\| \begin{pmatrix} f(A,t_1-s) \\ \vdots \\ f(A,t_m-s) \end{pmatrix} \right\|^\beta ds\,\pi(dA) \\
&\qquad = \int_{M_d^-}\int_{\mathbb{R}} \max\left\{\|f(A,t_1-s)\|,\ldots,\|f(A,t_m-s)\|\right\}^\beta ds\,\pi(dA) \\
&\qquad \leq \int_{M_d^-}\int_{\mathbb{R}} \|f(A,t_1-s)\|^\beta + \ldots + \|f(A,t_m-s)\|^\beta \, ds\,\pi(dA) \\
&\qquad < \infty,
\end{aligned}$$

since $f \in \mathbb{L}^\beta(\lambda \times \pi)$. If the existence of X_t is ensured by Theorem 4.2.7 (ii) or (iii), this implies that for the existence and regular variation of $(X_{t_1}^T,\ldots,X_{t_m}^T)^T$ a simple application of Theorem 4.3.2 and Theorem 4.2.7 conclude. However, in general we note that assuming existence of X_t in the sense of Theorem 4.2.6 implies that each of the m individual integrals of $(X_{t_1}^T,\ldots,X_{t_m}^T)^T$ exists as a limit of approximating sums in probability. From these individual approximating sums one easily constructs a sequence of approximating sums for $\int_{M_d^-}\int_{\mathbb{R}} \left\| \left(f(A,t_1-s)^T,\ldots,f(A,t_m-s)^T\right)^T \right\|^\beta \Lambda(dA,ds)$ converging in probability. Hence, the necessary and sufficient existence conditions of Theorem 4.2.6 are satisfied and Theorem 4.3.2 shows the regular variation of $(X_{t_1}^T,\ldots,X_{t_m}^T)^T$.

□

A very important class of heavy tailed distributions are α-stable distributions with $\alpha \in (0,2)$. See Samorodnitsky and Taqqu (1994) for a detailed introduction. In Theorem 4.2.7 we have already given a criterion for the existence of MMA processes with stable driving Lévy process. Similar to Theorem 4.3.2, there is also a well-known link between stability of the driving Lévy measure and stability of the MMA process.

Lemma 4.3.6. *If the driving Lévy process of an MMA process X_t is α-stable and its Lévy measure is non-degenerate, then X_t is also α-stable.*

Proof. From Theorem 14.3 in Sato (2002) we have the result that α-stability of an infinitely divisible distribution is equivalent to

$$\Sigma = 0 \quad \text{and} \quad \nu(\cdot) = b^{-\alpha}\,\nu(b^{-1}\,\cdot) \quad \text{for all } b > 0.$$

Using the assumption together with Theorem 4.2.6, we immediately have $\Sigma_{X_t} = 0$ and $\nu_{X_t}(\cdot) = b^{-\alpha}\, \nu_{X_t}(b^{-1}\, \cdot)$.

□

Now we apply this result to multivariate continuous-time autoregressive moving average (MCARMA) processes.

Example 4.3.7 (MCARMA Processes). Univariate Lévy-driven CARMA processes have been introduced by Brockwell (2001) and they have been extended to multivariate CARMA (MCARMA) processes by Marquardt and Stelzer (2007). A d-dimensional MCARMA(p,q) process, $p > q$, driven by a two-sided square integrable Lévy process $(L_t)_{t\in\mathbb{R}}$ with $\mathbb{E}(L_1) = 0$ and $\mathbb{E}(L_1 L_1^T) = \Sigma_L$ can be formally interpreted as the stationary solution to the p-th order d-dimensional differential equation

$$P(D)Y_t = Q(D)DL_t,$$

where D denotes the differentiation operator with respect to t. The autoregressive and moving average polynomials are given by

$$P(z) = I_d z^p + A_1 z^{p-1} + \ldots + A_p \quad \text{and} \quad Q(z) = B_0 z^q + B_1 z^{q-1} + \ldots + B_q$$

with $A_1, \ldots, A_p,\ B_0, \ldots, B_q \in M_d$ such that $B_q \neq 0$ and $\{z \in \mathbb{C} : det(P(z)) = 0\} \subset \mathbb{R}\backslash\{0\} + i\mathbb{R}$. The MCARMA process Y_t can be represented as a moving average process

$$Y_t = \int_{\mathbb{R}} f(t-s)dL_s$$

with kernel function $f : \mathbb{R} \mapsto M_d$ given by

$$f(t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{iut} P(iu)^{-1} Q(iu) du.$$

Obviously, MCARMA processes are MMA processes and thus we can apply Lemma 4.3.6 to obtain an α-stable MCARMA process by using an α-stable driving Lévy process. Furthermore, by Proposition 3.32 of Marquardt and Stelzer (2007) we know that in the case $p > q+1$ MCARMA processes have continuous sample paths and these are $p-q-1$ times differentiable. This shows that in the case $\alpha \in (0,2)$ and $p > q+1$ we can get heavy tailed MCARMA processes, where the heavy tails come from the jumps of the underlying Lévy process, but the paths of the observed process are continuous and may even be differentiable.

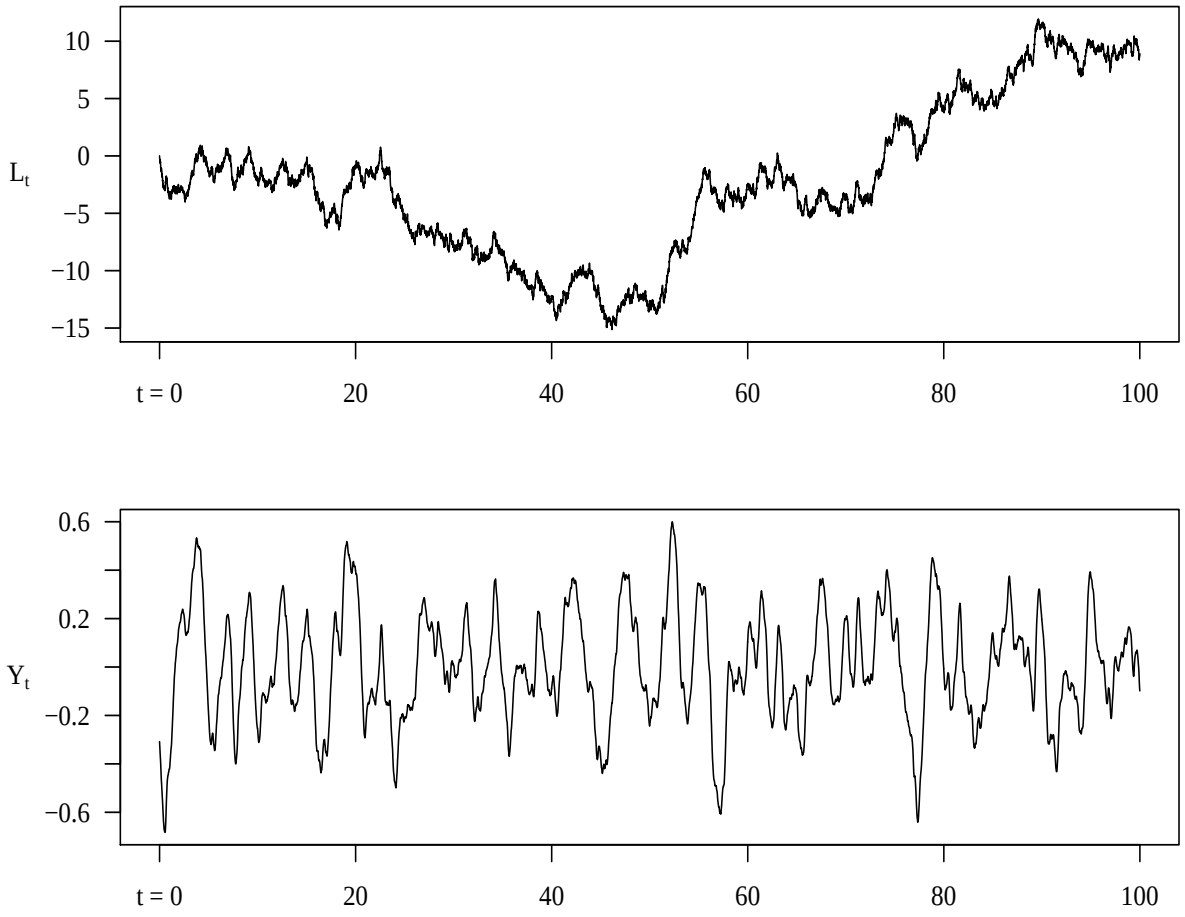


Figure 4.3.1: Simulations of one path of the driving Lévy process L_t and the CARMA(3,1) process Y_t in the α-stable case with $\alpha = 2$

To illustrate this, we simulated several univariate CARMA(3,1) processes. They are given by the autoregressive and moving average polynomials

$$p(z) = z^3 + 4.5z^2 + 6.5z + 3 \quad \text{and} \quad q(z) = z.$$

The CARMA(3,1) process can then given in its state space representation (see Marquardt and Stelzer (2007), Theorem 3.12)

$$G(t) = \int_{-\infty}^{t} e^{A(t-u)}\, \beta\, dL_u,$$

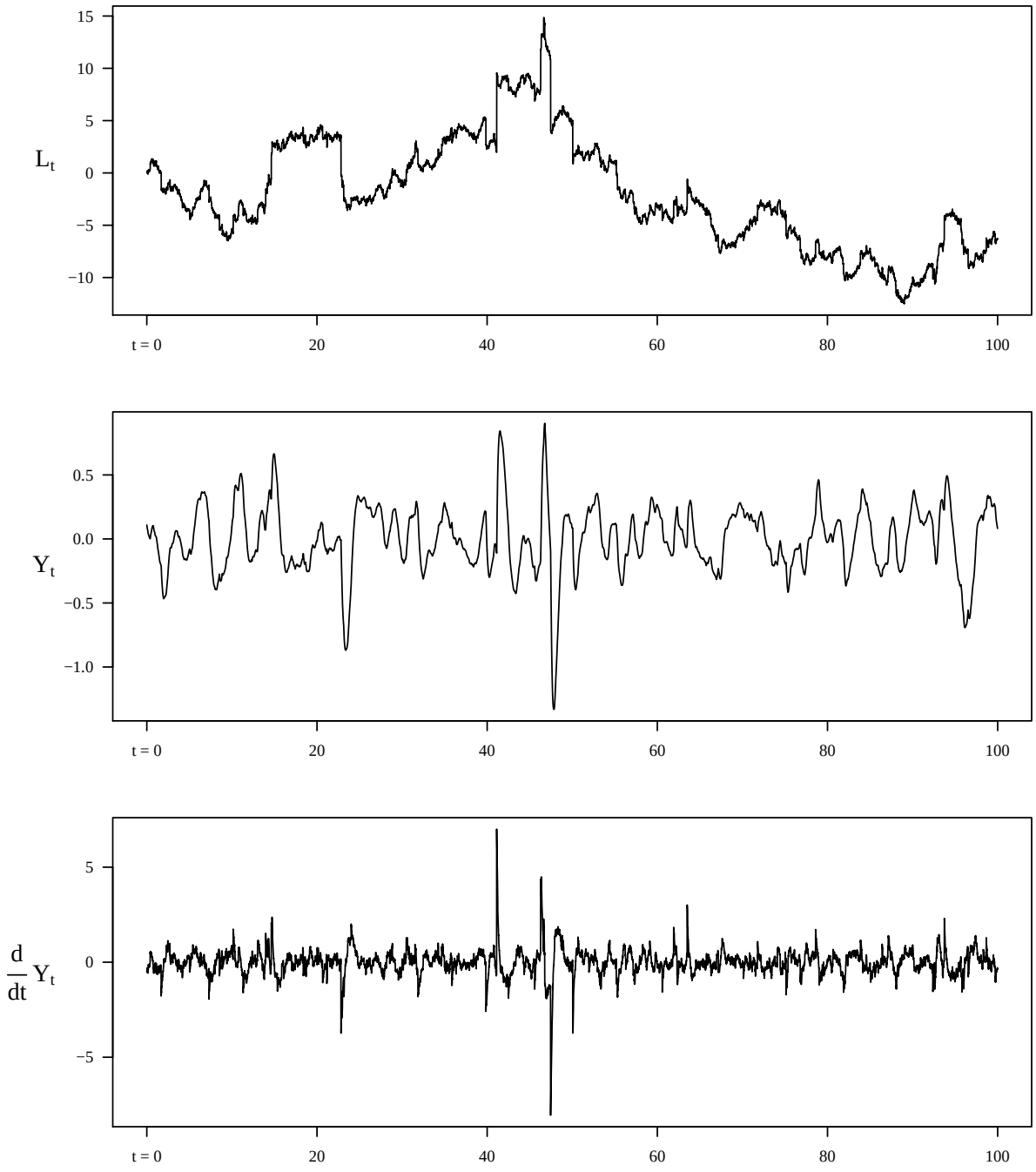


Figure 4.3.2: Simulations of one path of the driving Lévy process L_t, the CARMA(3,1) process Y_t and its derivative in the α-stable case with $\alpha = 1.5$

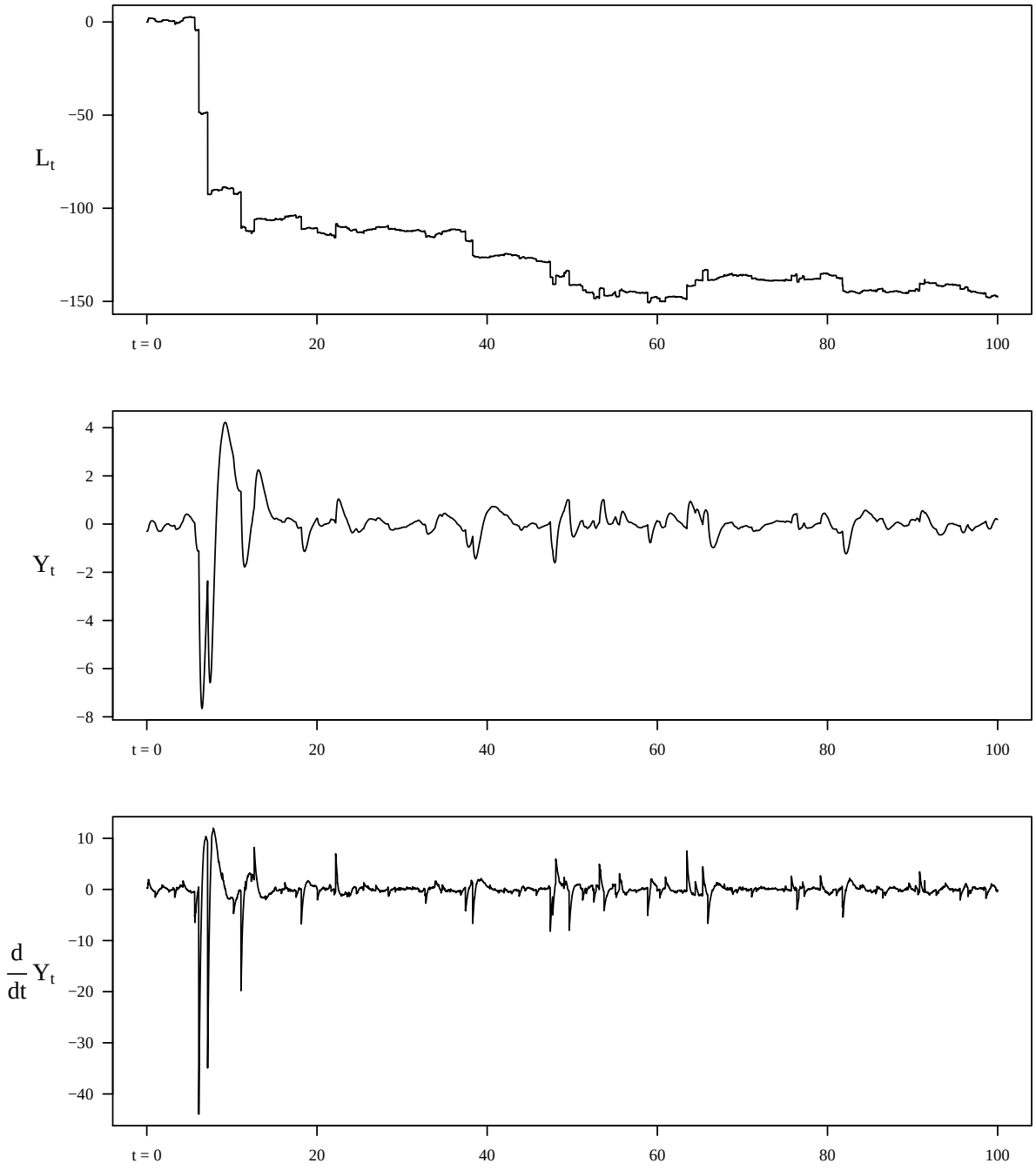


Figure 4.3.3: Simulations of one path of the driving Lévy process L_t, the CARMA(3,1) process Y_t and its derivative in the α-stable case with $\alpha = 1$

where

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -3 & -6.5 & -4.5 \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} 0 \\ 1 \\ -4.5 \end{pmatrix}.$$

This representation has the advantage that it applies also in the multivariate setting and it directly includes the derivatives of the CARMA process, as long as they exist. In that case, we have $Y_t = G_1(t)$ and $\frac{d}{dt}Y_t = G_2(t)$. Due to the foregoing results G is regularly varying with index α (resp. α-stable) if L is so.

For the driving Lévy process L_t we used a symmetric α-stable Lévy motion without skewness and with α-values of 2 (Brownian Motion), 1.5 and 1 (both heavy-tailed). Furthermore, we plotted the simulated values after a burn-in period of 1000 to ensure stationarity. In all three cases, one can see nicely, how the tail behavior of the driving Lévy process determines the tail behavior of the continuous CARMA(3,1) process.

In Figure 4.3.1 we see the case $\alpha = 2$ where the integrator is a light tailed Brownian Motion and the resulting CARMA process is also light tailed. In the cases $\alpha = 1.5$ (see Figure 4.3.2) and $\alpha = 1$ (see Figure 4.3.3) the driving process is heavy tailed and, as α is decreasing, the process is more and more determined by only a few very large jumps. The respective CARMA process is also heavy tailed and oscillates around the mean except for some large, but continuous shocks. For these two cases we also plotted the first derivatives of the paths of the CARMA process, which are not continuous anymore, but jointly α-stable together with the process itself.

4.4 Application to SupOU Processes

One example of MMA processes are superpositions of Ornstein-Uhlenbeck processes, or supOU processes for short. They are especially useful in modeling the stochastic volatility in continuous time models or long range dependent time series. For an introduction to univariate supOU processes see Barndorff-Nielsen (2001) and for the extension to multivariate supOU processes we refer to Barndorff-Nielsen and Stelzer (2011a).

Definition 4.4.1 ($\mathbb{R}^d$-Valued SupOU Process). Let Λ be an $\mathbb{R}^d$-valued Lévy basis on $M_d^- \times \mathbb{R}$. If the process

$$X_t := \int_{M_d^-} \int_{-\infty}^{t} e^{A(t-s)} \Lambda(dA, ds)$$

exists for all $t \in \mathbb{R}$, it is called $\mathbb{R}^d$-*valued supOU process*.

We easily see that supOU processes are MMA processes with special kernel function

$$f(A,s) = e^{As}\,\mathbb{1}_{[0,\infty)}(s).$$

Consequently, existence of supOU processes is covered by Theorem 4.2.6. But if we take the special properties of supOU processes into account, some more accessible sufficient conditions for the existence can be given.

Theorem 4.4.2 (Barndorff-Nielsen and Stelzer (2011a), Theorem 3.1). *Let X_t be an $\mathbb{R}^d$-valued supOU process as defined in Definition 4.4.1. If*

$$\int_{\|x\|>1} \ln(\|x\|)\nu(dx) < \infty$$

and there exist measurable functions $\rho : M_d^- \mapsto \mathbb{R}^+\backslash\{0\}$ and $\kappa : M_d^- \mapsto [1,\infty)$ such that

$$\left\|e^{As}\right\| \leq \kappa(A)e^{-\rho(A)s}\ \forall s \in \mathbb{R}^+\ \pi\text{-almost surely and } \int\limits_{M_d^-} \frac{\kappa(A)^2}{\rho(A)}\pi(dA) < \infty,$$

then the supOU process $X_t = \int_{M_d^-}\int_{-\infty}^t e^{A(t-s)}\Lambda(dA,ds)$ is well defined for all $t \in \mathbb{R}$ and stationary. Furthermore, the stationary distribution of X_t is infinitely divisible with characteristic triplet $(\gamma_X, \Sigma_X, \nu_X)$ given by Theorem 4.2.6.

Now we want to go one step further and analyze the tail behavior, but regular variation of the supOU processes follows directly from Theorem 4.3.2.

Corollary 4.4.3. *Let $\Lambda \in \mathbb{R}^d$ be a Lévy basis on $M_d^- \times \mathbb{R}$ with generating quadruple $(\gamma, \Sigma, \nu, \pi)$ and let $\nu \in RV(\alpha, l, \mu_\nu)$. If the conditions of Theorem 4.4.2 hold and additionally*

$$\int\limits_{M_d^-} \frac{\kappa(A)^\alpha}{\rho(A)}\,\pi(dA) < \infty,$$

then $X_0 = \int_{M_d^-}\int_{\mathbb{R}^+} e^{As}\Lambda(dA,ds) \in RV(\alpha, l, \mu_X)$ with Radon measure

$$\mu_X(\cdot) := \int\limits_{M_d^-}\int\limits_{\mathbb{R}^+}\int\limits_{\mathbb{R}^d} \mathbb{1}_{(\cdot)}\left(e^{As}x\right)\mu_\nu(dx)ds\pi(dA).$$

Proof. Using the given conditions, we have

$$\int_{M_d^-}\int_{\mathbb{R}^+} \left\|e^{As}\right\|^\alpha ds\pi(dA) \leq \int_{M_d^-}\int_{\mathbb{R}^+} \kappa(A)^\alpha e^{-\alpha\rho(A)s}ds\pi(dA)$$

$$= \alpha^{-1} \int_{M_d^-} \frac{\kappa(A)^\alpha}{\rho(A)} \pi(dA)$$
$$< \infty$$

and thus $e^{As} \in \mathbb{L}^\alpha(\lambda \times \pi)$. It is left to show that $\mu_\nu(f^{-1}(A,s)\mathbb{R}^n) = 0$ does not hold for $\pi \times \lambda$ almost-every (A,s), but since

$$\mu_\nu\left(e^{-As}\,\mathbb{R}^d\right) = \mu_\nu\left(\mathbb{R}^d\right)$$

for any (A,s), this follows simply from μ_ν being a nonzero measure.

□

For illustration, let us now calculate the measures μ_X of regular variation in some special cases.

Example 4.4.4 (Measure of Regular Variation of OU Processes). SupOU processes with probability measure π being a one-point measure (i.e. $\pi(A) = 1$ for some $A \in M_d^-$) are called *Ornstein-Uhlenbeck* (OU) processes. They can be defined as stochastic integrals of the form

$$\int_{-\infty}^{t} e^{A(t-s)}\, dL_s,$$

where (L_t) is a Lévy process with values in $\mathbb{R}^d$ and $A \in M_d^-$ is the parameter matrix. Applying Corollary 4.4.3, their measure of regular variation can be given by

$$\mu_X(B) := \int_{\mathbb{R}^+} \mu_\nu\left(e^{-As}B\right) ds.$$

We consider several examples in the case $d = 2$. Let us first assume that the mass of the measure μ_ν is concentrated on a straight line, i.e. on the points of the form $h = (a(1,b)^T)_{a\in\mathbb{R}\setminus\{0\}}$ for $b \in \mathbb{R}$, see Figure 4.4.1 for an example with $b = 0.5$.

1. If $A = c\,I_d, c \in \mathbb{R}^-$, is a multiple of the identity matrix, then

$$\mu_X(B) = \int_{\mathbb{R}^+} \mu_\nu\left(e^{-cs}B\right) ds = \int_{\mathbb{R}^+} e^{c\alpha s}\mu_\nu\left(B\right) ds = -\frac{\mu_\nu(B)}{c\,\alpha}.$$

   Consequently, μ_X has its mass in the same directions as μ_ν and thus its mass is also concentrated on h.

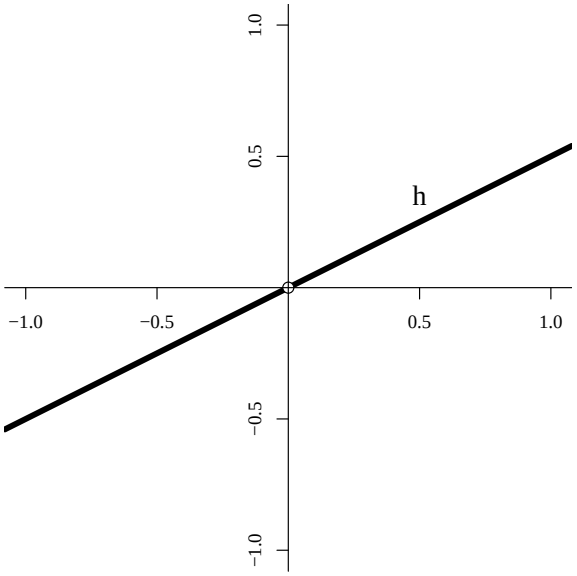


Figure 4.4.1: Mass of the measures μ_ν and μ_X in case *1.*

2. If $A = diag(a_1, a_2)$ is a diagonal matrix, then the mass of μ_X is concentrated on the cones between the straight line h and one of the two axes, see Figure 4.4.2. The mass is drawn to the horizontal axis if $a_2 > a_1$ and to the vertical axis if $a_1 > a_2$ (i.e. to the axis associated with the slower exponential decay rate). Intuitively this happens as follows. An extreme jump $(x_1, x_2)^T$ occurred at some time u in the past and had direction s. This causes an extreme value $(e^{d_1(t-u)}x_1, e^{d_2(t-u)}x_2)^T$ at a later time t. Since one of the components decays slower, this extreme event is now in a direction closer to the direction with the slowest exponential decay.

3. If A is real diagonalizable, i.e. $A = UDU^{-1}$ with $D = diag(d_1, d_2)$, then the mass is drawn to the eigenspace e that belongs to the biggest eigenvalue $\max(d_1, d_2)$. This means that the mass is concentrated on the cone between e and h, see Figure 4.4.3. This follows immediately by a change of the basis from the last case.

However, if the support of μ_ν is the whole space $\mathbb{R}^d$, then the support of μ_X is also $\mathbb{R}^d$, regardless of the choice of A in any of the three cases above.

Like in the general MMA case, we shall again have a closer look at the essential condition

$$\int_{M_d^-} \frac{\kappa(A)^\alpha}{\rho(A)}\, \pi(dA) < \infty.$$

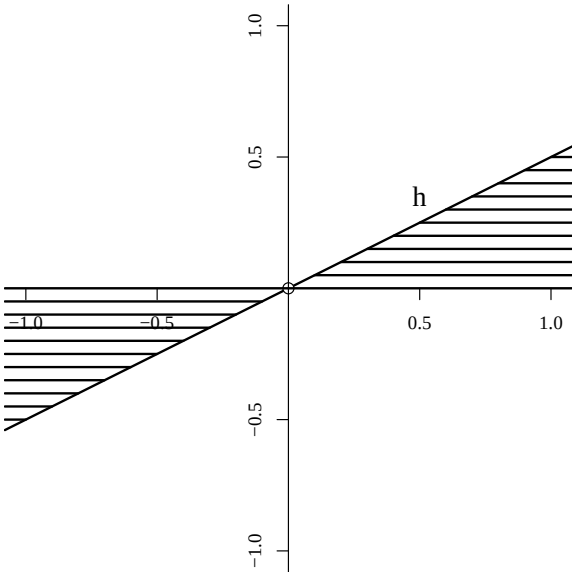


Figure 4.4.2: Mass of the measure μ_X in case *2*.

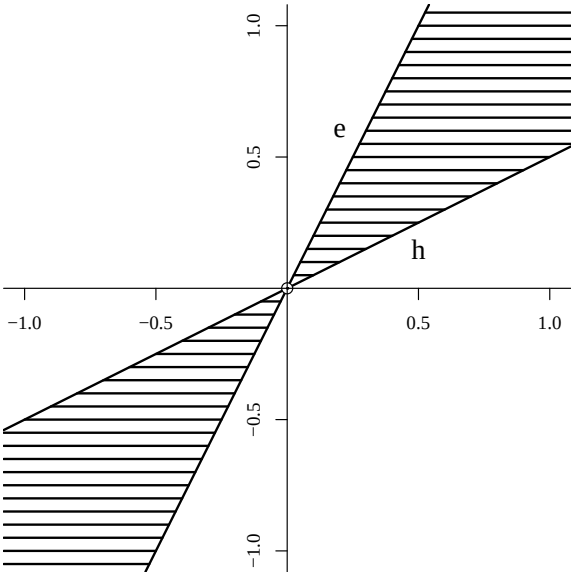


Figure 4.4.3: Mass of the measure μ_X in case *3*.

Using the modulus of injectivity, we can derive necessary conditions similar to the previous chapter, see also Barndorff-Nielsen and Stelzer (2011a), Proposition 3.3, where comparable necessary conditions are given for the existence of supOU processes.

Corollary 4.4.5. *Let $\Lambda \in \mathbb{R}^d$ be a Lévy basis on $M_d^- \times \mathbb{R}$ with generating quadruple $(\gamma, \Sigma, \nu, \pi)$, let $\nu \in RV(\alpha, l, \mu_\nu)$ and let $X_0 = \int_{M_d^-} \int_{\mathbb{R}^+} e^{As} \Lambda(dA, ds)$ exist following Theorem 4.4.2. Furthermore, assume there exist measurable functions $\tau : M_d^- \mapsto \mathbb{R}^+ \backslash \{0\}$ and $\vartheta : M_d^- \mapsto [1, \infty)$ such that*

$$j\left(e^{As}\right) \geq \vartheta(A) e^{-\tau(A)s} \; \forall s \in \mathbb{R}^+ \; \pi\text{-almost surely.}$$

Then

$$\int\limits_{M_d^-} \frac{\vartheta(A)^\alpha}{\tau(A)} \, \pi(dA) < \infty$$

is a necessary condition for $X_0 \in RV(\alpha, l, \mu_X)$ with

$$\mu_X(\cdot) := \int\limits_{M_d^-} \int\limits_{\mathbb{R}^+} \int\limits_{\mathbb{R}^d} \mathbb{1}_{(\cdot)}\left(e^{As}x\right) \mu_\nu(dx) ds \pi(dA).$$

Proof. Suppose

$$\int_{M_d^-} \frac{\vartheta(A)^\alpha}{\tau(A)}\,\pi(dA) = \infty.$$

Then

$$\begin{aligned}\int_{M_d^-}\int_{\mathbb{R}^+} j\left(e^{As}\right)^\alpha ds\pi(dA) &\geq \int_{M_d^-}\int_{\mathbb{R}^+} \vartheta(A)^\alpha e^{-\alpha\tau(A)s} ds\pi(dA)\\ &= \alpha^{-1}\int_{M_d^-}\frac{\vartheta(A)^\alpha}{\tau(A)}\pi(dA)\\ &= \infty.\end{aligned}$$

Consequently $e^{As} \notin \mathbb{J}^\alpha(\lambda\times\pi)$ and Theorem 4.3.4 yields the result.

□

Finally, as a consequence of Theorem 4.3.5, we also have regular variation of the finite dimensional distributions of the process.

Corollary 4.4.6. *Given the conditions of Corollary 4.4.3, the supOU process $(X_t)_{t\in\mathbb{R}}$ is also regularly varying with index α as a process.*

4.5 Stochastic Volatility Model

4.5.1 The Model

In this section we review and analyze the supOU type stochastic volatility model introduced in Barndorff-Nielsen and Stelzer (2011b). We consider a d-dimensional logarithmic stock price process $(X_t)_{t\in\mathbb{R}}$ given by an equation of the form

$$\begin{aligned} dX_t &= \Sigma_t^{1/2} dW_t \\ X_0 &= 0, \end{aligned} \tag{4.5.1}$$

where W is a d-dimensional Brownian motion and $\Sigma^{1/2}$ denotes the unique positive semidefinite square root. The stochastic volatility process $(\Sigma_t)_{t\in\mathbb{R}}$ is given by an $\mathbb{S}_d^+$-valued supOU process that is independent of the Brownian Motion W.

Definition 4.5.1 (Positive Semi-Definite SupOU Process). Let Λ be a Lévy basis on $M_d^-\times\mathbb{R}$ with values in $\mathbb{S}_d$. If the process

$$\Sigma_t := \int_{M_d^-}\int_{-\infty}^{t} e^{A(t-s)}\Lambda(dA,ds)e^{A^T(t-s)}$$

exists for all $t \in \mathbb{R}$, it is called a *positive semi-definite (or $\mathbb{S}_d^+$-valued) supOU process*.

The process $(X_t)_{t\in\mathbb{R}^+}$ being given by equation (4.5.1) with volatility process $(\Sigma_t)_{t\in\mathbb{R}}$ given by a positive semi-definite supOU process is called *multivariate supOU type stochastic volatility model* or *SVsupOU*.
The introduced model is of course only the most basic version of a SVsupOU. We can easily enhance the model by adding a stochastic or deterministic drift a and a leverage term Ψ, see Barndorff-Nielsen and Stelzer (2011b) for details. The model is then given by the equation

$$\begin{aligned} dX_t &= a_t dt + \Sigma_{t-}^{1/2} dW_t + \Psi(dL_t) \\ X_0 &= 0, \end{aligned}$$

where a is an $\mathbb{R}^d$-valued predictable process, W is the d-dimensional Brownian motion, L is the Lévy process associated with Λ and $\Psi : \mathbb{S}_d \mapsto \mathbb{R}^d$ is a linear operator. The stochastic volatility process $(\Sigma_t)_{t\in\mathbb{R}}$ is again a matrix-valued supOU process.
However, the drift term and the leverage term are usually dominating the tail behavior if they are non-vanishing. The leverage term is determined by the behavior of the Lévy process L and as we always assume the driving Lévy measure to be regularly varying with index α, the leverage term is also regularly varying with index α. A popular choice for the drift term is

$$a_t = \mu + \beta\Sigma_t$$

with $\beta : \mathbb{S}_d \mapsto \mathbb{R}^d$ being a linear operator and in this case a_t is regularly varying with index α, as we will show below. This means that if such a drift or leverage term exists, they dominate the Brownian term which will turn out to be regularly varying with index 2α. For this reason, we will only consider the simple model in this chapter.
Let us start with analyzing the volatility process. Existence of the positive semi-definite supOU processes is given similarly to the existence of $\mathbb{R}^d$-valued supOU processes.

Theorem 4.5.2 (Barndorff-Nielsen and Stelzer (2011a), Theorem. 4.1). *Let Λ be an $\mathbb{S}_d$-valued Lévy basis with generating quadruple $(\gamma, 0, \nu, \pi)$ and with $\gamma_0 := \gamma - \int_{\|x\|\leq 1} x\nu(dx) \in \mathbb{S}_d^+$, $\nu(\mathbb{S}_d \backslash \mathbb{S}_d^+) = 0$,*

$$\int_{\|x\|>1} \ln(\|x\|)\,\nu(dx) < \infty \quad \text{and} \quad \int_{\|x\|\leq 1} \|x\|\nu(dx) < \infty.$$

Furthermore, assume the existence of measurable functions $\rho : M_d^- \mapsto \mathbb{R}^+$ *and* $\kappa : M_d^- \mapsto [1, \infty)$ *such that*

$$\left\|e^{As}\right\| \leq \kappa(A)e^{-\rho(A)s} \; \forall s \in \mathbb{R}^+ \; \pi\text{-almost surely and } \int\limits_{M_d^-} \frac{\kappa(A)^2}{\rho(A)}\pi(dA) < \infty.$$

Then the positive-semidefinite supOU process

$$\Sigma_t = \int\limits_{M_d^-} \int\limits_{-\infty}^{t} e^{A(t-s)}\Lambda(dA, ds)e^{A^T(t-s)}$$

is well-defined for all $t \in \mathbb{R}$*, has values in* $\mathbb{S}_d^+$ *for all* $t \in \mathbb{R}$ *and its distribution is stationary and infinitely divisible. Moreover, the vector representation has the form*

$$vec\left(\Sigma_t\right) = \int\limits_{M_d^-} \int\limits_{-\infty}^{t} e^{(A\otimes I_d + I_d \otimes A)(t-s)} vec(\Lambda)(dA, ds)$$

and the distribution of Σ_t *is infinitely divisible with characteristic function*

$$\exp\left(i\mathrm{tr}(u\Sigma_t)\right) = \exp\left(i\mathrm{tr}\left(u\gamma_\Sigma\right) + \int\limits_{\mathbb{S}_d} \left(e^{i\mathrm{tr}(ux)} - 1\right)\nu_\Sigma(dx)\right)$$

for $u \in \mathbb{S}_d$ *with*

$$\gamma_\Sigma = \int\limits_{M_d^-} \int\limits_{0}^{\infty} e^{As}\,\gamma_0\,e^{A^T s}\,ds\pi(dA) \quad \text{and}$$

$$\nu_\Sigma(B) = \int\limits_{M_d^-} \int\limits_{0}^{\infty} \int\limits_{\mathbb{S}_d^+} \mathbb{1}_B\left(e^{As}xe^{A^T s}\right)\nu(dx)ds\pi(dA)$$

for all Borel sets $B \subseteq \mathbb{S}_d$*.*

Note that in the above theorem $vec(\Lambda)$ is defined by $vec(\Lambda)(A) := vec((\Lambda(A))$ and it is a Lévy basis in $\mathbb{R}^{d^2}$.

Based on this theorem, we can now analyze the tail behavior of the volatility process. Therefore, we have to define regular variation in a matrix-valued setting, which is just a translation of $\mathbb{R}^d$-valued regular variation. A random matrix $X \in M_d$ is said to be regularly

varying with index $\alpha > 0$ if there exists a slowly varying function $l : \mathbb{R} \mapsto \mathbb{R}$ and a nonzero Radon measure μ defined on $\mathcal{B}(\overline{M}_d \backslash \{0\})$ with $\mu(\overline{M}_d \backslash M_d) = 0$ such that, as $u \to \infty$,

$$u^\alpha l(u) P(u^{-1} X \in \cdot) \xrightarrow{v} \mu(\cdot)$$

on $\mathcal{B}(\overline{M}_d \backslash \{0\})$ and we write $X \in RV(\alpha, l, \mu)$. Of course, for a random matrix $X \in M_d$ there exists the straightforward connection that $X \in RV(\alpha, l, \mu)$ if and only if $vec(X) \in RV(\alpha, l, \mu^v)$, where $\mu^v(vec(A)) = \mu(A)$. Given this relationship, we can then analyze the tail behavior of the volatility process, where regular variation can be derived using the results of the previous sections.

Corollary 4.5.3. *Let $\Lambda \in \mathbb{S}_d$ be a Lévy basis on $M_d^- \times \mathbb{R}$ with generating quadruple $(\gamma, 0, \nu, \pi)$ and let $\nu \in RV(\alpha, l, \mu_\nu)$. If the conditions of Theorem 4.5.2 hold and additionally*

$$\int\limits_{M_d^-} \frac{\kappa(A)^{2\alpha}}{\rho(A)} \, \pi(dA) < \infty,$$

then $\Sigma_0 = \int_{M_d^-} \int_{\mathbb{R}^+} e^{As} \Lambda(dA, ds) e^{A^T s} \in RV(\alpha, l, \mu_\Sigma)$ with Radon measure

$$\mu_\Sigma(\cdot) := \int\limits_{M_d^-} \int\limits_{\mathbb{R}^+} \int\limits_{\mathbb{R}^d} \mathbb{1}_{(\cdot)} \left(e^{As} x e^{A^T s} \right) \mu_\nu(dx) ds \pi(dA).$$

Furthermore, the supOU process $(\Sigma_t)_{t \in \mathbb{R}}$ is also regularly varying with index α as a process.

Proof. From Theorem 4.5.2 we have

$$vec\,(\Sigma_t) = \int\limits_{M_d^-} \int\limits_{-\infty}^{t} e^{(A \otimes I_d + I_d \otimes A)(t-s)} vec(\Lambda)(dA, ds)$$

and thus the vectorized volatility process $vec(\Sigma_t)$ is an MMA process with kernel function $f(A, s) = e^{(A \otimes I_d + I_d \otimes A)s} \mathbb{1}_{[0,\infty)}(s)$. In order to apply Theorem 4.3.2, it is left to show that $f \in \mathbb{L}^\alpha(\lambda \times \pi)$. For that purpose, we make use of the relations $e^{(A \otimes I_d + I_d \otimes A)s} = e^{As} \otimes e^{As}$ and $\|e^{As} \otimes e^{As}\| = \|e^{As}\|^2$ (see Horn and Johnson (1991), Chapter 4.2, Problem 28, and Chapter 6.2, Problem 14) and obtain

$$\begin{aligned} \int\limits_{M_d^-} \int\limits_{\mathbb{R}^+} \|e^{(A \otimes I_d + I_d \otimes A)s}\|^\alpha ds \pi(dA) &\leq \int\limits_{M_d^-} \int\limits_{\mathbb{R}^+} \|e^{As}\|^{2\alpha} ds \pi(dA) \\ &\leq \int\limits_{M_d^-} \int\limits_{\mathbb{R}^+} \kappa(A)^{2\alpha} e^{-2\alpha\rho(A)s} ds \pi(dA) \end{aligned}$$

$$= \frac{1}{2\alpha} \int_{M_d^-} \frac{\kappa(A)^{2\alpha}}{\rho(A)} \pi(dA)$$
$$< \infty.$$

□

Now we want to go one step further and analyze the tail behavior of the logarithmic stock price process. Therefore, we use the independence between W and Λ yielding the equality

$$\int_0^t \Sigma_s^{1/2} dW_s \stackrel{d}{=} \left(\int_0^t \Sigma_s ds \right)^{1/2} W_1. \tag{4.5.2}$$

We immediately see that it is necessary to analyze the integrated volatility

$$\Sigma_t^+ := \int_0^t \Sigma_s ds$$

in order to obtain regular variation of the stock price process. We start with its existence.

Theorem 4.5.4 (Barndorff-Nielsen and Stelzer (2011a), Theorem 4.3). *Let Σ be a positive semi-definite supOU process as given in Definition 4.5.1 that exists according to Theorem 4.5.2. Then $\Sigma_t(\omega)$ is measurable as a function of $t \in \mathbb{R}$ and $\omega \in \Omega$. If also*

$$\int_{M_d^-} \kappa(A)^2 \pi(dA) < \infty,$$

then the paths of Σ are uniformly bounded in t and the integrated process Σ_t^+ exists for all $t \in \mathbb{R}^+$ and the equation

$$\Sigma_t^+ = \int_{M_d^-} \int_{-\infty}^t g_A^{-1} \left(e^{A(t-s)} \Lambda(dA, ds) e^{A^T(t-s)} \right) + \int_{M_d^-} \int_{-\infty}^0 g_A^{-1} \left(e^{-As} \Lambda(dA, ds) e^{-A^T s} \right)$$
$$- \int_{M_d^-} \int_0^t g_A^{-1} \left(\Lambda(dA, ds) \right)$$

holds with the function $g_A : \mathbb{S}_d \to \mathbb{S}_d, X \mapsto AX + XA^T$.

Another important and closely related characteristic of a time series are observed log returns over given time periods of length $\Delta \in \mathbb{R}^+$ (representing for example observation

intervals, trading periods etc.) given by

$$Z_n := X_{n\Delta} - X_{(n-1)\Delta} = \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \stackrel{d}{=} \left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s ds \right)^{1/2} W_1. \tag{4.5.3}$$

Existence of the related integrated volatilities

$$\Sigma_n^+ := \int_{(n-1)\Delta}^{n\Delta} \Sigma_s ds$$

is given by the previous theorem and conditions for regular variation of Σ_n^+ and of the integrated volatility Σ_t^+ can be derived simultaneously.

Corollary 4.5.5. *Let $\Lambda \in \mathbb{S}_d$ be a Lévy basis on $M_d^- \times \mathbb{R}$ with generating quadruple $(\gamma, 0, \nu, \pi)$ and let $\nu \in RV(\alpha, l, \mu_\nu)$. If the conditions of Theorem 4.5.4 hold and additionally*

$$\int_{M_d^-} \frac{\kappa(A)^{2\alpha}}{\rho(A)^{\alpha+1}} \pi(dA) < \infty,$$

then $\Sigma_n^+ \in RV(\alpha, l, \mu_{\Sigma_n^+})$ with Radon measure

$$\mu_{\Sigma_n^+}(B) := \int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \mathbb{1}_B \left(\int_{u \vee (n-1)\Delta}^{n\Delta} e^{A(s-u)} x e^{A^T(s-u)} \mathbb{1}_{(-\infty, n\Delta]}(u) ds \right) \mu_\nu(dx) du \pi(dA)$$

and $\Sigma_t^+ \in RV(\alpha, l, \mu_{\Sigma_t^+})$ with

$$\mu_{\Sigma_t^+}(B) := \int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \mathbb{1}_B \left(\int_{u \vee 0}^{t} e^{A(s-u)} x e^{A^T(s-u)} \mathbb{1}_{(-\infty, t]}(u) ds \right) \mu_\nu(dx) du \pi(dA).$$

Furthermore, the process $(\Sigma_t^+)_{t \in \mathbb{R}^+}$ is also regularly varying with index α as a process.

Proof. Again, we use the vector representation of the process and get from the proof of Theorem 3.12 in Barndorff-Nielsen and Stelzer (2011a)

$$vec\left(\Sigma_n^+\right) = \int_{(n-1)\Delta}^{n\Delta} \int_{M_d^-} \int_{-\infty}^{s} e^{(A \otimes I_d + I_d \otimes A)(s-u)} vec(\Lambda)(dA, du) ds$$

$$= \int_{\mathbb{R}} \int_{M_d^-} g(A,t,u) vec(\Lambda)(dA,du)$$

with

$$g(A,t,u) := \int_{u\vee(n-1)\Delta}^{n\Delta} e^{(A\otimes I_d + I_d \otimes A)(s-u)} \mathbb{1}_{(-\infty,n\Delta]}(u) ds.$$

As before, it is left to show that $g(A,t,u) \in \mathbb{L}^\alpha(\lambda \times \pi)$ in order to apply Theorem 4.3.2. Therefore, we estimate

$$\begin{aligned}
\|g(A,t,u)\| &\leq \mathbb{1}_{(-\infty,n\Delta]}(u) \int_{u\vee(n-1)\Delta}^{n\Delta} \left\| e^{(A\otimes I_d + I_d \otimes A)(s-u)} \right\| ds \\
&\leq \mathbb{1}_{(-\infty,n\Delta]}(u) \int_{u\vee(n-1)\Delta}^{n\Delta} \left\| e^{A(s-u)} \right\|^2 ds \\
&\leq \mathbb{1}_{(-\infty,n\Delta]}(u) \int_{u\vee(n-1)\Delta}^{n\Delta} \kappa(A)^2 e^{-2\rho(A)(s-u)} ds \\
&= \frac{\kappa(A)^2}{-2\rho(A)} \Big(e^{-2\rho(A)(n\Delta-u)} \mathbb{1}_{(-\infty,n\Delta]}(u) - \mathbb{1}_{((n-1)\Delta,n\Delta)}(u) \\
&\qquad - e^{-2\rho(A)((n-1)\Delta-u)} \mathbb{1}_{(-\infty,(n-1)\Delta]}(u) \Big).
\end{aligned}$$

For the first term of the sum we get

$$\begin{aligned}
\int_{M_d^-} \int_{-\infty}^{n\Delta} \left| \frac{\kappa(A)^2 e^{-2\rho(A)(n\Delta-u)}}{-2\rho(A)} \right|^\alpha du\, \pi(dA) &= \int_{M_d^-} \int_{-\infty}^{n\Delta} \frac{\kappa(A)^{2\alpha} e^{-2\alpha\rho(A)(n\Delta-u)}}{2^\alpha \rho(A)^\alpha} du\, \pi(dA) \\
&= \frac{1}{2^{\alpha+1}\alpha} \int_{M_d^-} \frac{\kappa(A)^{2\alpha}}{\rho(A)^{\alpha+1}} \pi(dA) \\
&< \infty.
\end{aligned}$$

The second summand is in $\mathbb{L}^\alpha(\lambda \times \pi)$, since the function has bounded support, and for the last term in the sum we simply substitute n by $(n-1)$ in the first term. The result for Σ_t^+ follows directly setting $\Delta = t$ and $n = 1$.

□

Note that this result also gives us regular variation with index α of a possible drift term $a_t = \mu + \beta\Sigma_t$.

The next step is to derive the tail behavior of the square root $(\Sigma^+)^{1/2}$ of the integrated volatility process.

Lemma 4.5.6. *Let Σ be a random variable with values in $\mathbb{S}_d^+$ and let $\Sigma^{1/2}$ be its square root. Then $\Sigma \in RV(\alpha, l, \mu_\Sigma)$ if and only if $\Sigma^{1/2} \in RV(2\alpha, l^{1/2}, \mu_\Sigma^{1/2})$ with $l^{1/2}(x) := l(x^2)$ and $\mu_\Sigma^{1/2}(B) := \mu(B^2)$.*

Proof. Note that the square root of a matrix in $\mathbb{S}_d^+$ is a bijective mapping and is thus well defined. Since both functions, the square as well as the square root, map compacts to compacts, we can apply Proposition 3.18 of Resnick (1987).

□

Now we can consider the log-returns and the logarithmic stock price process regarding their tail behavior.

Theorem 4.5.7. *Let $(X_t)_{t\in\mathbb{R}}$ be the stock price process given by equation (4.5.1), let Z_n be the log-returns given by (4.5.3) and let Σ_n^+ be the increments of a positive semi-definite supOU process $(\Sigma_t)_{t\in\mathbb{R}^+}$. Furthermore, let the conditions of Corollary 4.5.5 hold. Then $Z_n \in RV(2\alpha, l^{1/2}, \mu_Z)$ with Radon measure*

$$\mu_Z(B) := \mathbb{E}\left(\mu_{\Sigma_n^+}^{1/2}\left(W_1^{-1}(B)\right)\right)$$

and $X_t \in RV(2\alpha, l^{1/2}, \mu_X)$ with

$$\mu_X(B) := \mathbb{E}\left(\mu_{\Sigma_t^+}^{1/2}\left(W_1^{-1}(B)\right)\right),$$

where $W_1 : M_d^- \mapsto \mathbb{R}^d$ is considered to be a random linear mapping with $W_1(x) := x \cdot W_1 \overset{d}{=} xN(0, I_d)$. Furthermore, $(X_t)_{t\in\mathbb{R}}$ is also regular varying with index 2α as a process.

Proof. Since W and Λ are independent, we have

$$Z_n \overset{d}{=} \left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s ds\right)^{1/2} W_\Delta \quad \text{and} \quad \int_0^t \Sigma_s^{1/2} dW_s \overset{d}{=} \left(\int_0^t \Sigma_s ds\right)^{1/2} W_t.$$

From Corollary 4.5.5 we know that

$$\Sigma_n^+ \in RV(\alpha, l, \mu_{\Sigma_n^+}) \quad \text{and} \quad \Sigma_t^+ \in RV(\alpha, l, \mu_{\Sigma_t^+})$$

and, together with Lemma 4.5.6, this yields

$$(\Sigma_n^+)^{1/2} \in RV(2\alpha, l^{1/2}, \mu_{\Sigma_n^+}^{1/2}) \quad \text{and} \quad (\Sigma_t^+)^{1/2} \in RV(2\alpha, l^{1/2}, \mu_{\Sigma_t^+}^{1/2}).$$

Finally, we use the fact that the Brownian Motion has finite moments to apply the multivariate version of Breiman's Lemma (see Theorem 2.1.18), which yields the result.

□

4.5.2 Relevance and Applications in Finance

Let us conclude with some final remarks. First, we easily see that the model allows for heavy tails, in the volatility as well as in logarithmic stock prices and log-returns. This is a useful fact, since observed market data often shows heavy tails. Furthermore, we see that there is a direct connection between the indexes of regular variation of the driving Lévy measure on the one hand and the volatility, log-prices and log-returns on the other hand. We can also calculate the concrete measure μ of regular variation in order to describe the spatial structure of the extremes.
Second, we note that all the results given above hold also in the case of an Ornstein-Uhlenbeck type stochastic volatility model where the volatility is modeled by an $\mathbb{S}_d^+$-valued OU process. This is obvious, since OU processes are special cases of supOU processes with π being a Dirac measure.
In a financial context, the results can now be used for a statistical analysis of observed data. We can use one of the well established estimators (see Embrechts et al. (1997) or Resnick (2007)) to estimate the index of regular variation of the given data (logarithmic stock prices or log-returns). The result can then be compared with the estimated index of regular variation of the integrated volatility. If they do not match by the factor of 2, this is a hint for the existence of a leverage or drift term of the form specified in this chapter. Yet, there is still some future work to be done to analyze and estimate the (index of regular variation of the) integrated volatility, since it cannot be observed directly. If we make additional assumptions on the different terms (leverage, drift) to exist or not, we can calculate the index of regular variation of the log-prices or log-returns from the index of the volatility and vice versa.
It would also be very interesting to generalize the stochastic volatility model by substituting the Brownian motion W_t by a more general Lévy process $\widetilde{L}_t$. However, as there is then no analogue of (4.5.2) available, it will be much more difficult to get results for this case and different methods will be needed.

Modelling the Correlation Breakdown

Applied research in financial mathematics and econometrics has often noted that one typically encounters what has been dubbed "correlation breakdown" in times of severe crisis. This notion means that when extreme negative events potentially affecting the whole (or large parts of the) economy occur, basically all traded stocks are loosing tremendously in value simultaneously and the correlations between them are seemingly more or less one.

Moreover, after such an event the variances are typically extremely high. Models employed in financial institutions (for risk management) clearly need to include this feature in order to be realistic and provide accurate predictions.

The results of this chapter on the (sup)OU stochastic volatility model allow us to understand how to incorporate such effects into the model. Clearly, the most extreme movements in crisis can typically not come from the Brownian term but have to come from jumps. Hence, Ψ needs to be chosen non-zero and such that when all variances have a jump upwards all prices have jumps downwards. Assume now that the positive semidefinite Lévy basis is regularly varying with index α and the measure μ_ν is concentrated on the rank one matrices with all diagonal elements being non-zero and all correlations being 1. The extremes of Σ are now caused by a single big jump which will be such that it is (almost) a rank one matrix with all diagonal elements being non-zero. After the occurrence of the jump the process Σ will be almost equal to the value of this jump and hence all correlations will be very close to one for quite some time afterwards. Moreover, by the choice of parameters the prices will simultaneously have a huge jump downward. Clearly, this would model the "correlation breakdown". The results actually show that Σ would be regularly varying with index α and μ_Σ would be concentrated on the rank one matrices with all diagonal elements being non-zero as this class of matrices is preserved by the mappings $X \mapsto e^{As} X e^{A^T s}$ for all $A \in M_d(\mathbb{R})$ and $s \in \mathbb{R}^+$. Likewise the log prices would be regularly varying with index α (unless we had a drift with heavier tails which seems not reasonable) and the measure of regular variation for the log prices follows easily, because what matters is only the linear transformation Ψ of the driving positive semidefinite Lévy basis. Note that the measure of regular variation of the log prices will in general still have a completely non-degenerate support (in the positive d-dimensional cone).

In practice the above explained model can only form an important building block of a realistic and suitable model, since not all extreme events affect the whole economy, some only affect individual sectors of industry or companies. However, also for such extensions (which basically demand regular variation of Λ with appropriate μ_Λ) the results of this chapter provide the necessary insight into the resulting tail behavior.

Chapter 5

Functional Regular Variation of Lévy-Driven Multivariate Mixed Moving Average Processes

In Chapter 4 we have introduced multivariate mixed moving average (MMA) processes and given conditions for regular variation of their finite-dimensional distributions if the driving Lévy basis is regularly varying. In this chapter, we extend the results to functional regular variation in the space $\mathbb{D}$ of càdlàg functions. We give sufficient conditions for an MMA process (X_t) to have càdlàg sample paths. Furthermore, we prove that (X_t) is regularly varying in $\mathbb{D}$ if the finite-dimensional distributions are regularly varying and the kernel function f satisfies certain conditions. The conditions are also validated in the special case of supOU processes.

5.1 Introduction

In many applications of stochastic processes, the center of the distributions involved and related quantities (e.g. sample means, variances etc.) can be modeled quite well. In view of the central limit theorem, Gaussian distributions play an important role in that field. However, this needs not to be true for the tail of the distribution which is of great importance in many areas of application. Possible examples are severe crisis in stock markets or extreme weather events which can cause huge losses to the financial industry, insurance companies and also to private people. Therefore, it is of great importance to model the distribution tail and related quantities (e.g. quantiles, exceedances, maxima etc.) correctly.

A very well established concept to model extreme values is regular variation. It has

its origin in classical extreme value theory, where limit distributions for sample maxima are derived. The maximum domains of attraction of two of the three possible standard extreme value distributions (Fréchet and Weibull) can be described by regular variation of functions, meaning functions behaving like a power law in the limit, see also Embrechts et al. (1997) and Resnick (2007).

Moreover, regular variation can intuitively be extended to a multivariate setup. It is then formulated in terms of vague convergence of measures given by

$$nP(a_n^{-1}X \in \cdot) \xrightarrow{v} \mu(\cdot), \tag{5.1.1}$$

where X is a multivariate random vector, (a_n) an increasing sequence and μ is a Radon measure. Since μ is homogeneous, multivariate regular variation of X can be interpreted as convergence of the radial part $\|X\|$ to a univariate regularly varying random variable Y and of the spherical part $X/\|X\|$ to a random variable Z on the unit sphere, which is independent of Y and can be described by the measure μ. Detailed introductions to multivariate regular variation can be found in Resnick (2007) and Hult and Lindskog (2006b).

Finally, Hult and Lindskog (2005) extended the definition (5.1.1) to the space of multivariate stochastic processes with sample paths in the space $\mathbb{D}$ of càdlàg functions, i.e. right-continuous functions with limits from the left. The formulation of regular variation in such generality has the advantage that, in addition to functionals based on the values of a stochastic process at fixed time points, one can also analyze functionals acting on the complete sample paths of the process. This is a very powerful tool for the analysis of extremal properties of a process, especially in combination with methods for weak convergence of point processes which are closely related to regular variation (see Section 5.6).

In this chapter we apply the concept of regular variation on $\mathbb{D}$ to multivariate mixed moving average (MMA) processes with càdlàg sample paths. MMA processes have been first introduced by Surgailis et al. (1993) in the univariate stable case and are given as integrals of the form

$$X_t = \int_{M_d^-} \int_{\mathbb{R}} f(A, t-s)\Lambda(dA, ds),$$

where Λ is a multivariate Lévy basis. The class of (multivariate) MMA processes covers a wide range of processes which are well known and extensively used in applications. Examples include Ornstein-Uhlenbeck processes (cf. Barndorff-Nielsen and Shephard (2001) and Pigorsch and Stelzer (2009)), superpositions of Ornstein-Uhlenbeck (supOU) processes

(cf. Barndorff-Nielsen (2001) and Barndorff-Nielsen and Stelzer (2011a)), (fractionally integrated) CARMA processes (cf. Brockwell (2004), Marquardt (2007)) and increments of fractional Lévy processes (cf. Marquardt (2006), Bender et al. (2011) and references therein).

Regular variation of the finite-dimensional distributions of MMA processes has already been proved in Section 4.3, given that the underlying Lévy basis is regularly varying and the kernel function satisfies the integrability condition $f \in \mathbb{L}^\alpha$. In this chapter we give additional integrability and continuity conditions on the kernel function f such that the MMA process is functionally regularly varying on $\mathbb{D}$. Furthermore, we also analyze the special case of multivariate supOU processes given by

$$X_t = \int_{M_d^-} \int_{-\infty}^{t} e^{A(t-s)} \Lambda(dA, ds)$$

and give some more accessible sufficient conditions for supOU processes to be functionally regularly varying.

The chapter is organized as follows. In Section 5.2.2 we introduce the notion of multivariate regular variation and related properties. In Section 5.2.3 we recall the definition of MMA processes and the related integration theory. Furthermore, we review the conditions for existence of MMA processes and for regular variation of their finite dimensional distributions. The sample path behavior of MMA processes is discussed in Section 5.3. We give an overview over the relevant literature and derive sufficient conditions for MMA processes to have càdlàg sample paths in the case where the driving Lévy process is of finite variation. In Section 5.4 we introduce the notion of functional regular variation and prove that MMA processes are regularly varying on $\mathbb{D}$, given certain conditions. In Section 5.5 we verify these conditions in the special case of supOU processes. Finally, in Section 5.6 we show the connection between functional regular variation and point process convergence and discuss the relevance of the results to the extremal analysis of MMA processes.

5.2 Preliminaries

5.2.1 Notation

Let $\mathbb{R}$ be the real numbers, $\mathbb{R}^+$ the positive and $\mathbb{R}^-$ the negative real numbers, both without 0. $\mathbb{N}$ is the set of positive integers and $\mathbb{Q}$ are the rational numbers. The Borel sets are denoted by $\mathcal{B}$ and $\mathcal{B}_b$ are the bounded Borel sets. λ is the Lebesgue measure on

$\mathbb{R}$ and $B_r(x) := \{y \in \mathbb{R}^d : \|y - x\| \leq r\}$ is the closed ball of radius r centered at x. $\mathbb{D}$ is the space of càdlàg (right-continuous with left limits) functions $x : [0,1] \to \mathbb{R}^d$ and $S_{\mathbb{D}} = \{x \in \mathbb{D} : \sup_{t\in[0,1]} \|x_t\| = 1\}$ is the unit sphere in $\mathbb{D}$.

For matrices, $M_{n,d}$ is the set of all $n \times d$ matrices and M_d the set of all $d \times d$ matrices. M_d^- is the set of all $d \times d$ matrices with all eigenvalues having strictly negative real part. I_d is the $d \times d$ identity matrix. We write A^T for the transposed of a matrix A and $\|A\|$ for the matrix norm induced by the Euclidean norm.

If random variables, vectors, processes, measures etc. are considered, they are given as measurable mappings with respect to a complete probability space $(\Omega, \mathcal{F}, P)$.

Vague convergence is defined in terms of convergence of Radon measures and it is denoted by $\xrightarrow{v}$. It is defined on the one-point uncompactification $\overline{\mathbb{R}}^d\backslash\{0\}$, which assures that the sets bounded away from the origin can be referred to as the relatively compact sets in the vague topology. Similarly, $\hat{w}$-convergence is given by the convergence of boundedly finite measures and is defined on $\overline{\mathbb{D}}_0 = (0, \infty] \times S_{\mathbb{D}}$, which can be viewed as the one-point uncompactification in $\mathbb{D}$.

5.2.2 Multivariate Regular Variation

The tail behavior of multivariate random variables is often described by the property of regular variation. This concept is derived from the notion of regularly varying functions, i.e. functions behaving like a power law in the limit, and has an intuitive extension to the space of real-valued and multivariate random variables. Regular variation on $\mathbb{R}^d$ is expressed in terms of vague convergence of measures and several different, but equivalent, definitions exist. For a detailed introduction we refer to Section 2.1. Very good overviews on regular variation can also be found in Bingham et al. (1987), Resnick (1987), Resnick (2007) and Lindskog (2004).

Definition 5.2.1 (Multivariate Regular Variation). A random vector $X \in \mathbb{R}^d$ is *regularly varying* if there exists a sequence $(a_n)_{n\in\mathbb{N}}$, $0 < a_n \nearrow \infty$, and a nonzero Radon measure μ on $\mathcal{B}(\overline{\mathbb{R}}^d\backslash\{0\})$ such that $\mu(\overline{\mathbb{R}}^d\backslash\mathbb{R}^d) = 0$ and, as $n \to \infty$,

$$nP(a_n^{-1}X \in \cdot) \xrightarrow{v} \mu(\cdot)$$

on $\mathcal{B}(\overline{\mathbb{R}}^d\backslash\{0\})$. Similarly, we call a Radon measure ν regularly varying if (a_n) and μ exist as above such that, as $n \to \infty$,

$$n\,\nu(a_n^{-1}\cdot) \xrightarrow{v} \mu(\cdot).$$

The limiting measure μ of the definition is homogeneous, i.e. it necessarily satisfies the condition

$$\mu(tB) = t^{-\alpha}\mu(B)$$

for all $B \in \mathcal{B}(\overline{\mathbb{R}}^d \backslash \{0\})$ and $t > 0$. Hence, we write $X \in RV(\alpha, (a_n), \mu)$ or $\nu \in RV(\alpha, (a_n), \mu)$ respectively. In the special case of an infinitely divisible random vector $X \in \mathbb{R}^d$ with Lévy measure ν we know that $X \in RV(\alpha, (a_n), \mu)$ if and only if $\nu \in RV(\alpha, (a_n), \mu)$ (see Theorem 2.3.10). This result is very useful throughout this chapter, since MMA processes are infinitely divisible, just as the driving Lévy bases are. Detailed introductions to infinitely divisible distributions and Lévy processes can be found in Sato (2002) and Section 2.3.

5.2.3 Multivariate Mixed Moving Average Processes

In this section we shortly recall the definition and main properties of multivariate *mixed moving average processes* (short MMA processes) which have already been introduced and analyzed in Section 4.

A multivariate ($\mathbb{R}^n$-valued) MMA process (X_t) can be defined as an integral over a measurable kernel function $f : M_d^- \times \mathbb{R} \mapsto M_{n,d}$ with respect to an $\mathbb{R}^d$-valued Lévy basis Λ on $M_d^- \times \mathbb{R}$, i.e.

$$X_t := \int_{M_d^-} \int_{\mathbb{R}} f(A, t-s)\Lambda(dA, ds).$$

An $\mathbb{R}^d$-valued *Lévy basis* $\Lambda = (\Lambda(B))$ with $B \in \mathcal{B}_b(M_d^- \times \mathbb{R})$ is a random measure which is

- *infinitely divisible*, i.e. the distribution of $\Lambda(B)$ is infinitely divisible for all $B \in \mathcal{B}_b(M_d^- \times \mathbb{R})$,
- *independently scattered*, i.e. for any $n \in \mathbb{N}$ the random variables $\Lambda(B_1), \ldots, \Lambda(B_n)$ are independent for pairwise disjoint sets $B_1, \ldots, B_n \in \mathcal{B}_b(M_d^- \times \mathbb{R})$ and
- *σ-additive*, i.e. for any pairwise disjoint sets $(B_i)_{i\in\mathbb{N}} \in \mathcal{B}_b(M_d^- \times \mathbb{R})$ with $\bigcup_{n\in\mathbb{N}} B_n \in \mathcal{B}_b(M_d^- \times \mathbb{R})$ we have $\Lambda(\bigcup_{n\in\mathbb{N}} B_n) = \sum_{n\in\mathbb{N}} \Lambda(B_n)$ almost surely.

Thus, Lévy bases are also called infinitely divisible independently scattered random measures (i.d.i.s.r.m.). Following Section 4.2.3 and Remark 4.2.5 we only consider time-homogeneous and factorisable Lévy bases, i.e. Lévy bases with characteristic function

$$\mathbb{E}\left(e^{iu^T\Lambda(B)}\right) = e^{\varphi(u)\Pi(B)} \tag{5.2.1}$$

for all $u \in \mathbb{R}^d$ and $B \in \mathcal{B}_b(M_d^- \times \mathbb{R})$, where $\Pi = \lambda \times \pi$ is the product of a probability measure π on $M_d^-(\mathbb{R})$ and the Lebesgue measure λ on $\mathbb{R}$ and

$$\varphi(u) = iu^T\gamma - \frac{1}{2}u^T\Sigma u + \int_{\mathbb{R}^d} \left(e^{iu^Tx} - 1 - iu^Tx\mathbb{1}_{[-1,1]}(\|x\|)\right)\nu(dx)$$

is the characteristic function of an infinitely divisible distribution with characteristic triplet (γ, Σ, ν). The distribution of the Lévy basis is then completely determined by $(\gamma, \Sigma, \nu, \pi)$ which is therefore called the *generating quadruple*. By L we denote the *underlying Lévy process* which is given by $L_t = \Lambda(M_d^- \times (0,t])$ and $L_{-t} = \Lambda(M_d^- \times [-t,0))$ for $t \in \mathbb{R}^+$. For more details on Lévy bases see Rajput and Rosiński (1989) and Pedersen (2003).

As in Section 4.3, we also mention that the set M_d^- in the definition of MMA processes can be replaced by M_d or basically any other Borel set of matrices. The choice of M_d^- is motivated by the special case of supOU processes, where this is the canonical choice.

Necessary and sufficient conditions for the existence of MMA processes are given by the multivariate extension of Theorem 2.7 in Rajput and Rosiński (1989).

Theorem 5.2.2. *Let Λ be an $\mathbb{R}^d$-valued Lévy basis with characteristic function of the form (4.2.1) and let $f : M_d^- \times \mathbb{R} \mapsto M_{n,d}$ be a measurable function. Then f is Λ-integrable as a limit in probability in the sense of Rajput and Rosiński (1989) if and only if*

$$\int_{M_d^-}\int_{\mathbb{R}} \left\| f(A,s)\gamma + \int_{\mathbb{R}^d} f(A,s)x\left(\mathbb{1}_{[0,1]}(\|f(A,s)x\|) - \mathbb{1}_{[0,1]}(\|x\|)\right)\nu(dx)\right\| ds\pi(dA) < \infty, \quad (5.2.2)$$

$$\int_{M_d^-}\int_{\mathbb{R}} \|f(A,s)\Sigma\, f(A,s)^T\| ds\pi(dA) < \infty \quad \text{and} \quad (5.2.3)$$

$$\int_{M_d^-}\int_{\mathbb{R}}\int_{\mathbb{R}^d} \left(1 \wedge \|f(A,s)x\|^2\right)\nu(dx)ds\pi(dA) < \infty. \quad (5.2.4)$$

If f is Λ-integrable, the distribution of $X_0 = \int_{M_d^-}\int_{\mathbb{R}^+} f(A,s)\Lambda(dA,ds)$ is infinitely divisible with characteristic triplet $(\gamma_{int}, \Sigma_{int}, \nu_{int})$ given by

$$\gamma_{int} = \int_{M_d^-}\int_{\mathbb{R}} \left(f(A,s)\gamma + \int_{\mathbb{R}^d} f(A,s)x\left(\mathbb{1}_{[0,1]}(\|f(A,s)x\|) - \mathbb{1}_{[0,1]}(\|x\|)\right)\nu(dx)\right) ds\pi(dA),$$

$$\Sigma_{int} = \int_{M_d^-}\int_{\mathbb{R}} f(A,s)\Sigma\, f(A,s)^T ds\pi(dA) \quad \text{and}$$

$$\nu_{int}(B) = \int_{M_d^-}\int_{\mathbb{R}}\int_{\mathbb{R}^d} \mathbb{1}_B(f(A,s)x)\nu(dx)ds\pi(dA) \quad \text{for all Borel sets } B \subseteq \mathbb{R}^d.$$

However, since the focus of this section is on regularly varying processes, we will also use some more convenient conditions for this special setting. They have been derived in Theorem 4.2.7 and are based on integrability conditions described by the set

$$\mathbb{L}^\delta(\lambda \times \pi) := \left\{ f : M_d^- \times \mathbb{R} \mapsto M_{n,d} \text{ measurable}, \int_{M_d^-} \int_{\mathbb{R}} \|f(A,s)\|^\delta ds\pi(dA) < \infty \right\}.$$

Theorem 5.2.3. *Let Λ be a Lévy basis with values in $\mathbb{R}^d$ and characteristic quadruple $(\gamma, \Sigma, \nu, \pi)$. Furthermore, let ν be regularly varying with index α and let $f : M_d^- \times \mathbb{R} \mapsto M_{n,d}$ be a measurable function. Then f is Λ-integrable in the sense of Rajput and Rosiński (1989) and X_t is well defined for all $t \in \mathbb{R}$, stationary and infinitely divisible with known characteristic triplet if one of the following conditions is satisfied:*

(i) L_1 is α-stable with $\alpha \in (0,2)\backslash\{1\}$ and $f \in \mathbb{L}^\alpha \cap \mathbb{L}^1$.

(ii) f is bounded and $f \in \mathbb{L}^\delta$ for some $\delta < \alpha$, $\delta \leq 1$.

(iii) f is bounded, $\mathbb{E}\, L_1 = 0$, $\alpha > 1$ and $f \in \mathbb{L}^\delta$ for some $\delta < \alpha$, $\delta \leq 2$.

Regular Variation of X_t for fixed $t \in \mathbb{R}$ as well as regular variation of the finite dimensional distributions of the process (X_t) have been derived in Theorem 4.3.2 and Corollary 4.3.5 under similar conditions.

Theorem 5.2.4. *Let Λ be an $\mathbb{R}^d$-valued Lévy basis on $M_d^- \times \mathbb{R}$ with generating quadruple $(\gamma, \Sigma, \nu, \pi)$ and let $\nu \in RV(\alpha, (a_n), \mu_\nu)$. If $X_0 = \int_{M_d^-} \int_{\mathbb{R}} f(A,s)\Lambda(dA, ds)$ exists (in the sense of Theorem 5.2.2), $f \in \mathbb{L}^\alpha(\lambda \times \pi)$ and $\mu_\nu(f^{-1}(A,s)(\mathbb{R}^n\backslash\{0\})) = 0$ does not hold for $\pi \times \lambda$ almost-every (A,s), then $X_0 \in RV(\alpha, (a_n), \mu_X)$ with*

$$\mu_X(B) := \int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \mathbb{1}_B\left(f(A,s)x\right) \mu_\nu(dx) ds\pi(dA).$$

Furthermore, the finite dimensional distributions $(X_{t_1}, \ldots, X_{t_k})$, $t_i \in \mathbb{R}$ and $k \in \mathbb{N}$, are also regularly varying with index α and a given limiting measure $\mu_{t_1,\ldots,t_k}$.

Comparable necessary conditions for regular variation do also exist, see Theorem 4.3.4 for details.

Next we introduce a result which allows to decompose a Lévy basis into a drift, a Brownian part, a part with bounded jumps and a part with finite variation. This is the extension of the Lévy-Itô decomposition of Theorem 2.3.6 to Lévy bases.

Theorem 5.2.5 (Barndorff-Nielsen and Stelzer (2011a), Theorem 2.2). *Let Λ be a Lévy basis on $M_d^- \times \mathbb{R}$ with characteristic function of the form* (5.2.1) *and generating quadruple $(\gamma, \Sigma, \nu, \pi)$. Then there exists a modification $\widetilde{\Lambda}$ of Λ which is also a Lévy basis with the same characteristic quadruple $(\gamma, \Sigma, \nu, \pi)$ such that there exists an $\mathbb{R}^d$-valued Lévy basis $\widetilde{\Lambda}^G$ on $M_d^- \times \mathbb{R}$ with generating quadruple $(0, \Sigma, 0, \pi)$ and an independent Poisson random measure N on $\mathbb{R}^d \times M_d^- \times \mathbb{R}$ with intensity measure $\nu \times \pi \times \lambda$ such that*

$$\widetilde{\Lambda}(B) = \gamma(\pi\times\lambda)(B) + \widetilde{\Lambda}^G(B) + \int_{\|x\|\leq 1} \int_B x(N(dx, dA, ds) - \pi(dA)ds\nu(dx)) + \int_{\|x\|>1} \int_B xN(dx, dA, ds)$$

for all $B \in \mathcal{B}_b(M_d^- \times \mathbb{R})$ and $\omega \in \Omega$. If, additionally, $\int_{\|x\|\leq 1} \|x\|\nu(dx) < \infty$, then

$$\widetilde{\Lambda}(B) = \gamma_0(\pi \times \lambda)(B) + \widetilde{\Lambda}^G(B) + \int_{\mathbb{R}^d} \int_B xN(dx, dA, ds)$$

for all $B \in \mathcal{B}_b(M_d^- \times \mathbb{R})$, where

$$\gamma_0 := \gamma - \int_{\|x\|\leq 1} x\nu(dx).$$

Moreover, the Lebesgue integral exists with respect to N for all $\omega \in \Omega$.

In the case of an underlying Lévy process with finite variation, Theorem 5.2.5 and Theorem 5.2.2 can be combined to obtain integrability conditions for this special setting. Note that by Theorem 2.3.7 finite variation of (L_t) is equivalent to $\Sigma = 0$ and $\int_{\|x\|\leq 1} \|x\| \, \nu(dx) < \infty$.

Proposition 5.2.6 (Barndorff-Nielsen and Stelzer (2011a), Prop. 2.4). *Let Λ be a Lévy basis on $M_d^- \times \mathbb{R}$ with characteristic function of the form* (5.2.1) *and generating quadruple $(\gamma, 0, \nu, \pi)$ such that $\int_{\|x\|\leq 1} \|x\| \, \nu(dx) < \infty$. Let γ_0 and N be as defined in Theorem 5.2.5. If $f \in \mathbb{L}^1$ and*

$$\int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} (1 \wedge \|f(A, s)x\|) \, \nu(dx) \, ds \, \pi(dA) < \infty,$$

then

$$X = \int_{M_d^-} \int_{\mathbb{R}} f(A, s)\Lambda(dA, ds) = \int_{M_d^-} \int_{\mathbb{R}} f(A, s) \, \gamma_0 \, ds\pi(dA) + \int_{\mathbb{R}^d} \int_{M_d^-} \int_{\mathbb{R}} f(A, s) \, x \, N(dx, dA, ds)$$

and the integrals on the right hand side exist as Lebesgue integrals for every $\omega \in \Omega$. Moreover, the distribution of X is infinitely divisible with characteristic function

$$\mathbb{E}\left(e^{iu^T X}\right) = \exp\left(iu^T \gamma_{int,0} + \int_{\mathbb{R}^d} \left(e^{iu^T x} - 1\right) \nu_{int}(dx)\right),$$

where

$$\gamma_{int,0} = \int_{M_d^-} \int_{\mathbb{R}} f(A,s)\, \gamma_0 \, ds \, \pi(dA) \quad \textit{and}$$

$$\nu_{int}(B) = \int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \mathbb{1}_B(f(A,s)x)\, \nu(dx)\, ds\, \pi(dA)$$

for all Borel sets $B \subseteq \mathbb{R}^d$.

5.3 Sample Path Behavior

In Section 5.4 we review the concept of regular variation for càdlàg processes and apply it to MMA processes. Therefore, we first have to discuss the sample path behavior of MMA processes.

Many examples of results for MMA processes to have càdlàg sample paths exist in the special case where the underlying Lévy process is of finite variation, i.e. $\Sigma = 0$ and $\int_{\|x\|\leq 1} \|x\| \, \nu(dx) < \infty$. In this case, the sample path behavior of the driving Lévy process transfers to the sample paths of the MMA process. For example, define for any Lévy process L_t the corresponding *filtered Lévy process* X_t by

$$X_t = \int_0^t f(t,s)\, dL_s \tag{5.3.1}$$

for $t \in [0,1]$. If X_t exists, L_t is of finite variation and the kernel function f is bounded and continuous, then X_t has càdlàg sample paths (cf. Hult and Lindskog (2005), Lemma 28).

A more general result for supOU processes is given by Theorem 3.12 in Barndorff-Nielsen and Stelzer (2011a). This result can be extended to the general case of MMA processes. For that reason, we introduce the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}}$, where for $t \in \mathbb{R}$ the σ-algebra $\mathcal{F}_t$ is generated by the sets

$$\left\{\Lambda(B) : B \in \mathcal{B}(M_d^- \times (-\infty, t])\right\}.$$

That way it is ensured that the paths $X_t(\omega)$, understood as a function of $t \in \mathbb{R}$ and $\omega \in \Omega$, are adapted to the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}}$.

Theorem 5.3.1. *Let Λ be a Lévy basis on $M_d^- \times \mathbb{R}$ with characteristic function of the form (5.2.1) and generating quadruple $(\gamma, 0, \nu, \pi)$ such that $\int_{\|x\|\leq 1} \|x\| \, \nu(dx) < \infty$. Suppose that the kernel function $f(A, s)$ is continuous and differentiable in s for all $s \in \mathbb{R}\backslash\{0\}$ and $f(A, 0^-) = \lim_{s\nearrow 0} f(A, s) = C_1 \in M_{n,d}$ as well as $f(A, 0^+) = \lim_{s\searrow 0} f(A, s) = C_2 \in M_{n,d}$ for all $A \in M_d^-$. Set*

$$f'(A, s) := \begin{cases} \frac{d}{ds} f(A, s) & \text{if } s \neq 0, \\ \lim_{s\searrow 0} \frac{d}{ds} f(A, s) & \text{if } s = 0 \end{cases}$$

and assume that for some $\delta > 0$ and for every $t_1, t_2 \in \mathbb{R}$ such that $t_1 \leq t_2$ and $t_2 - t_1 \leq \delta$ the function $\sup_{t\in[t_1,t_2]} \|f'(A, t - s)\|$ satisfies the conditions of Proposition 5.2.6, where $(\gamma, 0, \nu, \pi)$ is replaced by $(\|\gamma\|, 0, \nu_T, \pi)$ and the Lévy measure $\nu_T(\cdot) = \nu(T^{-1}(\cdot))$ is transformed by $T(x) = \|x\|$. If the processes $X_t = \int_{M_d^-} \int_{\mathbb{R}} f(A, t - s)\Lambda(dA, ds)$ and

$$Z_t := \int_{M_d^-} \int_{\mathbb{R}} f'(A, t - s)\Lambda(dA, ds)$$

exist (in the sense of Proposition 5.2.6), then

$$X_t = X_0 + \int_0^t Z_u \, du + (C_1 - C_2) \, L_t$$

and consequently X_t has sample paths in $\mathbb{D}$ which are of finite variation on compacts.

Proof. We follow the proof of Theorem 3.12 in Barndorff-Nielsen and Stelzer (2011a) and begin by showing that Z_t is locally uniformly bounded on compacts. Note that by Proposition 5.2.6 the processes X_t and Z_t can be given as integrals with respect to a Poisson measure and $\pi \times \lambda$. For $\delta > 0$ and every $t_1, t_2 \in \mathbb{R}$ such that $t_1 \leq t_2$ and $t_2 - t_1 \leq \delta$ we obtain

$$\begin{aligned} \sup_{t\in[t_1,t_2]} \|Z_t\| &= \sup_{t\in[t_1,t_2]} \left\| \int_{M_d^-} \int_{\mathbb{R}} f'(A, t - s)\Lambda(dA, ds) \right\| \\ &\leq \int_{M_d^-} \int_{\mathbb{R}} \sup_{t\in[t_1,t_2]} \|f'(A, t - s)\| \Lambda_T(dA, ds), \end{aligned}$$

where $T : \mathbb{R}^d \to \mathbb{R}$ is given by $T(x) = \|x\|$ and Λ_T is the transformed Lévy basis with characteristic triplet $(\|\gamma\|, 0, \nu_T, \pi)$. Existence of the right hand side is covered by Proposition 5.2.6. Thus Z_t is locally uniformly bounded and it follows by Fubini that

$$\int_0^t Z_u du = \int_0^t \int_{M_d^-} \int_{-\infty}^u f'(A, u - s)\Lambda(dA, ds) \, du + \int_0^t \int_{M_d^-} \int_u^\infty f'(A, u - s)\Lambda(dA, ds) \, du$$

$$
\begin{aligned}
&= \int_{M_d^-} \int_{-\infty}^{t} \int_{0\vee s}^{t} f'(A, u-s)\, du\, \Lambda(dA, ds) + \int_{M_d^-} \int_{0}^{\infty} \int_{0}^{t\wedge s} f'(A, u-s)\, du\, \Lambda(dA, ds) \\
&= \int_{M_d^-} \int_{-\infty}^{t} f(A, u-s)\Big|_{u=0\vee s}^{t} \Lambda(dA, ds) + \int_{M_d^-} \int_{0}^{\infty} f(A, u-s)\Big|_{u=0}^{t\wedge s} \Lambda(dA, ds) \\
&= \int_{M_d^-} \int_{-\infty}^{t} f(A, t-s)\Lambda(dA, ds) - \int_{M_d^-} \int_{-\infty}^{0} f(A, 0-s)\Lambda(dA, ds) \\
&\quad - \int_{M_d^-} \int_{0}^{t} f(A, 0^+)\Lambda(dA, ds) + \int_{M_d^-} \int_{t}^{\infty} f(A, t-s)\Lambda(dA, ds) \\
&\quad + \int_{M_d^-} \int_{0}^{t} f(A, 0^-)\Lambda(dA, ds) - \int_{M_d^-} \int_{0}^{\infty} f(A, 0-s)\Lambda(dA, ds) \\
&= X_t - X_0 + (C_1 - C_2)\, L_t.
\end{aligned}
$$

□

Remark 5.3.2. The inclusion of kernel functions with a discontinuity at $s = 0$ is motivated by the class of causal MMA processes where the kernel function is of the form $f(A, s)\mathbb{1}_{[0,\infty)}(s)$. For example, in the supOU case the kernel function is $e^{As}\,\mathbb{1}_{[0,\infty)}(s)$ and the limits at $s = 0$ can be given directly by $C_1 = \mathbf{0}$ and $C_2 = I_d$ yielding

$$X_t = X_0 + \int_0^t Z_u\, du - L_t.$$

This coincides with the result of Theorem 3.12 in Barndorff-Nielsen and Stelzer (2011a), where the existence conditions on X_t and Z_t as well as the integrability condition on $\sup_{t\in[t_1,t_2]} \|f'(A, t-s)\|$ are expressed in terms of the special properties of supOU processes with driving Lévy process of finite variation.

If $C_1 - C_2 = 0$ in the above theorem, further properties of the sample paths of X_t follow directly.

Corollary 5.3.3. *Assume that the conditions of Theorem 5.3.1 hold. If additionally $C_1 = C_2$, then the paths of $X_t = \int_{M_d^-} \int_{\mathbb{R}} f(A, t-s)\Lambda(dA, ds)$ are absolutely continuous and almost surely differentiable.*

Remark 5.3.4. The condition $C_1 = C_2$ holds if and only if $f(A, s)$ is continuous in $s = 0$ and $f(A, 0)$ is constant for all $A \in M_d^-$. This is satisfied, for example, by two-sided supOU processes which are MMA processes with kernel function

$$f(A, s) = e^{As}\mathbb{1}_{[0,\infty)}(s) + e^{-As}\mathbb{1}_{(-\infty,0)}(s).$$

In that case $C_1 = C_2 = I_d$. In the case of moving average processes, where π is a one-point measure, the condition only requires that f is continuous in $s = 0$. Processes of this class include, for example, two-sided CARMA and two-sided Ornstein-Uhlenbeck processes.

Similar results for the sample paths of MMA processes, where the driving Lévy process is not of finite variation, are in general not so easy to obtain. Basse and Pedersen (2009), Corollary 3.3, give necessary and sufficient conditions for filtered Lévy processes of the form (5.3.1) to have càdlàg sample paths of bounded variation even if the driving Lévy process itself has sample paths of unbounded variation. Furthermore, they also study *two-sided moving averages* of the form

$$X_t = \int_{-\infty}^{t} (f_1(t-s) - f_2(-s))\, dL_t,$$

where $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ are measurable kernel functions such that $f_1(s) = f_2(s) = 0$ for all $s \in (-\infty, 0)$. They give necessary and sufficient conditions for such processes to have càdlàg sample paths of finite variation. These conditions also allow the underlying Lévy process to be of infinite variation. Moreover, they also consider the special case where the driving Lévy process is symmetric α-stable with $\alpha \in (1, 2]$ (cf. Basse and Pedersen (2009), Lemma 5.2, Proposition 5.3 and Proposition 5.5). Conditions for α-stable MMA processes, $\alpha \in (0, 2)$, to have càdlàg sample paths are also given in Basse and Rosiński (2011), Section 4.

Additionally, there also exist some results for the stronger property of continuous sample paths. See Marcus and Rosiński (2005), Cambanis et al. (1990) and Rosiński (1989) for results on general MMA processes to have continuous sample paths. For the special case of α-stable MMA processes, see also Rosiński et al. (1991) and Rosiński (1986).

5.4 Functional Regular Variation

Following Hult and Lindskog (2005), regular Variation on $\mathbb{D}$ is given in terms of the $\hat{w}$-convergence of boundedly finite measures on $\overline{\mathbb{D}}_0$. A measure μ on a complete separable

metric space $\mathbb{E}$ is said to be *boundedly finite* if $\mu(B) < \infty$ for every bounded set $B \in \mathcal{B}(\mathbb{E})$. Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of boundedly finite measures on $\mathbb{E}$. Then (μ_n) converges to μ in the *$\hat{w}$-topology* if $\mu_n(B) \to \mu(B)$ for all bounded Borel sets $B \in \mathcal{B}(\mathbb{E})$ with $\mu(\partial B) = 0$. We write $\mu_n \xrightarrow{\hat{w}} \mu$. A complete introduction of the used topology can be found in Section 2.1.4 or Hult and Lindskog (2005). See also Daley and Vere-Jones (1988) and Kallenberg (1983) for details on $\hat{w}$-convergence and vague convergence.

Definition 5.4.1 (Regular Variation on $\mathbb{D}$). A stochastic process (X_t), $t \in [0,1]$, with sample paths in $\mathbb{D}$ is said to be *regularly varying* if there exists a positive sequence (a_n), $n \in \mathbb{N}$, with $a_n \nearrow \infty$ and a nonzero boundedly finite measure μ on $\mathcal{B}(\overline{\mathbb{D}}_0)$ with $\mu(\overline{\mathbb{D}}_0 \backslash \mathbb{D}) = 0$ such that, as $n \to \infty$,

$$nP(a_n^{-1}X \in \cdot) \xrightarrow{\hat{w}} \mu(\cdot) \quad \text{ on } \mathcal{B}(\overline{\mathbb{D}}_0).$$

As in the finite dimensional case, direct calculation shows that the measure μ is homogeneous, i.e. there exists a positive index $\alpha > 0$ such that $\mu(uB) = u^{-\alpha}\mu(B)$ for all $u > 0$ and for every $B \in \mathcal{B}(\overline{\mathbb{D}}_0)$. Thus, we say that the process (X_t) is *regularly varying with index* α and write $X \in RV_{\overline{\mathbb{D}}_0}(\alpha, (a_n), \mu)$.

Analogously to multivariate regular variation, several alternative definitions of regular variation on $\mathbb{D}$ exist, we will just state one example here.

Theorem 5.4.2 (Hult and Lindskog (2005), Theorem 4). *A process (X_t) with sample paths in $\mathbb{D}$ is regularly varying if and only if there exists an index $\alpha > 0$ and a probability measure σ on $\mathcal{B}(S_{\mathbb{D}})$ such that for every positive $x > 0$, as $u \to \infty$,*

$$\frac{P(\|X\|_\infty > ux, X/\|X\|_\infty \in \cdot)}{P(\|X\|_\infty > u)} \xrightarrow{\hat{w}} x^{-\alpha}\sigma(\cdot) \quad \text{ on } \mathcal{B}(S_{\mathbb{D}}),$$

where $\|X\|_\infty = \sup_{t\in[0,1]} \|X_t\|$.

The probability measure σ is called the *spectral measure* of X.

Example 5.4.3 (Lévy Process). Let (L_t) be a Lévy process. Then by definition (or Theorem 2.3.2 resp.) (L_t) has sample paths in $\mathbb{D}$. Furthermore, (L_t) is also a strong Markov process (cf. Sato (2002), Theorem 10.5 and Corollary 40.11). Now the results of Hult and Lindskog (2005), Section 3, can be applied. If $L_t \in RV(\alpha, (a_n), t\mu)$, then it follows by Theorem 13 of Hult and Lindskog (2005) that $(L_t) \in RV_{\overline{\mathbb{D}}_0}(\alpha, (a_n), \widetilde{\mu})$ for some measure $\widetilde{\mu}$. For details we refer to Hult and Lindskog (2005), Example 17.

We will now recall some useful results from Hult and Lindskog (2005) related to regular variation on $\mathbb{D}$. Since it is often of interest, how the regular variation property is preserved

under mappings, we look at a continuous mapping theorem. Therefore, for any function h from a metric space E to a metric space E' we introduce the set $disc(h)$ which consists of all discontinuities of h.

Theorem 5.4.4 (Hult and Lindskog (2005), Theorem 6). *Let (X_t) be a stochastic process with sample paths in $\mathbb{D}$ and let E' be a complete separable metric space. Assume that $X \in RV_{\overline{\mathbb{D}}_0}(\alpha, (a_n), \mu)$ and $h : \mathbb{D} \to E'$ is a measurable mapping such that $\mu(disc(h)) = 0$ and $h^{-1}(B)$ is bounded in $\overline{\mathbb{D}}_0$ for every bounded $B \in \mathcal{B}(E')$. Then, as $n \to \infty$,*

$$nP(h(a_n^{-1}X) \in \cdot) \xrightarrow{\hat{w}} \mu \circ h^{-1}(\cdot) \quad \text{on } \mathcal{B}(E').$$

There also exists a different version of the previous theorem for the special case of positively homogeneous mappings of order $\gamma > 0$, i.e. mappings $h : \mathbb{D} \to \mathbb{D}$ with $h(\lambda x) = \lambda^\gamma h(x)$ for all $\lambda \geq 0$ and $x \in \mathbb{D}$. See Theorem 2.1.21 for details.

The next theorem states some necessary and sufficient conditions for regular variation on $\mathbb{D}$. In the theorem, we use the notation

$$w(x, T_0) := \sup_{t_1, t_2 \in T_0} \|x_{t_1} - x_{t_2}\| \quad \text{and}$$
$$w''(x, \delta) := \sup_{t_1 \leq t \leq t_2;\, t_2 - t_1 \leq \delta} \min\{\|x_t - x_{t_1}\|, \|x_{t_2} - x_t\|\}$$

for $x \in \mathbb{D}$, $T_0 \subseteq [0,1]$ and $\delta \in [0,1]$.

Theorem 5.4.5 (Hult and Lindskog (2005), Theorem 10). *Let (X_t) be a stochastic process with sample paths in $\mathbb{D}$. Then the following statements are equivalent.*

(i) $X \in RV_{\overline{\mathbb{D}}_0}(\alpha, (a_n), \mu)$.

(ii) There exists a set $T \subseteq [0,1]$ containing 0, 1 and all but at most countably many points of $[0,1]$, a positive sequence $a_n \nearrow \infty$ and a collection $\{\mu_{t_1,\ldots,t_k} : t_i \in T, k \in \mathbb{N}\}$ of Radon measures on $\mathcal{B}(\overline{\mathbb{R}}^{dk} \backslash \{0\})$ with $\mu_{t_1,\ldots,t_k}(\overline{\mathbb{R}}^{dk} \backslash \mathbb{R}^{dk}) = 0$ and μ_t is nonzero for some $t \in T$ such that

$$nP(a_n^{-1}(X_{t_1}, \ldots, X_{t_k}) \in \cdot) \xrightarrow{v} \mu_{t_1,\ldots,t_k}(\cdot) \quad \text{on } \mathcal{B}(\overline{\mathbb{R}}^{dk} \backslash \{0\}) \tag{5.4.1}$$

holds for all $t_1, \ldots, t_k \in T$. Furthermore, for every $\varepsilon, \eta > 0$, there exist $\delta \in (0,1)$ and $n_0 \in \mathbb{N}$ such that, for all $n \geq n_0$,

$$nP(a_n^{-1} w(X, [0, \delta)) \geq \varepsilon) \leq \eta, \tag{5.4.2}$$
$$nP(a_n^{-1} w(X, [1-\delta, 1)) \geq \varepsilon) \leq \eta, \tag{5.4.3}$$
$$nP(a_n^{-1} w''(X, \delta) \geq \varepsilon) \leq \eta. \tag{5.4.4}$$

Remark 5.4.6. The theorem links regular variation of the process $(X_t)_{t\in[0,1]}$ with sample paths in $\mathbb{D}$ to regular variation of the finite dimensional distributions $(X_{t_1}, \dots, X_{t_k})$ of the the process. Key to that connection are the relative compactness criteria (5.4.2), (5.4.3) and (5.4.4) which restrict the oscillation of the process (X_t) in small areas. See Hult and Lindskog (2005), Example 11, for a process satisfying conditions (5.4.2) and (5.4.3), but not (5.4.4).

Now we will extend the finite dimensional regular variation of MMA processes in the sense of Theorem 5.2.4 to regular variation in $\mathbb{D}$ by applying Theorem 5.4.5. Therefore, we need to restrict the MMA process (X_t) as defined in Section 5.2.3 to the time interval $[0, 1]$. Note that a restriction to any other compact time interval $[a, b]$, $a < b$, would not change any of the results. Furthermore, we assume that (X_t) has sample paths in the space $\mathbb{D}$ of càdlàg functions. See Section 5.3 for possible conditions ensuring this. We start with the main theorem for functional regular variation of MMA processes.

Theorem 5.4.7. *Let Λ and Λ_2 be $\mathbb{R}^d$-valued Lévy bases on $M_d^- \times \mathbb{R}$ with generating quadruples $(\gamma, \Sigma, \nu, \pi)$ and $(0, 0, \nu|_{B_1(0)^c}, \pi)$ respectively such that $\nu \in RV(\alpha, (a_n), \mu_\nu)$. Assume that the kernel function $f(A, s)$ is bounded, $f \in \mathbb{L}^\alpha(\lambda \times \pi)$, $\mu_\nu(f^{-1}(A, s)(\mathbb{R}^n \backslash \{0\})) = 0$ does not hold for $\pi \times \lambda$ almost-every (A, s) and*

$$\int_{M_d^-} \int_{\mathbb{R}} \int_{\|x\|>1} (1 \wedge \|f(A, s)x\|)\, \nu(dx)\, ds\, \pi(dA) < \infty. \tag{5.4.5}$$

Moreover, suppose that the MMA process $X_t = \int_{M_d^-} \int_{\mathbb{R}} f(A, t - s)\Lambda(dA, ds)$ exists for $t \in [0, 1]$ (in the sense of Theorem 5.2.2) and that the processes X_t and $X_t^{(2)} = \int_{M_d^-} \int_{\mathbb{R}} f(A, t - s)\Lambda_2(dA, ds)$ have càdlàg sample paths. If the function f_δ given by

$$f_\delta(A, s) := \sup_{0\le t_1\le t_2\le 1;\ t_2-t_1\le\delta} \|f(A, t_2 - s) - f(A, t_1 - s)\|\, \mathbb{1}_{(t_1,t_2]^c}(s) \tag{5.4.6}$$

satisfies (5.4.5) and, as $\delta \to 0$,

$$\int_{M_d^-} \int_{\mathbb{R}} f_\delta(A, s)^\alpha\, ds\pi(dA) \to 0, \tag{5.4.7}$$

then

$$(X_t)_{t\in[0,1]} \in RV_{\overline{\mathbb{D}}_0}(\alpha, (a_n), \mu),$$

where μ is uniquely determined by the measures $\mu_{t_1,\dots,t_k}$ in Theorem 5.2.4.

The condition $\int_{M_d^-} \int_{\mathbb{R}} f_\delta(A,s)^\alpha \, ds\pi(dA) \to 0$ is closely linked to the behavior of the function f in small areas. It restricts the amplitude of jumps and continuous oscillations for arbitrarily small values of δ.

Lemma 5.4.8. *Let π be a probability measure and $f : M_d^- \times \mathbb{R} \mapsto M_{n,d}$ be a measurable kernel function. Assume that the function $f_\delta(A,s)$ given by (5.4.6) satisfies $f_\delta \in \mathbb{L}^\alpha$ for some $\delta > 0$. Then*

$$\lim_{\delta \to 0} \int_{M_d^-} \int_{\mathbb{R}} f_\delta(A,s)^\alpha \, ds\pi(dA) = 0$$

if and only if $\lim_{\delta \to 0} f_\delta(A,s) \to 0$ for $\pi \times \lambda$ almost every (A,s).

Proof. Let $f_\delta(A,s) \to 0$ for $\pi \times \lambda$ almost every (A,s). Then

$$\int_{M_d^-} \int_{\mathbb{R}} f_\delta(A,s)^\alpha \, ds\pi(dA) \to 0$$

follows by dominated convergence and the assumption $f_\delta \in \mathbb{L}^\alpha$ for some $\delta > 0$. On the other hand, suppose that the set

$$\tilde{B} := \left\{ (A,s) \in \mathcal{B}_b(M_d^- \times \mathbb{R}) : f_\delta(A,s) \to 0 \text{ as } \delta \to 0 \right\}$$

satisfies $\pi \times \lambda \left(\tilde{B}^c\right) = C > 0$. Then the monotonicity of f_δ in δ implies $\lim_{\delta \to 0} f_\delta(A,s) > 0$ for every $(A,s) \in \tilde{B}^c$ and thus

$$\lim_{\delta \to 0} \int_{M_d^-} \int_{\mathbb{R}} f_\delta(A,s)^\alpha \, ds\pi(dA) > 0.$$

□

Remark 5.4.9. From the definition of f_δ we see that the condition $\lim_{\delta \to 0} f_\delta(A,s) \to 0$ for $\pi \times \lambda$ almost every (A,s) is equivalent to the kernel function $f(A,s)$ being continuous in s for all $s \in \mathbb{R} \backslash \{0\}$. Now we also see the importance of the restriction $\mathbb{1}_{(t_1,t_2]^c}(s)$ in the definition of f_δ because it allows for $f(A,s)$ being discontinuous at $s = 0$. Without such a restriction, condition (5.4.7) would be violated by many examples of the class of causal MMA processes which have a kernel function of the type $f(A,s)\mathbb{1}_{[0,\infty)}(s)$. Causal MMA processes with $f(A,0) \neq 0$ include CARMA and supOU processes as well as many other well-known examples of MMA processes.

We can also give sufficient conditions for a general function $f : M_d^- \times \mathbb{R} \to M_{n,d}$ to satisfy condition (5.4.5).

Lemma 5.4.10. *Let $f : M_d^- \times \mathbb{R} \to M_{n,d}$ be a measurable function. Then condition* (5.4.5) *holds if one of the following two conditions are satisfied:*

(i) $f \in \mathbb{L}^1$ *and* $\alpha > 1$.

(ii) $f \in \mathbb{L}^{\alpha-\varepsilon}$ *for one* $\varepsilon \in (0, \alpha)$ *and* $\alpha \leq 1$.

Proof. For (i) we calculate

$$\begin{aligned}
&\int_{M_d^-} \int_{\mathbb{R}} \int_{\|x\|>1} (1 \wedge \|f(A,s)x\|)\, \nu(dx)\, ds\, \pi(dA) \leq \\
&\qquad \leq \int_{M_d^-} \int_{\mathbb{R}} \int_{\|x\|>1} \|f(A,s)\| \|x\|\, \nu(dx)\, ds\, \pi(dA) \\
&\qquad = \int_{M_d^-} \int_{\mathbb{R}} \|f(A,s)\|\, ds\, \pi(dA) \int_{\|x\|>1} \|x\|\, \nu(dx) \\
&\qquad < \infty
\end{aligned}$$

by Sato (2002), Corollary 25.8, and similarly for (ii) we obtain

$$\begin{aligned}
&\int_{M_d^-} \int_{\mathbb{R}} \int_{\|x\|>1} (1 \wedge \|f(A,s)x\|)\, \nu(dx)\, ds\, \pi(dA) \leq \\
&\qquad \leq \int_{M_d^-} \int_{\mathbb{R}} \int_{\|x\|>1} (1 \wedge \|f(A,s)x\|^{\alpha-\varepsilon})\, \nu(dx)\, ds\, \pi(dA) \\
&\qquad \leq \int_{M_d^-} \int_{\mathbb{R}} \int_{\|x\|>1} \|f(A,s)\|^{\alpha-\varepsilon} \|x\|^{\alpha-\varepsilon}\, \nu(dx)\, ds\, \pi(dA) \\
&\qquad = \int_{M_d^-} \int_{\mathbb{R}} \|f(A,s)\|^{\alpha-\varepsilon}\, ds\, \pi(dA) \int_{\|x\|>1} \|x\|^{\alpha-\varepsilon}\, \nu(dx) \\
&\qquad < \infty.
\end{aligned}$$

□

Remark 5.4.11. The conditions of Lemma 5.4.10 are only sufficient, not necessary, similar to the ones of Theorem 5.2.3. Thus in general we will only demand the weaker condition (5.4.5) which is also one of the existence conditions for MMA processes with driving Lévy process of finite variation in Proposition 5.2.6. Furthermore, we see from Lemma 5.4.10(i) that condition (5.4.5) in Proposition 5.2.6 can be dropped if $\alpha > 1$.

Proof of Theorem 5.4.7

Let (X_t) be an MMA process as given in Theorem 5.4.7, i.e. (X_t) exists for $t \in [0,1]$ (in the sense of Theorem 5.2.2), the kernel function f is bounded by $C \in \mathbb{R}^+$ and the regular variation conditions of Theorem 5.2.4 hold. Then there exists a positive sequence $a_n \nearrow \infty$ and a collection $\{\mu_{t_1,\dots,t_k} : t_i \in T, k \in \mathbb{N}\}$ of Radon measures on $\mathcal{B}(\overline{\mathbb{R}}^{dk}\backslash\{0\})$ with $\mu_{t_1,\dots,t_k}(\overline{\mathbb{R}}^{dk}\backslash\mathbb{R}^{dk}) = 0$ and μ_t is nonzero for some $t \in T$ such that

$$nP(a_n^{-1}(X_{t_1},\dots,X_{t_k}) \in \cdot) \xrightarrow{v} \mu_{t_1,\dots,t_k}(\cdot) \quad \text{on } \mathcal{B}(\overline{\mathbb{R}}^{dk}\backslash\{0\}).$$

Applying Theorem 5.4.5, it is left to show that the conditions (5.4.2), (5.4.3) and (5.4.4) hold.

Using the Lévy-Itô decomposition we have two independent Lévy bases Λ_1 and Λ_2 such that $\Lambda = \Lambda_1 + \Lambda_2$, Λ_1 has generating quadruple $(\gamma, \Sigma, \nu_1, \pi)$ and Λ_2 has generating quadruple $(0, 0, \nu_2, \pi)$, where $\nu_1 = \nu|_{B_1(0)}$ and $\nu_2 = \nu|_{B_1(0)^c}$. This yields

$$X_t = X_t^{(1)} + X_t^{(2)}, \tag{5.4.8}$$

where

$$X_t^{(1)} = \int_{M_d^-} \int_{\mathbb{R}} f(A, t-s)\Lambda_1(dA, ds) \tag{5.4.9}$$

and

$$X_t^{(2)} = \int_{M_d^-} \int_{\mathbb{R}} f(A, t-s)\Lambda_2(dA, ds). \tag{5.4.10}$$

Note that the term $X_t^{(2)}$ can be written in the form

$$X_t^{(2)} = \int_{\|x\|\geq 1} \int_{M_d^-} \int_{\mathbb{R}} f(A, t-s)\, x\, N(dx, dA, ds),$$

where N is a Poisson random measure with mean measure $\nu \times \pi \times \lambda$. Before we proceed, we need to ensure the existence of $X_t^{(1)}$ and $X_t^{(2)}$. Therefore, we give conditions for ω-wise existence of $X_t^{(2)}$ as a Lebesgue integral. Then the existence of $X_t^{(1)} = X_t - X_t^{(2)}$ follows from the existence of X_t and $X_t^{(2)}$.

Proposition 5.4.12. *Let $X_t^{(2)}$ be the process given by* (5.4.10)*, where Λ_2 is a Lévy basis with generating quadruple $(0, 0, \nu_2, \pi)$. If*

$$\int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} (1 \wedge \|f(A, s)x\|)\, \nu_2(dx)\, ds\, \pi(dA) < \infty,$$

then $X_t^{(2)}$ exists as a Lebesgue integral for all $\omega \in \Omega$.

Proof. By definition, $X_t^{(2)}$ has no Gaussian component and $\int_{\|x\|\leq 1} \|x\|\nu_2(dx) = 0$ and thus we have an underlying Lévy process of finite variation. Now the result follows as a special case of Proposition 5.2.6, where the condition $f \in \mathbb{L}^1$ is obsolete due to the absence of a drift.

□

Like for X_t, we also assumed that $X_t^{(2)}$ has càdlàg sample paths. Then also $X_t^{(1)} = X_t - X_t^{(2)}$ has càdlàg sample paths. Appropriate conditions for MMA processes to have càdlàg sample paths have been given in Section 5.3.

Now we continue the proof of Theorem 5.4.7 by verifying the relative compactness conditions (5.4.2), (5.4.3) and (5.4.4). For the first condition (5.4.2) we obtain

$$\sup_{t_1,t_2\in[0,\delta)} \|X_{t_1} - X_{t_2}\| \leq \sup_{t_1,t_2\in[0,\delta)} \|X_{t_1}^{(1)} - X_{t_2}^{(1)}\| + \sup_{t_1,t_2\in[0,\delta)} \|X_{t_1}^{(2)} - X_{t_2}^{(2)}\|$$

and hence

$$\begin{aligned} &nP\Big(a_n^{-1} \sup_{t_1,t_2\in[0,\delta)} \|X_{t_1} - X_{t_2}\| \geq \varepsilon\Big) \leq \\ &\quad \leq nP\Big(a_n^{-1} \sup_{t_1,t_2\in[0,\delta)} \|X_{t_1}^{(1)} - X_{t_2}^{(1)}\| \geq \varepsilon/2\Big) + nP\Big(a_n^{-1} \sup_{t_1,t_2\in[0,\delta)} \|X_{t_1}^{(2)} - X_{t_2}^{(2)}\| \geq \varepsilon/2\Big). \end{aligned}$$

The analogue result for the second condition (5.4.3) can be obtained likewise. For the third condition (5.4.4) we estimate

$$\begin{aligned} &\sup_{t_1\leq t\leq t_2;\, t_2-t_1\leq\delta} \min\Big\{\|X_{t_2} - X_t\|, \|X_t - X_{t_1}\|\Big\} \leq \\ &\quad \leq \sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \|X_{t_1}^{(1)} - X_{t_2}^{(1)}\| + \sup_{t_1\leq t\leq t_2;\, t_2-t_1\leq\delta} \min\Big\{\|X_{t_2}^{(2)} - X_t^{(2)}\|, \|X_t^{(2)} - X_{t_1}^{(2)}\|\Big\}, \end{aligned}$$

and

$$\begin{aligned} &nP\Big(a_n^{-1} \sup_{t_1\leq t\leq t_2;\, t_2-t_1\leq\delta} \min\Big\{\|X_{t_2} - X_t\|, \|X_t - X_{t_1}\|\Big\} \geq \varepsilon\Big) \leq \\ &\qquad \leq nP\Big(a_n^{-1} \sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \|X_{t_1}^{(1)} - X_{t_2}^{(1)}\| \geq \varepsilon/2\Big) \\ &\qquad\quad + nP\Big(a_n^{-1} \sup_{t_1\leq t\leq t_2;\, t_2-t_1\leq\delta} \min\Big\{\|X_{t_2}^{(2)} - X_t^{(2)}\|, \|X_t^{(2)} - X_{t_1}^{(2)}\|\Big\} \geq \varepsilon/2\Big). \end{aligned}$$

For every $\varepsilon, \eta > 0$ we have to show that there exists $n_0 \in \mathbb{N}$ and $\delta > 0$ such that for $n \geq n_0$ these quantities can be bounded by η. Regarding the quantities based on $X_t^{(1)}$ we observe

$$nP\Big(a_n^{-1} \sup_{t_1,t_2\in[0,\delta)} \|X_{t_1}^{(1)} - X_{t_2}^{(1)}\| \geq \varepsilon/2\Big) \leq nP\Big(a_n^{-1} \sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \|X_{t_1}^{(1)} - X_{t_2}^{(1)}\| \geq \varepsilon/2\Big)$$

and for (5.4.3)

$$nP\Big(a_n^{-1} \sup_{t_1,t_2\in[1-\delta,1)} \|X_{t_1}^{(1)} - X_{t_2}^{(1)}\| \geq \varepsilon/2\Big) \leq nP\Big(a_n^{-1} \sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \|X_{t_1}^{(1)} - X_{t_2}^{(1)}\| \geq \varepsilon/2\Big)$$

and thus it is sufficient to prove the bound only for the right hand side of the inequality.

Proposition 5.4.13. *Let Λ_1 be the $\mathbb{R}^d$-valued Lévy basis on $M_d^- \times \mathbb{R}$ determined by the generating quadruple $(\gamma, \Sigma, \nu_1, \pi)$, where $\nu_1 = \nu|_{B_1(0)}$. Assume that the kernel function f is bounded, that the MMA process $X_t^{(1)}$ given by (5.4.9) exists for $t \in [0,1]$ and that $X_t^{(1)}$ has càdlàg sample paths. Moreover, suppose that $\nu \in RV(\alpha, (a_n), \mu_\nu)$. Then $X_t^{(1)}$ satisfies*

$$\lim_{n\to\infty} nP(a_n^{-1} \sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \|X_{t_1}^{(1)} - X_{t_2}^{(1)}\| \geq \varepsilon) = 0$$

for all $\delta \in (0,1)$ and $\varepsilon > 0$.

Proof. We start by observing that $X_t^{(1)}$ is càdlàg and thus also separable and hence we can estimate

$$\sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \|X_{t_1}^{(1)} - X_{t_2}^{(1)}\| \leq 2 \sup_{t\in[0,1]} \|X_t^{(1)}\| = 2 \sup_{t\in[0,1]\cap\mathbb{Q}} \|X_t^{(1)}\|.$$

Due to the equivalence of norms, we can now choose the matrix norm

$$\|A\| := \max\{|a_{ij}| : 1 \leq i \leq n \text{ and } 1 \leq j \leq d\}$$

for $A \in M_{n,d}$ and denote by $X_{t,i}^{(1)} \in \mathbb{R}$, $1 \leq i \leq n$, the i-th component of $X_t^{(1)}$, i.e.

$$X_t^{(1)} = \left(X_{t,1}^{(1)}, X_{t,2}^{(1)}, \ldots, X_{t,n}^{(1)}\right)^T.$$

Furthermore, define the (countable) set

$$\widetilde{T} := \{(t,i) : t \in [0,1] \cap \mathbb{Q} \text{ and } i \in \{1,\ldots,n\}\}.$$

Then we obtain

$$\sup_{t\in[0,1]\cap\mathbb{Q}} \|X_t^{(1)}\| = \sup_{t\in[0,1]\cap\mathbb{Q}} \max_{1\leq i\leq n} \|X_{t,i}^{(1)}\| = \sup_{s\in\widetilde{T}} \|X_s^{(1)}\|,$$

where $\sup_{s\in\widetilde{T}}$ is a subadditive functional on $\mathbb{R}^{\widetilde{T}}$. Furthermore, by Theorem 4.2.6 the processes $X_{t,i}^{(1)}$ are infinitely divisible with specified characteristic triplet $(\gamma_{t,i}, \Sigma_{t,i}, \nu_{t,i})$ and Lévy measure

$$\nu_{t,i}(B) = \int_{M_d^-} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \mathbb{1}_B(f_i(A, t-s)x)\nu_1(dx)ds\pi(dA)$$

for all $B \in \mathcal{B}(\mathbb{R})$, where f_i denotes the i-th row of f. It follows that $\mathbf{X^{(1)}} = \{X_s^{(1)} : x \in \widetilde{T}\}$ is infinitely divisible with characteristic triplet $(\widetilde{\gamma}, \widetilde{\Sigma}, \widetilde{\nu})$, where $\widetilde{\gamma}$, $\widetilde{\Sigma}$ and $\widetilde{\nu}$ are given as projective limits of the corresponding finite dimensional characteristics described by $(\gamma_{t,i}, \Sigma_{t,i}, \nu_{t,i})$ (cf. Maruyama (1970)). Moreover, the boundedness $\|f\| \leq C$ implies $\|f_i\| \leq C$ and this, together with the definition of $\nu_1 = \nu|_{B_1(0)}$, yields that the support of the Lévy measures $\nu_{t,i}$ and $\widetilde{\nu}$ can be bounded by C. Now we are able to apply Lemma 2.1 of Braverman and Samorodnitsky (1995) to obtain

$$\mathbb{E}\left(\exp\left(\varepsilon \sup_{s\in\widetilde{T}} \|X_s^{(1)}\|\right)\right) < \infty$$

for all $\varepsilon > 0$. Finally, the finite exponential moments in combination with Lemma 2.1.14 yield

$$\lim_{n\to\infty} nP\Big(a_n^{-1} \sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \|X_{t_1}^{(1)} - X_{t_2}^{(1)}\| \geq \varepsilon\Big) \leq \lim_{n\to\infty} nP\Big(a_n^{-1} \sup_{s\in\widetilde{T}} \|X_s^{(1)}\| \geq \varepsilon/2\Big) = 0$$

for all $\varepsilon > 0$.

□

Next we check the process $X_t^{(2)}$ with respect to the relative compactness conditions (5.4.2), (5.4.3) and (5.4.4).

Proposition 5.4.14. *Let Λ be an $\mathbb{R}^d$-valued Lévy basis on $M_d^- \times \mathbb{R}$ with generating quadruple $(\gamma, \Sigma, \nu, \pi)$ and let $\nu \in RV(\alpha, (a_n), \mu_\nu)$. Assume that the kernel function f is bounded, the MMA process $X_t^{(2)} = \int_{M_d^-}\int_{\mathbb{R}} f(A, t-s)\Lambda_2(dA, ds)$ satisfies the existence conditions of Proposition 5.4.12 and that the regular variation conditions of Theorem 5.2.4 hold. If the function f_δ given by* (5.4.6) *satisfies the existence condition of Proposition 5.4.12 and, as $\delta \to 0$,*

$$\int_{M_d^-}\int_{\mathbb{R}} f_\delta(A,s)^\alpha \, ds\pi(dA) \to 0,$$

then $X_t^{(2)}$ given by (5.4.10) *satisfies the relative compactness conditions* (5.4.2), (5.4.3) *and* (5.4.4).

Proof. We define the difference function $g_{t_1,t_2}(A,s) := f(A, t_1-s) - f(A, t_2-s)$ and mention that for every $t_1, t_2 \in [0,1]$ the random vector

$$X_{t_1}^{(2)} - X_{t_2}^{(2)} = \int_{\|x\|\geq 1}\int_{M_d^-}\int_{\mathbb{R}} g_{t_1,t_2}(A,s)\; x\; N(dx, dA, ds)$$

is again MMA and by Theorem 5.2.3 and Theorem 5.2.4 it exists and is regularly varying with index α.

Condition (5.4.2): We verify the condition by showing that, as $\delta \to 0$,

$$\lim_{n\to\infty} nP(a_n^{-1} \sup_{t_1,t_2\in[0,\delta)} \|X_{t_1}^{(2)} - X_{t_2}^{(2)}\| \geq \varepsilon) \to 0$$

for every $\varepsilon > 0$. We use the decomposition

$$\begin{aligned} &X_{t_1}^{(2)} - X_{t_2}^{(2)} = \\ &\quad = \int_{\|x\|\geq 1} \int_{M_d^-} \int_{(t_1,t_2]} g_{t_1,t_2}(A,s)\; x\; N(dx,dA,ds) + \int_{\|x\|\geq 1} \int_{M_d^-} \int_{(t_1,t_2]^c} g_{t_1,t_2}(A,s)\; x\; N(dx,dA,ds) \\ &\quad =: Z_{t_1,t_2}^{(1)} + Z_{t_1,t_2}^{(2)} \end{aligned} \tag{5.4.11}$$

which yields

$$\begin{aligned} &nP\Big(a_n^{-1} \sup_{t_1,t_2\in[0,\delta)} \|X_{t_1}^{(2)} - X_{t_2}^{(2)}\| \geq \varepsilon\Big) \leq \\ &\qquad \leq nP\Big(a_n^{-1} \sup_{t_1,t_2\in[0,\delta)} \|Z_{t_1,t_2}^{(1)}\| \geq \varepsilon/2\Big) + nP\Big(a_n^{-1} \sup_{t_1,t_2\in[0,\delta)} \|Z_{t_1,t_2}^{(2)}\| \geq \varepsilon/2\Big). \end{aligned} \tag{5.4.12}$$

With $\nu_2 = \nu|_{B_1(0)^c}$ and using the transformation $T : \mathbb{R}^d \to \mathbb{R}$ given by $T(x) = \|x\|$ together with the boundedness $f(A,s) \leq C$ for all $(A,s) \in M_d^- \times \mathbb{R}$ we can now calculate

$$\begin{aligned} \|Z_{t_1,t_2}^{(1)}\| &\leq \int_{\|x\|\geq 1} \int_{M_d^-} \int_{(t_1,t_2]} \|g_{t_1,t_2}(A,s)\|\, \|x\|\, N(dx,dA,ds) \\ &\leq 2\; C\; \Lambda_2^T(M_d^- \times (t_1,t_2]) \\ &= 2\; C\; (L_{t_2}^{(2)} - L_{t_1}^{(2)}), \end{aligned} \tag{5.4.13}$$

where Λ_2^T is a Lévy basis with generating quadruple $(0,0,{\nu_2}^T,\pi)$ and the transformed Lévy measure ${\nu_2}^T$ is given by ${\nu_2}^T(\cdot) = \nu_2(T^{-1}(\cdot))$. By $(L_t^{(2)})$ we denote the underlying Lévy process given by $L_t^{(2)} = \Lambda_2(M_d^- \times (0,t])$ for $t > 0$. Using a continuous mapping argument similar to Theorem 5.4.4 we see that $\nu \in RV(\alpha,(a_n),\mu_\nu)$ implies ${\nu_2}^T \in RV(\alpha,(a_n),\mu_{\nu^T})$ with μ_{ν^T} defined respectively. Thus by Theorem 4.2.3 $L_1^{(2)} \in RV(\alpha,(a_n),\mu_{\nu^T})$ and then by Example 5.4.3 also $(L_t^{(2)}) \in RV_{\mathbb{D}_0}(\alpha,(a_n),\widetilde{\mu})$ for some measure $\widetilde{\mu}$. Now another application of Theorem 5.4.5 yields that condition (5.4.2) holds for the process $(L_t^{(2)})$ and hence, as $\delta \to 0$,

$$\lim_{n\to\infty} nP\Big(a_n^{-1} \sup_{t_1,t_2\in[0,\delta)} \|Z_{t_1,t_2}^{(1)}\| \geq \varepsilon/2\Big) \leq \lim_{n\to\infty} nP\Big(a_n^{-1} \sup_{t_1,t_2\in[0,\delta)} (L_{t_2}^{(2)} - L_{t_1}^{(2)}) \geq \varepsilon/(4C)\Big) \to 0.$$

Similarly, the supremum of the second term $Z^{(2)}_{t_1,t_2}$ can be bounded by

$$\begin{aligned}\sup_{t_1,t_2\in[0,\delta)} \|Z^{(2)}_{t_1,t_2}\| &\leq \int_{\|x\|\geq 1}\int_{M_d^-}\int_{\mathbb{R}} \sup_{t_1,t_2\in[0,\delta)} \|g_{t_1,t_2}(A,s)\|\, \mathbb{1}_{(t_1,t_2]^c}(s)\, \|x\|\, N(dx,dA,ds)\\ &\leq \int_{\|x\|\geq 1}\int_{M_d^-}\int_{\mathbb{R}} f_\delta(A,s)\, \|x\|\, N(dx,dA,ds)\\ &= \int_{M_d^-}\int_{\mathbb{R}} f_\delta(A,s)\Lambda_2^T(dA,ds) =: Y.\end{aligned}$$

Then assumption (5.4.7) implies $f_\delta \in \mathbb{L}^\alpha$ for some $\delta > 0$ sufficiently small and another application of Theorem 5.2.4 yields $Y \in RV(\alpha,(a_n),\mu_Y)$ with

$$\mu_Y(B) := \int_{M_d^-}\int_{\mathbb{R}}\int_{\mathbb{R}^d} \mathbb{1}_B\left(f_\delta(A,s)\|x\|\right)\mu_\nu(dx)ds\pi(dA).$$

Finally, as $n\to\infty$, we obtain

$$\begin{aligned}nP\Big(a_n^{-1}\sup_{t_1,t_2\in[0,\delta)}\|Z^{(2)}_{t_1,t_2}\|\geq\varepsilon/2\Big) &\leq nP(a_n^{-1}Y\geq\varepsilon/2)\\ &\to \int_{M_d^-}\int_{\mathbb{R}}\mu_\nu(x: f_\delta(A,s)\|x\|\geq\varepsilon/2)\, ds\,\pi(dA)\\ &= \mu_\nu(x:\|x\|\geq\varepsilon/2)\int_{M_d^-}\int_{\mathbb{R}} f_\delta(A,s)^\alpha\, ds\,\pi(dA)\end{aligned}$$

and since μ_ν is a Radon measure the result follows by the assumption.

Condition (5.4.3)*:* The condition follows likewise to condition (5.4.2) (note also that the MMA process (X_t) is stationary).

Condition (5.4.4)*:* For the third condition we use (5.4.11) again and obtain

$$\begin{aligned}\sup_{t_1\leq t\leq t_2;\, t_2-t_1\leq\delta}&\min\left\{\|X_t^{(2)}-X_{t_2}^{(2)}\|,\|X_{t_1}^{(2)}-X_t^{(2)}\|\right\}\leq\\ &\leq \sup_{t_1\leq t\leq t_2;\, t_2-t_1\leq\delta}\min\left\{\|Z_{t,t_2}^{(1)}\|,\|Z_{t_1,t}^{(1)}\|\right\}+2\sup_{t_1\leq t_2;\, t_2-t_1\leq\delta}\|Z_{t_1,t_2}^{(2)}\|\end{aligned}$$

and

$$\begin{aligned}&nP\Big(a_n^{-1}\sup_{t_1\leq t\leq t_2;\, t_2-t_1\leq\delta}\min\left\{\|X_t^{(2)}-X_{t_2}^{(2)}\|,\|X_{t_1}^{(2)}-X_t^{(2)}\|\right\}\geq\varepsilon\Big)\leq\\ &\leq nP\Big(a_n^{-1}\sup_{t_1\leq t\leq t_2;t_2-t_1\leq\delta}\min\left\{\|Z_{t,t_2}^{(1)}\|,\|Z_{t_1,t}^{(1)}\|\right\}\geq\frac{\varepsilon}{2}\Big)+nP\Big(a_n^{-1}\sup_{t_1\leq t_2;t_2-t_1\leq\delta}\|Z_{t_1,t_2}^{(2)}\|\geq\frac{\varepsilon}{4}\Big).\end{aligned}$$

Applying (5.4.13) this implies, as $\delta \to 0$,

$$\lim_{n\to\infty} nP\Big(a_n^{-1} \sup_{t_1\le t\le t_2;\, t_2-t_1\le\delta} \min\big\{\|Z^{(1)}_{t,t_2}\|, \|Z^{(1)}_{t_1,t}\|\big\} \ge \varepsilon/2\Big) \le$$
$$\le \lim_{n\to\infty} nP\Big(a_n^{-1} \sup_{t_1\le t\le t_2;\, t_2-t_1\le\delta} \min\big\{\|L^{(2)}_{t_2} - L^{(2)}_t\|, \|L^{(2)}_t - L^{(2)}_{t_1}\|\big\} \ge \varepsilon/(4\,C)\Big) \to 0,$$

since this is exactly condition (5.4.4) for the Lévy process $L^{(2)}_t$ which is regularly varying in $\mathbb{D}$ and thus by Theorem 5.4.5 satisfies (5.4.4). Furthermore,

$$\sup_{t_1\le t_2;\, t_2-t_1\le\delta} \|Z^{(2)}_{t_1,t_2}\| \le \int_{M_d^-}\int_{\mathbb{R}} f_\delta(A,s)\Lambda_2^T(dA,ds) = Y$$

and consequently, as $\delta \to 0$,

$$\lim_{n\to\infty} nP\Big(a_n^{-1} \sup_{t_1\le t_2;\, t_2-t_1\le\delta} \|Z^{(2)}_{t_1,t_2}\| \ge \varepsilon/4\Big) \le nP\Big(a_n^{-1}\, Y \ge \varepsilon/4\Big) \to 0$$

as shown for condition (5.4.2).

□

This concludes the proof of Theorem 5.4.7.

5.5 Application to SupOU Processes

In Section 4.4 we have introduced superpositions of Ornstein-Uhlenbeck processes (supOU processes) which have useful properties and a wide range of applications. A supOU process (X_t) can be defined as an MMA process with kernel function

$$f(A,s) = e^{As}\mathbb{1}_{[0,\infty)}(s).$$

We will shortly recall the main results of Section 4.4. Sufficient conditions for the existence of supOU processes are given in the following theorem which takes the special properties of supOU processes into account.

Theorem 5.5.1 (Barndorff-Nielsen and Stelzer (2011a), Theorem 3.1). *Let X_t be an $\mathbb{R}^d$-valued supOU process as defined in Definition 4.4.1. If*

$$\int_{\|x\|>1} \ln(\|x\|)\nu(dx) < \infty$$

and there exist measurable functions $\rho : M_d^- \mapsto \mathbb{R}^+\backslash\{0\}$ and $\kappa : M_d^- \mapsto [1,\infty)$ such that

$$\left\|e^{As}\right\| \le \kappa(A)e^{-\rho(A)s}\ \forall s \in \mathbb{R}^+\ \pi\text{-almost surely and } \int_{M_d^-} \frac{\kappa(A)^2}{\rho(A)}\pi(dA) < \infty,$$

then the supOU process $X_t = \int_{M_d^-} \int_{-\infty}^t e^{A(t-s)} \Lambda(dA, ds)$ *is well defined for all* $t \in \mathbb{R}$ *and stationary. Furthermore, the stationary distribution of* X_t *is infinitely divisible with characteristic triplet* $(\gamma_X, \Sigma_X, \nu_X)$ *given by Theorem 4.2.6.*

Conditions for regular variation of X_t and of the finite-dimensional distributions of (X_t) were given in Corollary 4.4.3 and Corollary 4.4.6.

Corollary 5.5.2. *Let* $\Lambda \in \mathbb{R}^d$ *be a Lévy basis on* $M_d^- \times \mathbb{R}$ *with generating quadruple* $(\gamma, \Sigma, \nu, \pi)$ *and let* $\nu \in RV(\alpha, (a_n), \mu_\nu)$*. If the conditions of Theorem 5.5.1 hold and additionally*

$$\int_{M_d^-} \frac{\kappa(A)^\alpha}{\rho(A)} \, \pi(dA) < \infty,$$

then $X_0 = \int_{M_d^-} \int_{\mathbb{R}^+} e^{As} \Lambda(dA, ds) \in RV(\alpha, (a_n), \mu_X)$ *with Radon measure*

$$\mu_X(\cdot) := \int_{M_d^-} \int_{\mathbb{R}^+} \int_{\mathbb{R}^d} \mathbb{1}_{(\cdot)} \left(e^{As} x\right) \mu_\nu(dx) ds \pi(dA).$$

Furthermore, the finite dimensional distributions $(X_{t_1}, \ldots, X_{t_k})$*,* $t_i \in \mathbb{R}$ *and* $k \in \mathbb{N}$*, are also regularly varying with index* α *and given limiting measure* $\mu_{t_1,\ldots,t_k}$*.*

In order to apply Theorem 5.4.7 to obtain conditions for regular variation of supOU processes in $\mathbb{D}$, we state some useful sufficient conditions for the function $f_\delta(A, s) = \sup_{t_1 \le t_2;\, t_2 - t_1 \le \delta} \| f(A, t_2 - s) - f(A, t_1 - s) \| \, \mathbb{1}_{(t_1, t_2]^c}(s)$ to be an element of $\mathbb{L}^\alpha$ for $\alpha > 0$.

Proposition 5.5.3. *Let* $f(A, s) = e^{As} \mathbb{1}_{[0,\infty)}(s)$ *be the kernel function of a supOU process satisfying the conditions of Theorem 5.5.1 and let* f_δ *be given by* (5.4.6)*. If for some* $\alpha > 0$

$$\int_{M_d^-} \frac{\kappa(A)^\alpha}{\rho(A)} \, \pi(dA) < \infty,$$

and

$$\int_{M_d^-} \kappa(A)^\alpha \, \pi(dA) < \infty,$$

then $f_\delta \in \mathbb{L}^\alpha$*.*

Proof. We start with the observation

$$f_\delta(A, s)^\alpha = \sup_{t_1 \le t_2;\, t_2 - t_1 \le \delta} \left\| e^{A(t_2 - s)} - e^{A(t_1 - s)} \right\|^\alpha \mathbb{1}_{(-\infty, t_1]}(s)$$

$$\leq \sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \left\|e^{A(t_2-s)} - e^{A(t_1-s)}\right\|^\alpha \mathbb{1}_{(-\infty,0]}(s) + \sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \left\|e^{A(t_2-s)} - e^{A(t_1-s)}\right\|^\alpha \mathbb{1}_{(0,t_1]}(s). \tag{5.5.1}$$

Now for the first summand we obtain

$$\int_{M_d^-}\int_{\mathbb{R}} \sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \left\|e^{A(t_2-s)} - e^{A(t_1-s)}\right\|^\alpha \mathbb{1}_{(-\infty,0]}(s)\, ds\, \pi(dA) \leq$$

$$\leq \int_{M_d^-}\int_{-\infty}^{0} \sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \left(\left\|e^{A(t_2-s)}\right\| + \left\|e^{A(t_1-s)}\right\|\right)^\alpha ds\, \pi(dA)$$

$$\leq \int_{M_d^-}\int_{-\infty}^{0} \sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \kappa(A)^\alpha \left(e^{-\rho(A)(t_2-s)} + e^{-\rho(A)(t_1-s)}\right)^\alpha ds\, \pi(dA)$$

$$\leq 2^\alpha \int_{M_d^-}\int_{-\infty}^{0} \kappa(A)^\alpha\, e^{s\rho(A)\alpha}\, ds\pi(dA)$$

$$= \frac{2^\alpha}{\alpha} \int_{M_d^-} \frac{\kappa(A)^\alpha}{\rho(A)}\, \pi(dA)$$

$$< \infty$$

and the second summand of (5.5.1) can be bounded by

$$\int_{M_d^-}\int_{\mathbb{R}} \sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \left\|e^{A(t_2-s)} - e^{A(t_1-s)}\right\|^\alpha \mathbb{1}_{(0,t_1]}(s)\, ds\, \pi(dA) \leq$$

$$\leq \int_{M_d^-}\int_{\mathbb{R}} \sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \left(\left\|e^{A(t_2-s)}\right\| + \left\|e^{A(t_1-s)}\right\|\right)^\alpha \mathbb{1}_{(0,t_1]}(s)\, ds\, \pi(dA)$$

$$\leq \int_{M_d^-}\int_{\mathbb{R}} \sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \kappa(A)^\alpha \left(e^{-\rho(A)(t_2-s)} + e^{-\rho(A)(t_1-s)}\right)^\alpha \mathbb{1}_{(0,t_1]}(s)\, ds\, \pi(dA)$$

$$\leq \int_{M_d^-}\int_{\mathbb{R}} \kappa(A)^\alpha\, 2^\alpha\, \mathbb{1}_{[0,1]}(s)\, ds\pi(dA)$$

$$= 2^\alpha \int_{M_d^-} \kappa(A)^\alpha\, \pi(dA)$$

$$< \infty.$$

□

Remark 5.5.4. Requiring $f_\delta \in \mathbb{L}^\alpha$, the first condition

$$\int_{M_d^-} \frac{\kappa(A)^\alpha}{\rho(A)}\, \pi(dA) < \infty$$

is already needed for the regular variation condition $f \in \mathbb{L}^\alpha$ of Corollary 5.5.2. The additional condition $\int_{M_d^-} \kappa(A)^\alpha\, \pi(dA) < \infty$ is of importance if $\rho(A)$ decays very fast. In that case

$$\frac{\kappa(A)^\alpha}{\rho(A)} \ll \kappa(A)^\alpha$$

for values of $A \in M_d^-$ with high norm and therefore, the stronger integrability condition is needed. If we consider the case of a bounded kernel function f, i.e. in the supOU case κ is bounded, it follows directly that f_δ is also bounded. Moreover, the condition

$$\int_{M_d^-} \kappa(A)^\alpha\, \pi(dA) < \infty$$

is true for all $\alpha > 0$ and for all probability measures π.

Now we can use Proposition 5.5.3 to obtain conditions for functional regular variation of supOU processes with sample paths in $\mathbb{D}$. Therefore, we restrict the time interval to $t \in [0,1]$ and assume the supOU process to have càdlàg sample paths, see Section 5.3 and Barndorff-Nielsen and Stelzer (2011a), Theorem 3.12, for details on the sample path behavior of supOU processes.

Theorem 5.5.5. *Let Λ and Λ_2 be $\mathbb{R}^d$-valued Lévy bases on $M_d^- \times \mathbb{R}$ with generating quadruples $(\gamma, \Sigma, \nu, \pi)$ and $(0, 0, \nu|_{B_1(0)^c}, \pi)$ respectively such that $\nu \in RV(\alpha, (a_n), \mu_\nu)$. Assume that the supOU process (X_t) given by $X_t = \int_{M_d^-} \int_{-\infty}^t e^{A(t-s)} \Lambda(dA, ds)$ exists for $t \in [0,1]$ (in the sense of Theorem 5.5.1) and that the processes X_t and $X_t^{(2)} = \int_{M_d^-} \int_{\mathbb{R}} f(A, t-s)\Lambda_2(dA, ds)$ have càdlàg sample paths. Furthermore, suppose that κ is bounded and that*

$$\int_{M_d^-} \frac{\kappa(A)^\alpha}{\rho(A)}\, \pi(dA) < \infty.$$

If $f(A,s) = e^{As}\mathbb{1}_{[0,\infty)}(s)$ and the function f_δ given by (5.4.6) satisfy condition (5.4.5), then

$$(X_t)_{t\in[0,1]} \in RV_{\overline{\mathbb{D}}_0}(\alpha, (a_n), \mu),$$

where μ is uniquely determined by the measures $\mu_{t_1,\dots,t_k}$ in Corollary 5.5.2.

Proof. Applying Theorem 5.4.7, we start by observing that the supOU kernel function

$$f(A,s) = e^{As}\mathbb{1}_{[0,\infty)}(s) \leq \kappa(A)$$

is bounded if κ is. Next we show that, as $\delta \to 0$,

$$\int_{M_d^-} \int_{\mathbb{R}} f_\delta(A,s)^\alpha \, ds\pi(dA) \to 0.$$

The assumptions together with Proposition 5.5.3 and Remark 5.5.4 yield $f_\delta \in \mathbb{L}^\alpha$ and thus by Lemma 5.4.8 it is sufficient to show that $\lim_{\delta\to 0} f_\delta(A,s) \to 0$ for $\pi \times \lambda$ almost every (A,s). When considering differences of matrix exponentials, we can use the inequality

$$\begin{aligned}\left\|e^{A(t_2-s)} - e^{A(t_1-s)}\right\| &= \left\| \sum_{k=0}^{\infty} \frac{A^k((t_2-s)^k - (t_1-s)^k)}{k!} \right\| \leq \sum_{k=0}^{\infty} \frac{\|A\|^k((t_2-s)^k - (t_1-s)^k)}{k!} \\ &= e^{\|A\|(t_2-s)} - e^{\|A\|(t_1-s)}\end{aligned}$$

if $t_2 > t_1$. This yields

$$\begin{aligned} f_\delta(A,s) &= \sup_{t_1\leq t_2;\, t_2-t_1\leq\delta} \left\|e^{A(t_2-s)} - e^{A(t_1-s)}\right\| \mathbb{1}_{(-\infty,t_1]}(s) \\ &\leq \sup_{x\in[0,1-\delta]} \sup_{t_2,t_1\in[x,x+\delta];\, t_1\leq t_2} \left(e^{\|A\|(t_2-s)} - e^{\|A\|(t_1-s)}\right) \mathbb{1}_{(-\infty,t_1]}(s) \\ &\leq \sup_{x\in[0,1-\delta]} \left(e^{\|A\|(x+\delta-s)} - e^{\|A\|(x-s)}\right) \mathbb{1}_{(-\infty,t_1]}(s) \\ &\leq \sup_{x\in[0,1-\delta]} \left(e^{\|A\|\delta} - 1\right) e^{\|A\|(x-s)} \mathbb{1}_{(-\infty,t_1]}(s) \\ &\leq \left(e^{\|A\|\delta} - 1\right) e^{\|A\|(1-\delta-s)} \mathbb{1}_{(-\infty,1-\delta]}(s) \end{aligned}$$

and by the continuity of the exponential this term converges to 0 as $\delta \to 0$ for every $(A,s) \in \mathcal{B}(M_d^- \times \mathbb{R})$.

□

Conditions for f and f_δ to satisfy the existence condition (5.4.5), i.e.

$$\int_{M_d^-} \int_{\mathbb{R}} \int_{\|x\|>1} (1 \wedge \|f(A,s)x\|) \, \nu(dx) \, ds \, \pi(dA) < \infty$$

can be obtained by combining Lemma 5.5.1 with Lemma 5.4.10.

Corollary 5.5.6. *Let $f(A,s) = e^{As}\mathbb{1}_{[0,\infty)}(s)$ be the kernel function of a supOU process satisfying the conditions of Theorem 5.5.1 and let f_δ be given by* (5.4.6)*. Then f and f_δ satisfy condition* (5.4.5) *if one of the following two conditions are satisfied:*

(i) $\alpha > 1$ *as well as*

$$\int_{M_d^-} \frac{\kappa(A)}{\rho(A)}\, \pi(dA) < \infty \quad \text{and} \int_{M_d^-} \kappa(A)\, \pi(dA) < \infty.$$

(ii) $\alpha \leq 1$ *and there exists* $\varepsilon \in (0, \alpha)$ *such that*

$$\int_{M_d^-} \frac{\kappa(A)^{\alpha-\varepsilon}}{\rho(A)}\, \pi(dA) < \infty \quad \text{and} \int_{M_d^-} \kappa(A)^{\alpha-\varepsilon}\, \pi(dA) < \infty.$$

In correspondence with Remark 5.5.4 we mention that for all $\alpha > 0$, $\varepsilon \in (0, \alpha)$ and for all probability measures π the conditions

$$\int_{M_d^-} \kappa(A)\, \pi(dA) < \infty \quad \text{and} \int_{M_d^-} \kappa(A)^{\alpha-\varepsilon}\, \pi(dA) < \infty$$

are redundant if κ is bounded.

5.6 Point Process Convergence

In this section we discuss the use of the results of the previous two sections in combination with point process results for stochastic processes with sample paths in $\mathbb{D}$. Therefore, let $M_p(\overline{\mathbb{D}}_0)$ denote the space of all point measures on $\overline{\mathbb{D}}_0$ equipped with the $\hat{w}$-topology and let ε_x be the Dirac measure at the point x. Furthermore, let X_i, $i \in \mathbb{N}$, be a sequence of iid copies of a regularly varying stochastic process $X \in RV_{\overline{\mathbb{D}}_0}(\alpha, (a_n), \mu)$ with values in $\mathbb{D}$.

We start by stating the main result that links regular variation of X to weak convergence of the point processes

$$N_n = \sum_{i=1}^{n} \varepsilon_{a_n^{-1} X_i}, \quad n \in \mathbb{N}.$$

The following theorem is the extension of the classical result of Proposition 3.21 in Resnick (1987) to a state space which is not locally compact. Similar results have also been proved by de Haan and Lin (2001), Theorem 2.4, in the case of real-valued processes which are regularly varying with index 1 and by Davis and Mikosch (2008) for $\mathbb{D}$-valued random fields.

Theorem 5.6.1. *Let* $(X_i)_{i\in\mathbb{N}}$ *be an iid sequence of stochastic processes with values in* $\mathbb{D}$. *Then* $X_1 \in RV_{\overline{\mathbb{D}}_0}(\alpha, (a_n), \mu)$ *if and only if* $N_n \xrightarrow{d} N$ *in* $M_p(\overline{\mathbb{D}}_0)$, *where* N *is a Poisson random measure with mean measure* μ *(short PRM(*μ*)).*

Proof. See the proof of Theorem 2.2.18 in Section 2.2.2.

□

This result can now be combined with the results of Sections 5.4 and 5.5 to obtain functional point process convergence for MMA and supOU processes. Point processes of that kind include full information of the complete paths of the process X. In combination with the continuous mapping theorem (cf. Daley and Vere-Jones (1988), Proposition A2.3.V) this is an extremely powerful tool to analyze the extremal behavior of MMA and supOU processes. Using such methods, one gets a better understanding of the structure of the extreme values and their properties, e.g. the extremal clustering behavior or long memory effects.

In contrast to finite-dimensional point process results, functional point process convergence does not only allow to analyze, for example, the behavior of maxima at fixed time points, but also of functionals acting on the paths of the process in compact time intervals. Examples of such functionals are the subadditive functionals (e.g. suprema) studied by Rosiński and Samorodnitsky (1993) for a subexponential, by Braverman and Samorodnitsky (1995) for an exponential, and by Braverman et al. (2002) for a univariate regularly varying setting. Moreover, since point processes of suprema do not incorporate the directions of the extremes, it is also possible to include the directions into the analyzed point processes. Finally, we mention that basically in any field of extreme value theory, point process techniques are often very helpful to prove results.

For more reading on point processes we refer to Daley and Vere-Jones (1988) and Daley and Vere-Jones (2008). See also Kallenberg (1983) for a more general reading on random measures. Very good introductions to the use of point processes in extreme value theory can be found in Embrechts et al. (1997), Resnick (1987), Resnick (2007), Leadbetter et al. (1983) and de Haan and Ferreira (2006). For the exemplary use of functional point processes, see de Haan and Lin (2001) and Davis and Mikosch (2008).

Bibliography

D. Applebaum (2004). *Lévy processes and stochastic calculus*. Cambridge University Press, Cambridge.

O.E. Barndorff-Nielsen (2001). Superposition of Ornstein-Uhlenbeck type processes. *Theory Probab. Appl.*, 45:175–194.

O.E. Barndorff-Nielsen and N. Shephard (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics (with discussion). *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 63:167–241.

O.E. Barndorff-Nielsen and R. Stelzer (2011a). Multivariate supOU processes. *Ann. Appl. Probab.*, 21(1):140–182.

O.E. Barndorff-Nielsen and R. Stelzer (2011b). The multivariate supOU stochastic volatility model. *Math. Finance, accepted for publication*.

B. Basrak (2000). *The Sample Autocorrelation Function of Non-Linear Time Series*. PhD thesis, University of Groningen.

B. Basrak and J. Segers (2009). Regularly varying multivariate time series. *Stoch. Proc. Appl.*, 119:1055–1080.

B. Basrak, R. A. Davis and T. Mikosch (1999). The sample ACF of a simple bilinear process. *Stoch. Proc. Appl.*, 83:1–14.

B. Basrak, R. A. Davis and T. Mikosch (2002). Regular variation of GARCH processes. *Stoch. Proc. Appl.*, 99:95–115.

A. Basse and J. Pedersen (2009). Lévy driven moving averages and semimartingales. *Stoch. Proc. Appl.*, 119:2970–2991.

A. Basse and J. Rosiński (2011). On the uniform convergence of random series in skorohod space and representations of càdlàg infinitely divisible processes. *Submitted for Publication*.

C. Bender, A. Lindner and M. Schicks (2011). Finite variation of fractional Lévy processes. *J. Theor. Probab., accepted for publication.*

P. Billingsley (1968). *Convergence of Probability Measures.* Wiley, New York.

N. H. Bingham, C. M. Goldie and J. L. Teugels (1987). *Regular Variation.* Cambridge University Press, Cambridge.

T. Bollerslev (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, 31:307–327.

J. Boman and F. Lindskog (2009). Support theorems for the Radon transform and Cramér-Wold theorems. *J. Theor. Probab.*, 22:683–710.

M. Braverman and G. Samorodnitsky (1995). Functionals of infinitely divisible stochastic processes with exponential tails. *Stoch. Proc. Appl.*, 56:207–231.

M. Braverman, T. Mikosch and G. Samorodnitsky (2002). Tail probabilities of subadditive functionals acting on Lévy processes. *Ann. Appl. Probab.*, 12:69–100.

P. Brockwell (2001). Lévy-driven CARMA processes. *Ann. Inst. Stat. Math.*, 53:113–124.

P. Brockwell (2004). Representations of continuous-time ARMA processes. *J. Appl. Probab.*, 41:375–382.

S. Cambanis, J. P. Nolan and J. Rosiński (1990). On the oscillation of infinitely divisible and some other processes. *Stoch. Proc. Appl.*, 35:87–97.

D. Cline (1983). *Infinite series of random variables with regularly varying tail.* Tech. Report 83-24, Institute of Applied Mathematics and Statistics, University of British Columbia.

M. Csörgö and L. Horváth (1997). *Limit Theorems in Change Point Analysis.* Wiley, New York.

D. J. Daley and D. Vere-Jones (1988). *An Introduction to the Theory of Point Processes.* Springer, New York.

D. J. Daley and D. Vere-Jones (2008). *An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure.* Springer, New York.

R. A. Davis and T. Hsing (1995). Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Probab.*, 23:879–917.

R. A. Davis and T. Mikosch (1998). The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Ann. Statist.*, 26:2049–2080.

R. A. Davis and T. Mikosch (2001). Point process convergence of stochastic volatility processes with application to sample autocorrelation. *J. Appl. Probab.*, 38A:93–104.

R. A. Davis and T. Mikosch (2008). Extreme value theory for space-time processes with heavy-tailed distributions. *Stoch. Proc. Appl.*, 118:560–584.

R. A. Davis and S. I. Resnick (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.*, 13(1):179–195.

R. A. Davis and S. I. Resnick (1986). Limit theory for the sample covariance and correlation functions of moving averages. *Ann. Statist.*, 14(2):533–558.

R. A. Davis and S. I. Resnick (1996). Limit theory for bilinear processes with heavy-tailed noise. *Ann. Appl. Probab.*, 6:1191–1210.

L. de Haan and A. Ferreira (2006). *Extreme Value Theory*. Springer, New York.

L. de Haan and T. Lin (2001). On convergence toward an extreme value distribution in $C[0,1]$. *Ann. Probab.*, 29(1):467–483.

L. de Haan, S. I. Resnick, H. Rootzén and C. G. de Vries (1989). Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stoch. Proc. Appl.*, 32:213–224.

H. Dehling, O. Durieu and D. Volny (2009). New techniques for empirical processes of dependent data. *Stoch. Proc. Appl.*, 119:3699–3718.

P. Embrechts, C. Klüppelberg and T. Mikosch (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.

V. Fasen (2005). Extremes of regularly varying Lévy driven mixed moving average processes. *Adv. Appl. Probab.*, 37:993–1014.

V. Fasen (2009). Extremes of Lévy driven mixed MA processes with convolution equivalent distributions. *Extremes*, 12(3):265–296.

V. Fasen and C. Klüppelberg (2007). Extremes of supOU processes. In F.E. Benth, G. Di Nunno, T. Lindstrom, B. Øksendal and T. Zhang, editors, *Stochastic Analysis and Applications: The Abel Symposium 2005*, pages 340–359. Springer, New York.

W. Feller (1966). *An Introduction to Probability Theory and its Applications*. Wiley, New York, 2nd edition.

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

J. E. Griffin and M. F. J. Steel (2010). Bayesian inference with stochastic volatility models using continuous superpositions of non-Gaussian Ornstein-Uhlenbeck processes. *Comput. Statist. Data Anal.*, 54:2594–2608.

R. A. Horn and C. R. Johnson (1991). *Topics in Matrix Analysis*. Cambridge University Press, Cambridge.

H. Hult and F. Lindskog (2005). Extremal behavior of regularly varying stochastic processes. *Stoch. Proc. Appl.*, 115:249–274.

H. Hult and F. Lindskog (2006a). On regular variation for infinitely divisible random vectors and additive processes. *Adv. Appl. Probab.*, 38:134–148.

H. Hult and F. Lindskog (2006b). Regular variation for measures on metric spaces. *Publications de l'Institut Mathématique, Nouvelle Série*, 80:121–140.

H. Hult and G. Samorodnitsky (2008). Tail probabilities for infinite series of regularly varying random vectors. *Bernoulli*, 14:836–864.

J. Illian, A. Penttinen, H. Stoyan and D. Stoyan (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*. John Wiley & Sons, Chichester.

M. Jacobsen, T. Mikosch, J. Rosiński and G. Samorodnitsky (2009). Inverse problems for regular variation of linear filters, a cancellation property for σ-finite measures and identification of stable laws. *Ann. Appl. Probab.*, 19(1):210–242.

M. Juodis, A. Račkauskas and C. Suquet (2009). Hölderian invariance principle for linear processes. ALEA, 5:1–18.

O. Kallenberg (1983). *Random Measures*. Academic Press, New York.

O. Kallenberg (2002). *Foundation of Modern Probability*. Springer, New York.

J. Karamata (1930). Sur un mode de croissance régulière des fonctions. *Mathematica (Cluj)*, 4:38 – 53.

H. Kesten (1973). Random difference equations and renewal theory for products of random matrices. *Acta Math.*, 131:297–248.

A. Kyprianou (2006). *Introductory Lectures on Fluctuations of Lévy Processes with Applications*. Springer, Berlin.

M. R. Leadbetter, G. Lindgren and H. Rootzén (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, Berlin.

F. Lindskog (2004). *Multivariate Extremes and Regular Variation for Stochastic Processes*. PhD thesis, ETH Zurich.

M. B. Marcus and J. Rosiński (2005). Continuity and boundedness of infinitely divisible processes: A poisson point process approach. *J. Theor. Probab.*, 18:109–160.

T. Marquardt (2006). Fractional Lévy processes with an application to long memory moving average processes. *Bernoulli*, 12:1009–1126.

T. Marquardt (2007). Multivariate FICARMA processes. *J. Mult. Anal.*, 98:1705–1725.

T. Marquardt and R. Stelzer (2007). Multivariate CARMA processes. *Stoch. Proc. Appl.*, 117:96–120.

F. Maruyama (1970). Infinitely divisible processes. *Theory Probab. Appl.*, 15:1–22.

K. Maulik, S. I. Resnick and H. Rootzén (2002). Asymptotic independence and a network traffic model. *J. Appl. Probab.*, 39:671–699.

T. Mikosch and D. G. Konstantinides (2005). Large deviations and ruin probabilities for solutions to stochastic recurrence equations with heavy-tailed innovations. *Ann. Probab.*, 33:1992–2035.

T. Mikosch and M. Moser (2012). The limit distribution of the maximum increment of a random walk with dependent regularly varying jump sizes. *Probab. Theory Relat. Fields, accepted for publication.*

T. Mikosch and A. Račkauskas (2010). The limit distribution of the maximum increment of a random walk with regularly varying jump size distribution. *Bernoulli*, 16:1016–1038.

T. Mikosch and G. Samorodnitsky (2000a). The supremum of a negative drift random walk with dependent heavy-tailed steps. *Ann. Appl. Probab.*, 10:1025–1064.

T. Mikosch and G. Samorodnitsky (2000b). Ruin probability with claims modeled by a stationary ergodic stable process. *Ann. Probab.*, 28:1814–1851.

T. Mikosch and C. Stărică (2000). Limit theory for the sample autocorrelations and extremes of a GARCH$(1,1)$ process. *Ann. Statist.*, 28:1427–1451.

M. Moser and R. Stelzer (2011). Tail behavior of multivariate Lévy driven mixed moving average processes and related stochastic volatility models. *Adv. Appl. Prob.*, 43(4):1109–1135.

J. Pedersen (2003). The Lévy-Ito decomposition of an independently scattered random measure. *MaPhySto Research Report 2, MaPhySto, Århus. Available online at www.maphysto.dk*.

V. V. Petrov (1995). *Limit Theorems of Probability Theory. Sequences of Independent Random Variables*. Oxford Science Publications, New York.

C. Pigorsch and R. Stelzer (2009). A multivariate Ornstein-Uhlenbeck type stochastic volatility model. *Submitted for publication*.

A. Račkauskas and C. Suquet (2004). Hölder norm test statistics for epidemic change. *J. Statist. Plann. Inference*, 126:495–520.

A. Račkauskas and C. Suquet (2006). Testing epidemic changes of infinite-dimensional parameters. *Statist. Inference Stoch. Proc.*, 9:111–134.

A. Račkauskas and C. Suquet (2010). On limit theorems for Banach space valued linear processes. *Lith. Math. J.*, 50:71–87.

B. S. Rajput and J. Rosiński (1989). Spectral representations of infinitely divisible processes. *Probab. Theory Relat. Fields*, 82:451–487.

S. I. Resnick (1986). Point processes, regular variation and weak convergence. *Adv. Appl. Prob.*, 18:66–138.

S. I. Resnick (1987). *Extreme Values, Regular Variation and Point Processes*. Springer, New York.

S. I. Resnick (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, New York.

S. I. Resnick and H. Rootzén (2000). Self-similar communication models and very heavy tails. *Ann. Appl. Prob.*, 10:753–778.

H. Rootzén (1978). Extremes of moving averages of stable processes. *Ann. Probab.*, 6: 847–869.

J. Rosiński (1986). On stochastic integral representation of stable processes with sample paths in Banach spaces. *J. Mult. Anal.*, 20:277–302.

J. Rosiński (1989). On path properties of certain infinitely divisible processes. *Stoch. Proc. Appl.*, 33:73–87.

J. Rosiński and G. Samorodnitsky (1993). Distributions of subadditive functionals of sample paths of infinitely divisible processes. *Ann. Probab.*, 21:996–1014.

J. Rosiński, G. Samorodnitsky and M. S. Taqqu (1991). Sample path properties of stochastic processes represented as multiple stable integrals. *J. Mult. Anal.*, 37:115–134.

G. Samorodnitsky and M. Taqqu (1994). *Stable Non-Gaussian Random Processes*. Chapman & Hall/CRC, Florida.

K. Sato (2002). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge.

D. Surgailis, J. Rosiński, V. Mandrekar and S. Cambanis (1993). Stable mixed moving averages. *Probab. Theory Relat. Fields*, 97:543–558.

Index

P

R

S

T

U

V

W