

Fasen, Vicky; Klüppelberg, Claudia; Lindner, Alexander M.

**Working Paper**

## Extremal behavior of stochastic volatility models

Discussion Paper, No. 427

**Provided in Cooperation with:**

Collaborative Research Center (SFB) 386: Statistical Analysis of discrete structures - Applications in Biometrics and Econometrics, University of Munich (LMU)

*Suggested Citation:* Fasen, Vicky; Klüppelberg, Claudia; Lindner, Alexander M. (2005) : Extremal behavior of stochastic volatility models, Discussion Paper, No. 427, Ludwig-Maximilians-Universität München, Sonderforschungsbereich 386 - Statistische Analyse diskreter Strukturen, München, <https://doi.org/10.5282/ubm/epub.1796>

This Version is available at:

<https://hdl.handle.net/10419/30995>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Extremal behavior of stochastic volatility models

Vicky Fasen \*    Claudia Klüppelberg \*    Alexander Lindner \*

## Abstract

Empirical volatility changes in time and exhibits tails, which are heavier than normal. Moreover, empirical volatility has - sometimes quite substantial - upwards jumps and clusters on high levels. We investigate classical and non-classical stochastic volatility models with respect to their extreme behavior. We show that classical stochastic volatility models driven by Brownian motion can model heavy tails, but obviously they are not able to model volatility jumps. Such phenomena can be modelled by Lévy driven volatility processes as, for instance, by Lévy driven Ornstein-Uhlenbeck models. They can capture heavy tails and volatility jumps. Also volatility clusters can be found in such models, provided the driving Lévy process has regularly varying tails. This results then in a volatility model with similarly heavy tails. As the last class of stochastic volatility models, we investigate a continuous time GARCH(1,1) model. Driven by an arbitrary Lévy process it exhibits regularly varying tails, volatility upwards jumps and clusters on high levels.

*2000 MSC Subject Classifications:* primary: 60G70, 91B70

secondary: 60G10, 91B84

*JEL Classifications:* C23, C51

*Keywords:* COGARCH, extreme value theory, generalized Cox-Ingersoll-Ross model, Lévy process, Ornstein-Uhlenbeck process, Poisson approximation, regular variation, stochastic volatility model, subexponential distribution, tail behavior, volatility cluster.

---

\*Center for Mathematical Sciences, Munich University of Technology, D-85747 Garching, Germany, email: {fasen, cklu, lindner}@ma.tum.de, [www.ma.tum.de/stat/](http://www.ma.tum.de/stat/)

# 1 Introduction

The classical pricing model is the Black-Scholes model given by the SDE

$$dS_t = rS_t dt + \sigma S_t dB_t, \quad S_0 = x \in \mathbb{R}, \quad (1.1)$$

where  $r \in \mathbb{R}$  is the stock-appreciation rate,  $\sigma > 0$  is the volatility and  $B$  is a standard Brownian motion. The Black-Scholes model is based on the assumption that the relative price changes of the asset form a Gaussian process with stationary and independent increments. The crucial parameter is the volatility  $\sigma$ , which is in this model assumed to be constant. However, empirical analysis of stock volatility has already shown in the 1970ies that volatility is not constant, quite the contrary, it is itself stochastic and varies in time.

This observation has led to a vast number of volatility models in discrete-time as well as in continuous-time. In this paper we concentrate on continuous-time volatility models. Moreover, we are concerned with the so-called stylized facts of volatility as e. g.

- volatility changes in time,
- volatility is random,
- volatility has heavy tails,
- volatility clusters on high levels.

Introducing a stochastic volatility extends the Black-Scholes model to

$$dS_t = rS_t dt + \sqrt{V_t} S_t dB_t,$$

where  $V$  can in principle be any positive stationary stochastic process.

Within the framework of SDEs quite natural models are easily defined. Common examples are the *Ornstein-Uhlenbeck (OU) process*

$$dV_t = -\lambda V_t dt + \sigma dZ_t, \quad (1.2)$$

where  $\lambda, \sigma > 0$  and  $Z$  is the driving process, often a second Brownian motion, independent of  $B$ . As this is a Gaussian model, it is not a positive process. Alternatively, a *Cox-Ingersoll-Ross (CIR) model* has been suggested as a volatility model, defined by

$$dV_t = \lambda(a - V_t)dt + \sigma\sqrt{V_t}dZ_t, \quad (1.3)$$

where  $\lambda, a, \sigma > 0$  and  $\lambda a \geq \sigma^2/2$ . The parameter  $a$  is the long-term mean of the process and  $\lambda$  the rate of mean reversion. Again in the classical model  $Z$  is a standard Brownian motion, independent of  $B$ .

Apart from the fact that Gaussian OU processes are not positive, another stylized fact is also violated: empirical volatility exhibits heavy tails, consequently, again the OU model as a Gaussian model seems not very appropriate. Changing the constant  $\sigma$  to a time dependent diffusion coefficient  $\sigma V_t^\gamma$  for  $\gamma \in [1/2, \infty)$  and including a linear drift yields positive stationary models with arbitrarily heavy tails. This has been shown in Borkovec and Klüppelberg [8]. Such models are called *generalized Cox-Ingersoll-Ross models*, a parameter  $\gamma = 1/2$  corresponds to the classical CIR model of (1.3).

On the other hand, a constant  $\sigma$  is attractive, and an alternative way to generate heavy tails in the volatility is to replace the driving Gaussian process in (1.2) by a Lévy process with heavier tailed increments. Furthermore, the upward jumps often observed in empirical volatility cannot be modelled by a continuous process. So Lévy processes with jumps as driving processes seem to be quite natural. Such an OU process is positive, provided the driving Lévy process has only positive increments and no Gaussian component; i. e. it is a subordinator. This is exactly what Barndorff-Nielsen and Shephard [4,5] have suggested, modelling the (right-continuous) volatility process as a Lévy driven *OU process*. Their stochastic volatility model is given by

$$dV_t = -\lambda V_t dt + \sigma dL_{\lambda t}, \quad (1.4)$$

where  $a, b \in \mathbb{R}$ ,  $\lambda > 0$  and  $L$  is a subordinator, called the *background driving Lévy process* (BDLP). The price process itself is then driven by an independent Brownian motion.

A completely different approach to obtain continuous-time volatility models starts with a GARCH model and derives from this discrete-time model a continuous-time model. A natural idea is a diffusion approximation; see e. g. Drost and Werker [16] and references therein. This approach leads to stochastic volatility models of the type

$$\begin{aligned} dS_t &= \sqrt{V_t} dB_t^{(1)}, \\ dV_t &= \lambda(a - V_t)dt + \sigma V_t dB_t^{(2)}, \end{aligned} \quad (1.5)$$

i. e.  $V$  is a generalized CIR model with parameter  $\gamma = 1$ . The two processes  $B^{(1)}, B^{(2)}$  are independent Brownian motions.

A different approach has been taken by Klüppelberg, Lindner and Maller [27], who started with a discrete-time GARCH(1,1) model and replaced the noise variables by a Lévy process  $L$  with jumps  $\Delta L_t = L_t - L_{t-}$ ,  $t \geq 0$ . This yields a stochastic volatility model of the type

$$\begin{aligned} dS_t &= \sqrt{V_t} dL_t, \\ dV_{t+} &= \beta dt + V_t e^{X_t} d(e^{-X_t}), \end{aligned} \tag{1.6}$$

where  $\beta > 0$  and  $V$  is left-continuous. The auxiliary càdlàg process  $X$  is defined by

$$X_t = \eta t - \sum_{0 < s \leq t} \log(1 + \lambda e^\eta (\Delta L_s)^2), \tag{1.7}$$

for  $\eta > 0$  and  $\lambda \geq 0$ . This continuous-time GARCH(1,1) model is called a COGARCH(1,1) model.

Our paper focuses on the extremal behavior of stationary continuous-time stochastic volatility models. This can be described by the tail behavior of the stationary distribution and by the behavior of the process above high thresholds.

The tail behavior models the size of the fluctuations of  $V$  and determines the maximum domain of attraction (MDA) of the model. The notion of MDA is defined in Fisher-Tippett's theorem; see Theorem A.5. We distinguish  $\text{MDA}(\Phi_\alpha)$ ,  $\text{MDA}(\Lambda)$  and  $\text{MDA}(\Psi_\alpha)$ , for  $\alpha > 0$ , respectively. Distribution functions in  $\text{MDA}(\Phi_\alpha)$  have regularly varying tails: they are heavy-tailed in the sense that not all moments are finite; see Definition A.1. Distribution functions in  $\text{MDA}(\Lambda)$  have tails ranging from semi-heavy tails to very light tails. Distribution functions in  $\text{MDA}(\Psi_\alpha)$  have support bounded to the right. Financial risk is usually considered as having unbounded support above, hence  $\text{MDA}(\Psi_\alpha)$  is inappropriate in our context and will play no further role in this paper.

The description of a continuous-time process above high thresholds depends on the sample path behavior of the process. When classical volatility models driven by a Brownian motion have continuous sample paths with infinite variation, some discrete-time skeleton is introduced. A standard concept is based on so-called  $\epsilon$ -upcrossings, see Definition 2.4, which is only valid for processes with continuous sample paths.

For Lévy driven models large jumps (for instance larger than 1) constitute a natural discrete-time skeleton, which can be utilized. Denote by  $(\Gamma_k)_{k \in \mathbb{N}}$  the random time points on  $[0, \infty)$ , where the driving Lévy process jumps and exceeds a given threshold. The bivariate process  $(\Gamma_k, V_{\Gamma_k})_{k \in \mathbb{N}}$  is interpreted as the coordinates of a

point process in  $[0, \infty) \times \mathbb{R}_+$ . As usual we define point processes via Dirac measures. Recall that for any Borel sets  $A \times B \subseteq [0, \infty) \times \mathbb{R}_+$  the measure  $\sum_{k=1}^{\infty} \varepsilon\{\Gamma_k, V_{\Gamma_k}\}(A \times B)$  counts how often  $\Gamma_k \in A$  and  $V_{\Gamma_k} \in B$ .

After appropriate normalization in time and space these point processes may converge and the limit process may allow for an interpretation, thus providing a description of the extreme behavior of the volatility process. Under weak dependence in the data we obtain as limit a Poisson random measure with mean measure  $\vartheta$  (PRM( $\vartheta$ )); see Definition B.4. Moreover, the two components of  $\vartheta$  are independent and consist of the Poisson measure in time and the negative logarithm of an extreme value distribution in space. Under strong dependence the limit is a cluster Poisson random measure. All these considerations concern the discrete-time skeleton only and ignore the fact that we deal with continuous-time processes.

In the case of a driving Lévy process with jumps, in principle also the small jumps can influence the extreme behavior. In a very close neighborhood of a jump time  $\Gamma_k$  infinitely many small jumps can happen; they may contribute to the extreme behavior around  $\Gamma_k$ . To investigate the influence of these small jumps and the Gaussian component, we consider the process  $V$  at each point  $\Gamma_k$  in a surrounding interval  $I_k$ . Finally, in certain situations we investigate also the process  $V$  after it has achieved a local supremum. With each point  $\Gamma_k$  an excursion of  $V$  over a high threshold starts. Interesting questions concern the length of the excursion, the rate of “decrease” after  $\Gamma_k$ , and we shall answer such questions at least for some models considered in this paper. This is done by attaching marks to the point process  $(\Gamma_k, V_{\Gamma_k})_{k \in \mathbb{N}}$ . In our model marks are a vector of values of the process  $V$  after  $\Gamma_k$ , hence it describes the finite dimensional distributions of  $V$  after  $\Gamma_k$ . The limit process turns out to be different in different regimes.

Our paper is organized as follows. In Section 2 we review the extremal behavior of the generalized CIR model, which can belong to different maximum domains of attractions; i. e. such models can have arbitrary tails. Unfortunately, they are not appropriate models in case of high level volatility clusters in the data.

Section 3 deals with Lévy driven OU volatility models. Their extremal behavior is characterized by the extremal behavior of the driving Lévy process, so that we have to distinguish between different classes of BDLPes. In Section 3.1 this is done for subexponential Lévy processes  $L = (L_t)_{t \geq 0}$ . According to whether  $L_1 \in \text{MDA}(\Phi_\alpha)$  for some  $\alpha > 0$  or  $L_1 \in \text{MDA}(\Lambda)$ , the extremal behavior of the OU process is quite different. Then, in Section 3.2 we study OU processes with exponential tails. As a prominent example we investigate the  $\Gamma$ -OU process, i. e. the stationary volatility

is gamma distributed. As an important larger class we study OU processes, whose BDLP belongs to  $\mathcal{S}(\gamma)$  for  $\gamma > 0$ . This class extends subexponential Lévy processes in a natural way; see Definition A.3. It turns out that for all OU processes in Section 3, high level volatility clusters are exhibited only in the case of regularly varying BDLPes.

The last class of models reviewed in this paper concerns the COGARCH process in Section 4. In contrast to the OU processes considered earlier, the COGARCH volatility has heavy tails under quite general conditions on the driving Lévy process  $L$ . Furthermore, the COGARCH exhibits high level volatility clusters.

Finally, a short conclusion is given in Chapter 5. Here we compare the models introduced in the different sections before. It turns out that there is a striking similarity concerning the extremal behaviour of models with the same stationary distribution. Here we also discuss briefly some further empirical facts of volatility not quite in the focus of our paper.

As not to disturb the flow of arguments we postpone classical definitions and concepts to an Appendix. Throughout this paper we shall use the following notation. We abbreviate distribution function by d.f. and random variable by r.v. For any d.f.  $F$  we denote its tail  $\bar{F} = 1 - F$ . For two r.v.s  $X$  and  $Y$  with d.f.s  $F$  and  $G$  we write  $X \stackrel{d}{=} Y$  if  $F = G$ , and by  $\xrightarrow{T \rightarrow \infty} \Rightarrow$  we denote weak convergence for  $T \rightarrow \infty$ . For two functions  $f$  and  $g$  we write  $f(x) \sim g(x)$  as  $x \rightarrow \infty$ , if  $\lim_{x \rightarrow \infty} f(x)/g(x) = 1$ . We also denote  $\mathbb{R}_+ = (0, \infty)$ . For  $x \in \mathbb{R}$ , we let  $x^+ = \max\{x, 0\}$  and  $\log^+(x) = \log(\max\{x, 1\})$ . Integrals of the form  $\int_a^b$  will be interpreted as the integral taken over the interval  $(a, b]$ .

## 2 Extremal behavior of generalized Cox-Ingersoll-Ross models

In this section we summarize some well-known results on classical volatility models as defined in (1.3) and (1.5) driven by a standard Brownian motion. This section is based on Borkovec and Klüppelberg [8]; for a review see [26], Section 3.

These models fall into the framework of *generalized Cox-Ingersoll-Ross models* (GCIR) models. We restrict ourselves to stationary solutions of the SDE

$$dV_t = \lambda(a - V_t)dt + \sigma V_t^\gamma dB_t, \tag{2.1}$$

where  $\gamma \in [\frac{1}{2}, \infty)$ . For  $\lambda, a, \sigma > 0$  (in the case  $\gamma = 1/2$  additionally  $\lambda a \geq \sigma^2/2$  is needed) these models are ergodic with state space  $\mathbb{R}_+$  and have a stationary density.

Associated with the diffusion (2.1) is the scale function  $s$  and the speed measure  $m$ . The *scale function* is defined as

$$s(x) = \int_z^x \exp\left(-\frac{2\lambda}{\sigma^2} \int_z^y \frac{a-t}{t^{2\gamma}} dt\right) dy, \quad x \in \mathbb{R}_+, \quad (2.2)$$

where  $z$  is any interior point of  $\mathbb{R}_+$  whose choice does not affect the extremal behavior. For the *speed measure*  $m$  we know that it is finite for the GCIR model. Moreover,  $m$  is absolutely continuous with Lebesgue density

$$m'(x) = \frac{2}{\sigma^2 x^{2\gamma} s'(x)}, \quad x \in \mathbb{R}_+,$$

where  $s'$  is the Lebesgue density of  $s$ . Then the stationary density of  $V$  is given by

$$f(x) = m'(x)/m(\mathbb{R}_+), \quad x \in \mathbb{R}_+. \quad (2.3)$$

**Proposition 2.1.** *Let  $V$  be a GCIR model given by equation (2.1) and define  $M(T) := \sup_{t \in [0, T]} V_t$  for  $T > 0$ . Then for any initial value  $V_0 = y \in \mathbb{R}_+$  with corresponding distribution  $\mathbb{P}_y$  and any  $u_T \uparrow \infty$ ,*

$$\lim_{T \rightarrow \infty} |\mathbb{P}_y(M(T) \leq u_T) - H^T(u_T)| = 0,$$

where  $H$  is a d.f., defined for any  $z \in \mathbb{R}_+$  by

$$H(x) = \exp\left(-\frac{1}{m(\mathbb{R}_+)s(x)}\right), \quad x > z. \quad (2.4)$$

The function  $s$  and the quantity  $m(\mathbb{R}_+)$  depend on the choice of  $z$ . □

**Corollary 2.2 (Running maxima).**

*Let the assumptions of Proposition 2.1 hold. Assume further that  $H \in \text{MDA}(G)$  for  $G \in \{\Phi_\alpha, \alpha > 0, \Lambda\}$  with norming constants  $a_T > 0, b_T \in \mathbb{R}$ . Then*

$$\lim_{T \rightarrow \infty} \mathbb{P}(a_T^{-1}(M(T) - b_T) \leq x) = G(x), \quad x \in \mathbb{R}.$$

It is clear that the d.f.  $H$  decides about the extremal behavior of  $V$ . We present four cases.

**Example 2.3 (Tail behavior of GCIR models).**

Let  $V$  be a stationary GCIR model given by equation (2.1) with stationary density  $f$ ,



corresponding d. f.  $F$ , and d. f.  $H$  as given in (2.4). Recall that a  $\Gamma(\mu, \gamma)$  distributed r. v. has probability density

$$p(x) = \frac{\gamma^\mu}{\Gamma(\mu)} x^{\mu-1} e^{-\gamma x}, \quad x > 0, \quad (2.5)$$

for  $\mu > 1$  and  $\gamma > 0$ .

(1)  $\gamma = \frac{1}{2}$ : The stationary density of  $V$  is  $\Gamma\left(\frac{2\lambda a}{\sigma^2}, \frac{2\lambda}{\sigma^2}\right)$ . The tail of  $H$  behaves like

$$\overline{H}(x) \sim \frac{2\lambda^2 a}{\sigma^2} x \overline{F}(x), \quad x \rightarrow \infty,$$

so that the tail of  $H$  is that of a  $\Gamma\left(\frac{2\lambda a}{\sigma^2} + 1, \frac{2\lambda}{\sigma^2}\right)$  distribution. Hence  $H \in \text{MDA}(\Lambda)$  with norming constants

$$a_T = \frac{\sigma^2}{2\lambda} \quad \text{and} \quad b_T = \frac{\sigma^2}{2\lambda} \left[ \log T + \frac{2\lambda a}{\sigma^2} \log \log T - \log \left( \frac{\lambda}{\Gamma(2\lambda a/\sigma^2)} \right) \right].$$

(2)  $\frac{1}{2} < \gamma < 1$ : The stationary density of  $V$  is given by

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

with some normalizing factor  $A > 0$ . The d. f.  $H$  has tail

$$\overline{H}(x) \sim Bx^{2(1-\gamma)} \overline{F}(x), \quad x \rightarrow \infty, \quad (2.6)$$

where  $B > 0$ . Hence  $H \in \text{MDA}(\Lambda)$  with norming constants

$$a_T = \frac{\sigma^2}{2\lambda} \left( \frac{\sigma^2(1-\gamma)}{\lambda} \log T \right)^{\frac{2\gamma-1}{2-2\gamma}},$$

$$b_T = \left( \frac{\sigma^2(1-\gamma)}{\lambda} \log T \right)^{\frac{1}{(2-2\gamma)}} \left( 1 - \frac{2\gamma-1}{(2-2\gamma)^2} \frac{\log \left( \frac{\sigma^2(1-\gamma)}{\lambda} \log T \right)}{\log T} \right) + a_T \log \left( \frac{2\lambda}{A\sigma^2} \right).$$

(3)  $\gamma = 1$ : The stationary density of  $V$  is inverse gamma, i. e.

$$f(x) = \left( \frac{\sigma^2}{2\lambda a} \right)^{-\frac{2\lambda}{\sigma^2}-1} \left( \Gamma \left( \frac{2\lambda}{\sigma^2} + 1 \right) \right)^{-1} x^{-\frac{2\lambda}{\sigma^2}-2} \exp \left( -\frac{2\lambda a}{\sigma^2} x^{-1} \right), \quad x > 0,$$

so that  $V_0 \in \mathcal{R}_{-2\lambda/\sigma^2-1}$ ; see Definition A.1. In this case  $\overline{H}(x) \sim cx^{-2\lambda/\sigma^2-1}$  for  $x \rightarrow \infty$  and some  $c > 0$ . Hence  $H \in \text{MDA}(\Phi_\alpha)$  for  $\alpha = 2\lambda/\sigma^2 + 1$  with norming constants  $a_T = (cT)^{\sigma^2/(2\lambda+\sigma^2)}$  and  $b_T = 0$ .

(4)  $\gamma > 1$ : The stationary density  $f$  of  $V$  has the same form as in (2), but is regularly varying of index  $-2\gamma + 1$ . Now the tail of  $H$  becomes very extreme:  $\overline{H}(x) \sim cx^{-1}$ . Hence  $H \in \text{MDA}(\Phi_1)$  with  $a_T = cT$  and  $b_T = 0$ .  $\square$

Since all models (2.1) are driven by a Brownian motion, they have continuous sample paths; i. e. there is no natural discrete-time skeleton. We follow the standard approach to create a discrete-time skeleton of the process; see e. g. Leadbetter et al. [29], Chapter 12. This allows for a more profound extreme value analysis of  $V$ .

**Definition 2.4.** *Let  $V$  be a stationary version of the diffusion given by (2.1).  $V$  is said to have an  $\epsilon$ -upcrossing of the level  $u$  at a point  $\Gamma > 0$  if*

$$V_t < u \quad \text{for } t \in (\Gamma - \epsilon, \Gamma) \quad \text{and} \quad V_\Gamma = u.$$

With this definition we can formulate a further result describing the extreme behavior of a stationary GCIR model.

**Theorem 2.5 (Point process of  $\epsilon$ -upcrossings).**

*Let  $V$  be a stationary version of the diffusion given by (2.1) with d. f.  $H$  as in (2.4). Let  $a_T > 0$ ,  $b_T \in \mathbb{R}$  be the norming constants as given in Example 2.3. Let  $(\Gamma_{T,k})_{k \in \mathbb{N}}$  be the time points on  $\mathbb{R}_+$ , where the  $\epsilon$ -upcrossings of  $V$  of the level  $a_T x + b_T$  occur. Let  $(j_k)_{k \in \mathbb{N}}$  be the jump times of a Poisson process with intensity  $e^{-x}$  for  $x \in \mathbb{R}$ , if  $\gamma \in [1/2, 1)$  and  $x^{-\alpha}$  for  $x > 0$  with  $\alpha = 2\lambda/\sigma^2 + 1$  if  $\gamma = 1$  and  $\alpha = 1$  if  $\gamma > 1$ . Then*

$$\sum_{k=1}^{\infty} \varepsilon\{\Gamma_{T,k}/T\} \xrightarrow{T \rightarrow \infty} \sum_{k=1}^{\infty} \varepsilon\{j_k\}.$$

As is obvious from this result,  $\epsilon$ -upcrossings of  $V$  for high levels behave like exceedances of i. i. d. data, i. e. such models do not exhibit volatility clusters. They can, however, model heavy tails as the running maxima depend on the d. f.  $H$ .

### 3 Extremal behavior of OU volatility models

We start with a precise definition of a positive OU process as a solution of (1.4). For more information on Lévy processes we refer to the excellent monographs by Sato [35], Bertoin [6] and Cont and Tankov [15]. Let  $L$  be a subordinator; i. e.  $L$  is a Lévy process with increasing sample paths, hence they are of bounded variation, and we assume that they are càdlàg. The Laplace transform is then the natural transform and has for all  $t \geq 0$  the representation

$$\mathbb{E} \exp(-\lambda L_t) = \exp(-t\Psi(\lambda)), \quad \lambda \geq 0.$$

The Laplace exponent  $\Psi$  has representation

$$\Psi(\lambda) = m\lambda + \int_{(0,\infty)} (1 - e^{-\lambda x}) \nu(dx).$$

As there is no Gaussian component the characteristic triplet of arbitrary Lévy processes reduces to a pair  $(m, \nu)$ , where  $m > 0$  is the drift and the Lévy measure  $\nu$  has support on  $\mathbb{R}_+$  and satisfies

$$\int_{(0,\infty)} (1 \wedge x) \nu(dx) < \infty.$$

We denote for  $\lambda > 0$  by

$$V_t = e^{-\lambda t} V_0 + \int_0^t e^{-\lambda(t-s)} dL_{\lambda s}, \quad t \in \mathbb{R}, \quad (3.1)$$

the solution to the SDE (1.4). Then  $V$  becomes a càdlàg process.

A stationary solution of (1.4) exists if and only if  $\int_{s>1} \log(1+s) \nu(ds) < \infty$ . Note that this condition is only violated for Lévy measures with very heavy tails. As all models considered in this paper have tails which are regularly varying of some negative index or lighter, all our models satisfy this stationarity condition. Stationarity is then achieved, if  $V_0$  is taken to be independent of the driving Lévy process  $L$  and has distribution

$$V_0 \stackrel{d}{=} \int_0^\infty e^{-s} dL_s. \quad (3.2)$$

A convenient representation for the stationary version is

$$V_t = e^{-\lambda t} \int_{-\infty}^t e^{\lambda s} dL_{\lambda s}, \quad t \geq 0. \quad (3.3)$$

In this representation,  $L$  is extended to a Lévy process on the whole real line, by letting  $\tilde{L} = (\tilde{L}_t)_{t \geq 0}$  be an independent copy of  $(L_t)_{t \geq 0}$ , and defining  $L_t := -\tilde{L}_{-t}$  for  $t < 0$ . The parameter  $\lambda$  in the process  $L$  in (3.1) ensures that the stationary marginal distribution of  $V$  is independent of  $\lambda$ ; indeed it is given by (3.2).

The r. v.  $V_0$  is infinitely divisible with characteristic pair  $(m_V, \nu_V)$ , where  $m_V = m$  and

$$\nu_V[x, \infty) = \int_x^\infty u^{-1} \nu[u, \infty) du, \quad x > 0. \quad (3.4)$$

We are concerned with Lévy processes  $L$ , which are heavy or semi-heavy tailed; i. e. whose tails decrease not faster than exponentially. As indicated in (3.8) and

(3.11) below, this induces a similar tail behavior on  $V$ , which is in accordance with empirical findings.

The structure of an OU volatility process can be best understood when considering the following example.

**Example 3.1 (Positive shot noise process).**

Let  $L$  be a positive compound Poisson process with characteristic pair  $(0, \mu \mathbb{P}_F)$ , where  $\mu > 0$  and  $\mathbb{P}_F$  is a probability measure on  $\mathbb{R}_+$  with corresponding d.f.  $F$ . Then  $L$  has the representation

$$L_t = \sum_{j=1}^{N_t} \xi_j, \quad t > 0, \quad (3.5)$$

where  $(N_t)_{t \geq 0}$  is a Poisson process on  $\mathbb{R}_+$  with intensity  $\mu > 0$  and jump times  $(\Gamma_k)_{k \in \mathbb{N}}$ . The process  $N$  is independent of the i.i.d. sequence of positive r.v.s  $(\xi_k)_{k \in \mathbb{N}}$  with d.f.  $F$ .

The resulting volatility process is then the positive shot noise process

$$V_t = e^{-\lambda t} V_0 + \int_0^t e^{-\lambda(t-s)} dL_{\lambda s} = e^{-\lambda t} V_0 + \sum_{j=1}^{N_{\lambda t}} e^{-\lambda t + \Gamma_j} \xi_j, \quad t \geq 0,$$

and from (3.4) we get

$$\nu_V[x, \infty) = \mu \int_x^\infty u^{-1} \bar{F}(u) du, \quad x > 0.$$

If  $\mathbb{E} \log(1 + \xi_1) < \infty$ , a stationary solution exists in which case  $V$  can be represented as

$$V_t = e^{-\lambda t} \sum_{\substack{j=-\infty \\ j \neq 0}}^{N_{\lambda t}} e^{\Gamma_j} \xi_j, \quad t > 0. \quad (3.6)$$

Here, we let  $(\xi_k)_{k \in -\mathbb{N}_0}$  and  $(\Gamma_k)_{k \in -\mathbb{N}_0}$  be sequences of r.v.s such that  $(\xi_k)_{k \in \mathbb{Z}}$  and  $(\Gamma_k)_{k \in \mathbb{Z}}$  are independent. Furthermore,  $(\xi_k)_{k \in \mathbb{Z}}$  is an i.i.d. sequence and  $(-\Gamma_k)_{k \in -\mathbb{N}}$  are the jump times of a Poisson process on  $\mathbb{R}_+$  with intensity  $\mu$ , independent of  $(\Gamma_k)_{k \in \mathbb{N}}$ ; further, we define  $\Gamma_0 := 0$ .

The qualitative extreme behavior of this volatility process can be seen in Figure 1 in detail. The volatility jumps upwards, whenever  $(N_{\lambda t})_{t \geq 0}$  jumps and decreases exponentially fast between two jumps. This means in particular that  $V$  has local suprema exactly at the jump times  $\Gamma_k/\lambda$  (and  $t = 0$ ), i.e

$$V_t = V_{\Gamma_k/\lambda} e^{-\lambda t + \Gamma_k}, \quad t \in [\Gamma_k/\lambda, \Gamma_{k+1}/\lambda).$$

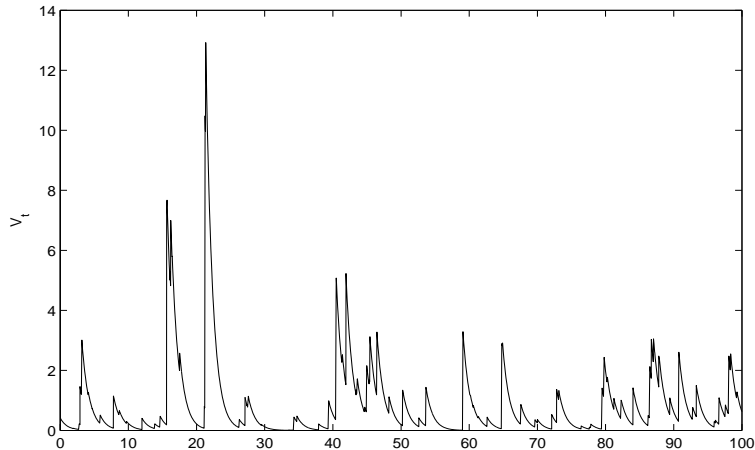


Figure 1: Sample path of an OU Weibull process with  $L_1$  as given in Example 3.3 with  $\lambda = 1$ ,  $\mu = 1$  and  $p = 1/2$ .

Consequently, it is the discrete-time skeleton of  $V$  at points  $\Gamma_k/\lambda$  that determines the extreme behavior of the volatility process.  $\square$

For a general subordinator  $L$  we decompose

$$L = L^{(1)} + L^{(2)} \quad (3.7)$$

into two independent Lévy processes, where  $L^{(1)}$  has characteristic pair  $(0, \nu_1)$  with  $\nu_1(x, \infty) = \nu(x, \infty) \mathbf{1}_{(1, \infty)}(x)$  and  $L^{(2)}$  has characteristic pair  $(m, \nu_2)$  with  $\nu_2(x, \infty) = \nu(x, 1] \mathbf{1}_{(0, 1]}(x)$ . Then again  $L^{(1)}$  is a compound Poisson process with intensity  $\nu(1, \infty)$  and jump sizes with d.f.  $\nu_1/\nu(1, \infty)$ . All the small jumps and the drift are summarized in  $L^{(2)}$ .

What is needed, however, are the precise asymptotic links between the tails of  $V$ ,  $L$  and the tail of the Lévy measure  $\nu(\cdot, \infty)$ . This implies immediately that we have to distinguish different regimes defined by the link between  $L$  and  $\nu(\cdot, \infty)$ . Any infinitely divisible distribution is asymptotically tail-equivalent to its Lévy measure, whenever it is convolution equivalent; see Theorem A.4. Definitions and results concerning subexponential and convolution equivalent distributions are summarized in Appendix 5.

The class  $\mathcal{S}(0) = \mathcal{S}$  of subexponential d.f.s contains all d.f.s with regularly varying tails, but is much larger. Subexponential distributions can be in two different maximum domains of attractions; see Theorem A.5. All d.f.s with regularly varying tails are subexponential and belong to  $\text{MDA}(\Phi_\alpha)$ . Other subexponential d.f.s, as for instance the lognormal and the semi-heavy tailed Weibull d.f.s (see Example 3.3),

belong to  $\text{MDA}(\Lambda)$ . On the other hand, d. f. s as the gamma distribution or d. f. s in  $\mathcal{S}(\gamma)$  for  $\gamma > 0$  belong to  $\text{MDA}(\Lambda)$ , but are lighter tailed than any subexponential distribution. Consequently, we also consider such exponential models below.

### 3.1 OU processes with subexponential tails

In this section we are concerned with the OU process given by (3.1), whose BDLP is subexponential. This section is based on Fasen [19, 20]; an additional reference is [18].

**Proposition 3.2 (Tail behavior of subexponential models).**

Let  $V$  be a stationary version of the OU process given by (3.1) and define  $M(h) := \sup_{t \in [0, h]} V_t$  for  $h > 0$ .

(a) If  $L_1 \in \mathcal{S} \cap \text{MDA}(\Phi_\alpha) = \mathcal{R}_{-\alpha}$ , then also  $V_0 \in \mathcal{R}_{-\alpha}$  and

$$\mathbb{P}(V_0 > x) \sim \alpha^{-1} \mathbb{P}(L_1 > x), \quad x \rightarrow \infty. \quad (3.8)$$

Moreover,

$$\mathbb{P}(M(h) > x) \sim [\lambda h + \alpha^{-1}] \mathbb{P}(L_1 > x), \quad x \rightarrow \infty. \quad (3.9)$$

(b) If  $L_1 \in \mathcal{S} \cap \text{MDA}(\Lambda)$ , then also  $V_0 \in \mathcal{S} \cap \text{MDA}(\Lambda)$  and

$$\mathbb{P}(V_0 > x) \sim \mathbb{P}(\exp(-U)L_1 > x), \quad x \rightarrow \infty, \quad (3.10)$$

where  $U$  is a uniform r. v. on  $(0, 1)$ , independent of  $L$ . In particular,  $\mathbb{P}(V_0 > x) = o(\mathbb{P}(L_1 > x))$  as  $x \rightarrow \infty$ . More precisely,

$$\mathbb{P}(V_0 > x) \sim \frac{a(x)}{x} \mathbb{P}(L_1 > x), \quad x \rightarrow \infty, \quad (3.11)$$

where  $a$  is the function from the representation (A.1):

$$\mathbb{P}(L_1 > x) \sim c \exp \left[ - \int_0^x \frac{1}{a(y)} dy \right], \quad x \rightarrow \infty,$$

for some  $c > 0$  and  $a : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is absolutely continuous with  $\lim_{x \rightarrow \infty} a'(x) = 0$  and  $\lim_{x \rightarrow \infty} a(x) = \infty$ . Finally,

$$\mathbb{P}(M(h) > x) \sim \lambda h \mathbb{P}(L_1 > x), \quad x \rightarrow \infty. \quad (3.12)$$

*Proof.* By (3.4) we have

$$\frac{\nu_V(x, \infty)}{\nu(x, \infty)} = \frac{\int_x^\infty u^{-1} \nu(u, \infty) du}{\nu(x, \infty)}, \quad x > 0. \quad (3.13)$$

Assume that  $L_1 \in \mathcal{R}_{-\alpha}$  for some  $\alpha > 0$ . Then by Theorem A.4 (i) we have  $\nu(\cdot, \infty)/\nu(1, \infty) \in \mathcal{R}_{-\alpha}$ . By Karamata's theorem (e.g. Embrechts et al. [17], Theorem A 3.6) we obtain

$$\lim_{x \rightarrow \infty} \frac{\nu_V(x, \infty)}{\nu(x, \infty)} = \frac{1}{\alpha}.$$

This implies in particular that also  $\nu_V(\cdot, \infty)/\nu_V(1, \infty) \in \mathcal{R}_{-\alpha}$  and hence, again by Theorem A.4 (i),  $V_0 \in \mathcal{R}_{-\alpha}$  and (3.8) holds.

If  $L_1 \in \text{MDA}(\Lambda) \cap \mathcal{S}$  we can only conclude from (3.13) that the right hand side tends to 0. To obtain a precise result we proceed as follows. Denote by  $\xi_1$  the jump distribution of the compound Poisson process  $L^{(1)}$  as given in (3.7). Taking  $\nu(\cdot, \infty)/\nu(1, \infty) \in \mathcal{R}_{-\infty}$  into account and applying l'Hospital's rule yields

$$\begin{aligned} \frac{\nu_V(x, \infty)}{\nu(1, \infty) \mathbb{P}(\exp(-U)\xi_1 > x)} &= \frac{\int_x^\infty u^{-1} \nu(u, \infty) du}{\int_0^1 \nu(e^s x, \infty) ds} = \frac{\int_x^\infty u^{-1} \nu(u, \infty) du}{\int_x^{xe} u^{-1} \nu(u, \infty) du} \\ &\sim \left[ 1 - \frac{\nu(ex, \infty)}{\nu(x, \infty)} \right]^{-1} \longrightarrow 1, \quad x \rightarrow \infty. \end{aligned}$$

The tail-equivalence (3.10) follows then from the fact that  $\nu(1, \infty) \mathbb{P}(\exp(-U)\xi_1 > x) \sim \mathbb{P}(\exp(-U)L_1 > x)$  as  $x \rightarrow \infty$  and Theorem A.4 (i).

For proving (3.11), by Theorem A.4 (i) we may assume without loss of generality that there exists a  $x_0 > 0$  such that

$$\nu(x, \infty) = c \exp \left[ - \int_{x_0}^x \frac{1}{a(y)} dy \right], \quad x \geq x_0.$$

Then  $\nu(dx) = (a(x))^{-1} \nu(x, \infty) dx$  and an application of l'Hospital's rule shows that

$$\begin{aligned} \frac{\nu_V(x, \infty)}{\nu(x, \infty) a(x)/x} &\sim \frac{-\nu(x, \infty)/x}{\nu(x, \infty)[a'(x) - a(x)/x]/x - (\nu(x, \infty)/a(x))a(x)/x} \\ &= [-a'(x) + a(x)/x + 1]^{-1} \\ &\rightarrow 1, \quad x \rightarrow \infty, \end{aligned}$$

since  $a(x)/x \sim a'(x)$  and  $a'(x) \rightarrow 0$  as  $x \rightarrow \infty$ . Theorem A.4 (i) then gives (3.11).

The results for  $M(h)$  are based on Theorem 2.1 of Rosinski and Samorodnitsky [34]. They show that for  $L_{\lambda h} + V_0 \in \mathcal{S}$

$$\mathbb{P}(M(h) > x) \sim \nu_{L_{\lambda h} + V_0}(x, \infty), \quad x \rightarrow \infty,$$

implying the result by Theorem A.4 (i).  $\square$

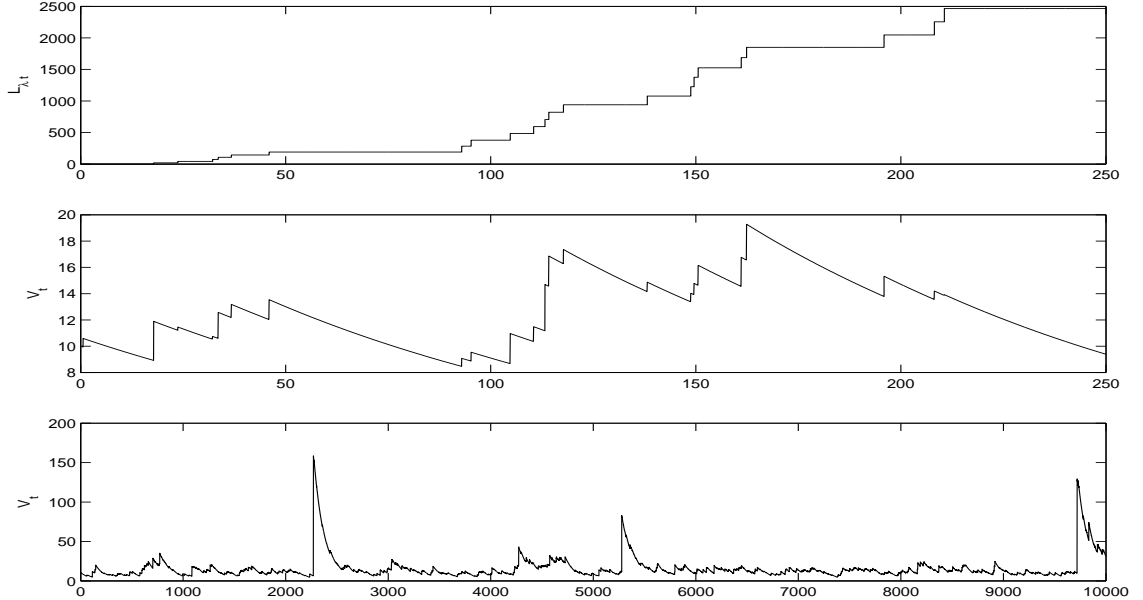


Figure 2: Sample path of an OU process driven by a regularly varying compound Poisson process with  $\mu = 8.5$  and  $\lambda = 0.01$ . The first plot shows  $(L_{\lambda t})_{0 \leq t \leq 250}$  and the second plot shows the corresponding volatility  $(V_t)_{0 \leq t \leq 250}$ , indicating the micro-behavior of the model. The third plot gives  $(V_t)_{0 \leq t \leq 10000}$  indicating the macro-behavior of the model.

### Example 3.3 (Semi-heavy tailed Weibull distribution).

Let  $L_1$  have distribution tail

$$\mathbb{P}(L_1 > x) \sim K \exp(-x^p), \quad x \rightarrow \infty,$$

for some  $K > 0$  and  $p \in (0, 1)$ . As  $a(x) = x^{1-p}/p$ , we obtain from (3.11) immediately

$$\mathbb{P}(V_0 > x) \sim \frac{K}{p} x^{-p} \exp(-x^p), \quad x \rightarrow \infty. \quad \square$$

Proposition 3.2 shows that in the regularly varying regime the tail of  $V_0$  is equivalent to the tail of  $L_1$ . In contrast to that, in the  $\mathcal{S} \cap \text{MDA}(\Lambda)$  case, the tail of  $V_0$  becomes lighter, due to the influence of  $\exp(-U)$ . But in both cases  $V_0$  is subexponential and the tail of  $M(h)$  is determined by the tail of  $L_1$ , only the constants differ.

The following result gives a complete account of the extreme behavior of the volatility process  $V$  driven by a subexponential Lévy process  $L$ . There are three



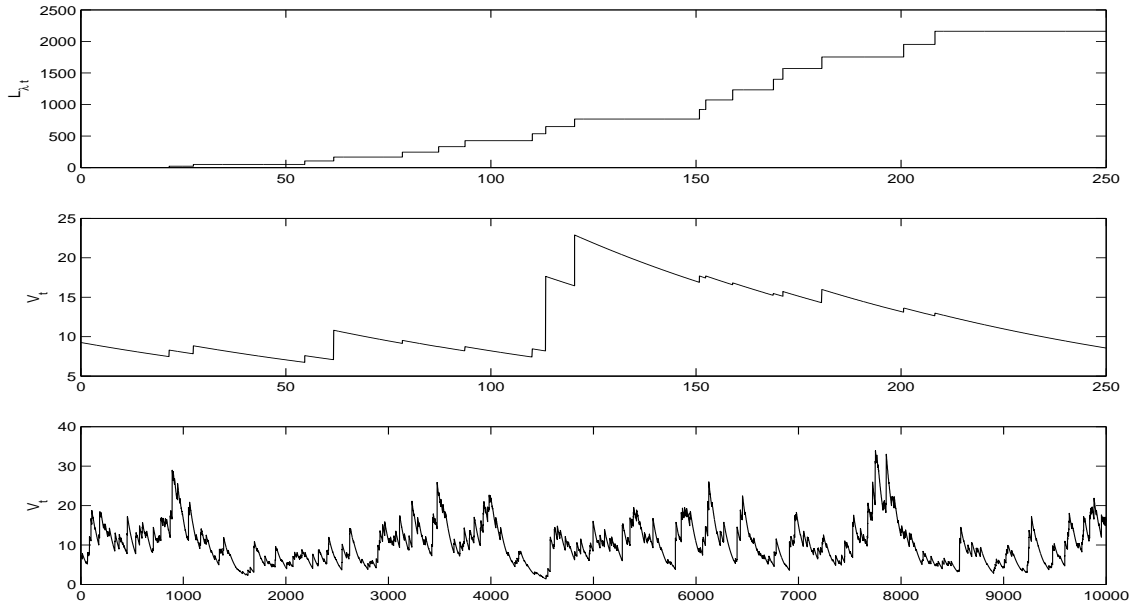


Figure 3: Sample path of an OU Weibull process with  $p = 0.7$ ,  $\mu = 8.5$  and  $\lambda = 0.01$ . The first plot shows  $(L_{\lambda t})_{0 \leq t \leq 250}$  and the second plot the corresponding volatility  $(V_t)_{0 \leq t \leq 250}$ , indicating the micro-behavior of the model. The third plot gives  $(V_t)_{0 \leq t \leq 10\,000}$  indicating the macro-behavior of the model.

components considered in (3.15) and (3.16). The first one is the scaled jump time process corresponding to the jumps of  $(L_{\lambda t})_{t \geq 0}$ , which are larger than 1. The second component is the normalized local supremum near that jump, and the third component is a vector of normalized values of  $V$  after the jump.

**Theorem 3.4 (Marked point process behavior of subexponential models).**

Let  $V$  be a stationary version of the OU process given by (3.1). Suppose  $\Gamma = (\Gamma_k)_{k \in \mathbb{N}}$  are the jump times of  $L^{(1)}$  given by (3.7) and  $I = (I_k)_{k \in \mathbb{N}}$ , where  $I_k = \frac{1}{2\lambda} [\Gamma_{k-1} + \Gamma_k, \Gamma_k + \Gamma_{k+1}]$ ,  $\Gamma_0 := 0$ . For  $m \in \mathbb{N}$  let  $0 = t_0 < t_1 < \dots < t_m$ .

(a) Assume that  $L_1 \in \mathcal{S} \cap \text{MDA}(\Phi_\alpha)$  with norming constants  $a_T > 0$  such that

$$\lim_{T \rightarrow \infty} T\mathbb{P}(L_1 > a_T x) = x^{-\alpha}, \quad x > 0. \quad (3.14)$$

Take  $\Gamma^{(k)} = (\Gamma_{k,i})_{i \in \mathbb{N}}$ ,  $k \in \mathbb{N}$ , as i. i. d. copies of the sequence  $\Gamma$  and set  $\Gamma_{k,0} = \Gamma_{k,-1} = 0$  for all  $k \in \mathbb{N}$ . Let  $\sum_{k=1}^{\infty} \varepsilon\{s_k, P_k\}$  be a PRM( $\vartheta$ ) with mean measure

$\vartheta(dt \times dx) = dt \times \alpha x^{-\alpha-1} \mathbf{1}_{(0,\infty)}(x) dx$ , independent of the sequence  $(\Gamma^{(k)})_{k \in \mathbb{N}}$ . Then,

$$\begin{aligned} & \sum_{k=1}^{\infty} \varepsilon \left\{ \frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1} \sup_{s \in I_k} V_s, \{a_{\lambda T}^{-1} V_{\Gamma_k/\lambda+t_i}\}_{i=0,\dots,m} \right\} \\ & \xrightarrow{T \rightarrow \infty} \sum_{k=1}^{\infty} \sum_{j=0}^{\infty} \varepsilon \{s_k, P_k \exp(-(\Gamma_{k,j-1} + \Gamma_{k,j})/2), \{P_k \exp(-\lambda t_i - \Gamma_{k,j})\}_{i=0,\dots,m}\}. \end{aligned} \quad (3.15)$$

(b) Assume that  $L_1 \in \mathcal{S} \cap \text{MDA}(\Lambda)$  with norming constants  $a_T > 0$ ,  $b_T \in \mathbb{R}$  such that

$$\lim_{T \rightarrow \infty} T\mathbb{P}(L_1 > a_T x + b_T) = \exp(-x), \quad x \in \mathbb{R}.$$

Let  $\sum_{k=1}^{\infty} \varepsilon \{s_k, P_k\}$  be a PRM( $\vartheta$ ) with  $\vartheta(dt \times dx) = dt \times e^{-x} dx$ . Then

$$\begin{aligned} & \sum_{k=1}^{\infty} \varepsilon \left\{ \frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1} (\sup_{s \in I_k} V_s - b_{\lambda T}), \{a_{\lambda T}^{-1} (V_{\Gamma_k/\lambda+t_i} - b_{\lambda T})\}_{i=0,\dots,m} \right\} \\ & \xrightarrow{T \rightarrow \infty} \sum_{k=1}^{\infty} \varepsilon \{s_k, P_k, (P_k, 0, \dots, 0)\}. \end{aligned} \quad (3.16)$$

Moreover,

$$\begin{aligned} & \sum_{k=1}^{\infty} \varepsilon \left\{ \frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1} (\sup_{s \in I_k} V_s - b_{\lambda T}), \{b_{\lambda T}^{-1} V_{\Gamma_k/\lambda+t_i}\}_{i=0,\dots,m} \right\} \\ & \xrightarrow{T \rightarrow \infty} \sum_{k=1}^{\infty} \varepsilon \{s_k, P_k, \{\exp(-\lambda t_i)\}_{i=0,\dots,m}\}. \end{aligned} \quad (3.17)$$

We first give an interpretation of (3.16). The limit relations of the first two components show that the local suprema of  $V$  around the  $\Gamma_k/\lambda$ , normalized with the constants determined via  $L_\lambda$ , converge weakly to the same extreme value distribution as  $L_\lambda$ . Moreover, the third component indicates that for  $t_0 = 0$  the second and third component have the same limiting behavior; i. e. the  $\sup_{s \in I_k} V_s$  behaves like  $V_{\Gamma_k/\lambda}$ . For all  $t_i > 0$  the last component is negligible, i. e. the process is considerably smaller away from  $V_{\Gamma_k/\lambda}$ .

In the second and third component of (3.15) all points  $\Gamma_{k,j}$  and not only  $\Gamma_{k,0} = 0$  like in (3.16) may influence the limit. This phenomenon has certainly its origin in the very large jumps caused by regular variation. Even though the volatility decreases between the jumps exponentially fast, huge jumps can have a long lasting influence on excursions above high thresholds. This is in contrast to the semi-heavy tailed case, where  $L$  is subexponential, but in  $\text{MDA}(\Lambda)$ .

Both relations (3.15) and (3.17) exhibit, however, a common effect in the third component: if the Lévy process  $L$  has an exceedance over a high threshold, then the OU process decreases after this event exponentially fast.

**Corollary 3.5 (Point process of exceedances).**

Let the assumptions of Theorem 3.4 hold.

(a) Assume that  $L_1 \in \mathcal{S} \cap \text{MDA}(\Phi_\alpha)$ . Let  $(j_k)_{k \in \mathbb{N}}$  be the jump times of a Poisson process with intensity  $x^{-\alpha}$  for  $x > 0$ . Let  $(\zeta_k)_{k \in \mathbb{Z}}$  be i. i. d. discrete r. v. s, independent of  $(j_k)_{k \in \mathbb{N}}$ , with probability distribution

$$\pi_k = \mathbb{P}(\zeta_1 = k) = \mathbb{E} \exp(-\alpha(\Gamma_{k-1} + \Gamma_k)/2) - \mathbb{E} \exp(-\alpha(\Gamma_k + \Gamma_{k+1})/2), \quad k \in \mathbb{N}.$$

Then,

$$\sum_{k=1}^{\infty} \varepsilon \left\{ \frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1} \sup_{s \in I_k} V_s \right\} (\cdot \times (x, \infty)) \xrightarrow{T \rightarrow \infty} \sum_{k=1}^{\infty} \zeta_k \varepsilon \{j_k\}.$$

(b) Assume that  $L_1 \in \mathcal{S} \cap \text{MDA}(\Lambda)$ . Let  $(j_k)_{k \in \mathbb{N}}$  be the jump times of a Poisson process with intensity  $e^{-x}$  for  $x \in \mathbb{R}$ . Then,

$$\sum_{k=1}^{\infty} \varepsilon \left\{ \frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1} (\sup_{s \in I_k} V_s - b_{\lambda T}) \right\} (\cdot \times (x, \infty)) \xrightarrow{T \rightarrow \infty} \sum_{k=1}^{\infty} \varepsilon \{j_k\}.$$

Again the qualitative difference of the two regimes is obvious. In the case of a regularly varying BDLP  $L$  the limiting process is a compound Poisson process, where at each Poisson point a cluster appears, whose size is random with distribution  $(\pi_k)_{k \in \mathbb{N}}$ . In contrast to this, in the  $\text{MDA}(\Lambda)$  case, the limit process is simply a homogeneous Poisson process; no clusters appear in the limit.

As the next result shows, the running maxima of the volatility process  $V$  have the same behavior as that of an i. i. d. sequence of copies of  $L_\lambda$ .

**Corollary 3.6 (Running maxima).**

Let  $V$  be a stationary version of the OU process given by (3.1), and define  $M(T) := \sup_{t \in [0, T]} V_t$  for  $T > 0$ .

(a) Assume that  $L_1 \in \mathcal{S} \cap \text{MDA}(\Phi_\alpha)$  with norming constants  $a_T > 0$  given by (3.14). Then

$$\lim_{T \rightarrow \infty} \mathbb{P}(a_{\lambda T}^{-1} M(T) \leq x) = \exp(-x^{-\alpha}), \quad x > 0.$$

(b) Assume that  $L_1 \in \mathcal{S} \cap \text{MDA}(\Lambda)$  with norming constants  $a_T > 0$ ,  $b_T \in \mathbb{R}$  given by (3.16). Then

$$\lim_{T \rightarrow \infty} \mathbb{P} \left( a_{\lambda T}^{-1} (M(T) - b_{\lambda T}) \leq x \right) = \exp(-e^{-x}), \quad x \in \mathbb{R}.$$

Finally, we investigate the possibility of volatility clusters in the OU process. As the concept of  $\epsilon$ -upcrossings is only defined for continuous-time processes, which does not fit into our framework, we shall introduce an appropriate method for describing clusters in continuous-time processes with jumps.

As our method will be motivated by the discrete-time skeleton of  $V$ , we recall that in a discrete-time process clusters are usually described by the extremal index  $\theta \in (0, 1]$ ; see Definition A.9. However, continuous-time processes are by nature dependent in small time intervals by the continuity and the structure of the process. Thus it is not adequate to adopt the extremal index concept for stochastic sequences to describe the dependence structure of the continuous-time process on a high level.

The following concept of an extremal index function has been introduced in Fasen [18].

**Definition 3.7.** Let  $(V_t)_{t \geq 0}$  be a stationary process. For  $h > 0$  define the sequence  $M_k(h) := \sup_{(k-1)h \leq t \leq kh} V_t$ ,  $k \in \mathbb{N}$ . Let  $\theta(h)$  be the extremal index of the sequence  $(M_k(h))_{k \in \mathbb{N}}$ . Then we call the function  $\theta : (0, \infty) \rightarrow [0, 1]$  extremal index function.

The idea is to divide the positive real line into blocks of length  $h$ . By taking local suprema of the process over these blocks the natural dependence of the continuous-time process is weakened, in certain cases it even disappears. However, for fixed  $h$  the extremal index function is a measure for the expected cluster sizes among these blocks. For an extended discussion on the extremal index in the context of discrete- and continuous-time processes see Fasen [18].

**Corollary 3.8 (Extremal index function).**

- (a) If  $L_1 \in \mathcal{S} \cap \text{MDA}(\Phi_\alpha)$ , then  $\theta(h) = h\alpha / (h\alpha + 1)$ ,  $h > 0$ .
- (b) If  $L_1 \in \mathcal{S} \cap \text{MDA}(\Lambda)$ , then  $\theta(h) = 1$ ,  $h > 0$ .

Regularly varying OU processes exhibit clusters among blocks, since  $\theta(h) < 1$ . So they have the potential to model both volatility features: heavy tails and high level clusters. This is in contrast to OU processes in  $\mathcal{S} \cap \text{MDA}(\Lambda)$ , where no clusters occur.

### 3.2 OU processes with exponential tails

In this section we investigate OU models having exponential tails, hence are lighter tailed than those considered in the previous section. More precisely, we will concentrate on two classes of models in  $\mathcal{L}(\gamma)$ ,  $\gamma > 0$ ; see Definition A.2. The first class concerns the *class of convolution equivalent distributions*  $\mathcal{S}(\gamma)$ ,  $\gamma > 0$  (Definition A.3). Here Theorem A.4 provides the necessary relationship between the tails of the infinitely divisible d. f. and of its Lévy measure, which leads to a comparison between the distribution tail of the stationary r. v.  $V_0$  and the increment  $L_1$  of the BDLP. An important family in  $\mathcal{S}(\gamma)$  are d. f. s with tail

$$\overline{F}(x) \sim x^{-\beta} l(x) e^{-\gamma x - cx^p}, \quad x \rightarrow \infty,$$

where  $\gamma, c \geq 0$ ,  $p < 1$ ,  $l(\cdot)$  is slowly varying, and if  $c = 0$ , then  $\beta > 1$  or  $\beta = 1$  and  $\int_1^\infty l(x)/x dx < \infty$  (Klüppelberg [25], Theorem 2.1, or Pakes [32], Lemma 2.3). The *generalized inverse Gaussian distribution* (GIGD) with probability density

$$p(x) = Kx^{\beta-1} \exp(-(\delta^2 x^{-1} + \gamma^2 x)/2), \quad x > 0,$$

where  $K$  is the normalizing constant,  $\beta < 0$  and  $\delta^2 > 0$ , is a prominent example in  $\mathcal{S}(\gamma^2/2)$ . Further examples for distributions in  $\mathcal{S}(\gamma)$  can be found e. g. in Cline [14].

The second class of processes with exponential tails, which we investigate in this section, are  $\Gamma$ -OU processes. These are defined as stationary OU processes, where  $V_0$  is  $\Gamma(\mu, \gamma)$  distributed with probability density as defined in (2.5) for  $\mu > 1$  and  $\gamma > 0$ . The gamma distribution is infinitely divisible with absolutely continuous Lévy measure given by its density

$$\nu_V(dx) = \mu x^{-1} e^{-\gamma x} dx, \quad x > 0.$$

Hence, by (3.4) the BDLP  $L$  has Lévy density

$$\nu(dx) = \mu \gamma e^{-\gamma x} dx, \quad x > 0.$$

Except for the factor  $\mu$  this is the probability density of an exponential d. f. Hence  $L$  is a positive compound Poisson process with Poisson rate  $\mu > 0$  and exponential jumps; for details see Barndorff-Nielsen and Shephard [2]. The exponential and gamma laws with scale parameter  $\gamma > 0$  belong to  $\mathcal{L}(\gamma)$  but not to  $\mathcal{S}(\gamma)$ .

In analogy to the  $\Gamma$ -OU process, also for OU- $\mathcal{S}(\gamma)$  processes with  $\gamma > 0$  we restrict our attention to positive compound Poisson processes as BDLPs; i. e. we

work in the framework of positive shot noise processes as defined in Example 3.1. Note that by Proposition A.6 (b) all d.f.s in  $\mathcal{L}(\gamma)$  for  $\gamma > 0$  belong to  $\text{MDA}(\Lambda)$ .

Some of the results in this section can be found in Albin [1], who studies the extremal behavior for a larger class of OU processes by purely analytic methods.

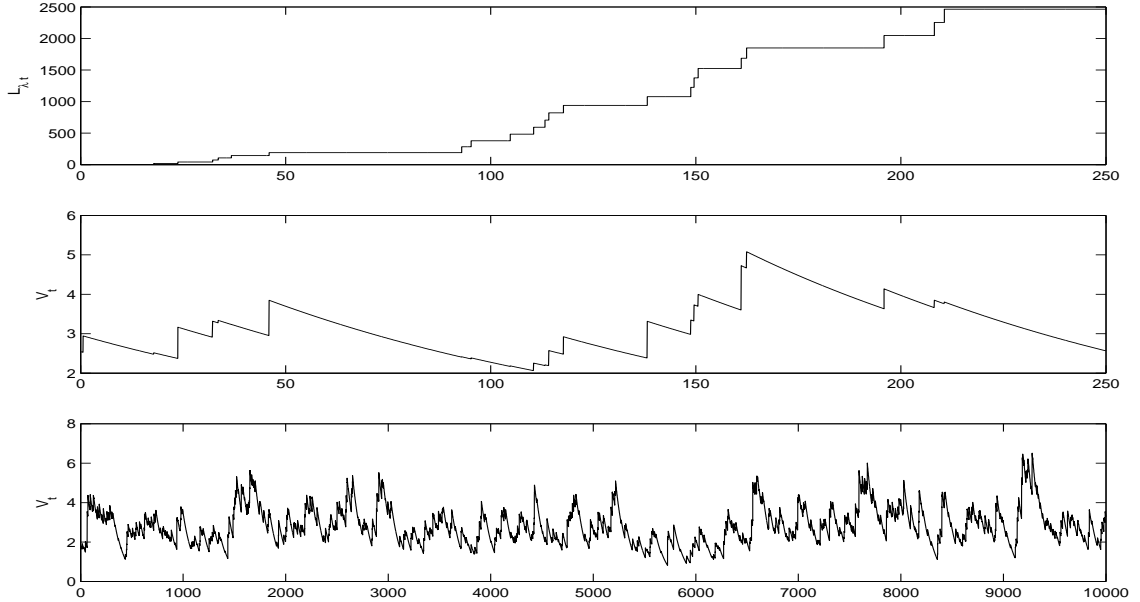


Figure 4: Sample path of a  $\Gamma$ -OU process with  $\gamma = 3$ ,  $\mu = 8.5$  and  $\lambda = 0.01$ . The first plot shows  $(L_{\lambda t})_{0 \leq t \leq 250}$  and the second plot the corresponding volatility  $(V_t)_{0 \leq t \leq 250}$ , indicating the micro-behavior of the model. The third plot gives  $(V_t)_{0 \leq t \leq 10000}$  indicating the macro-behavior of the model.

For BDLPs in  $\mathcal{S}(\gamma)$  for  $\gamma > 0$  the relation of the tail of the stationary d.f. and its Lévy measure are stated in the following proposition.

**Proposition 3.9 (Tail behavior of OU- $\mathcal{S}(\gamma)$  models for  $\gamma > 0$ ).**

Let  $V$  be a stationary version of the OU process given by (3.1).

(a) Suppose  $\nu(1, \cdot] / \nu(1, \infty) \in \mathcal{L}(\gamma)$ ,  $\gamma > 0$ . Then  $\nu_V(1, \cdot] / \nu_V(1, \infty) \in \mathcal{L}(\gamma)$  with

$$\nu_V(x, \infty) \sim \frac{1}{\gamma x} \nu(x, \infty), \quad x \rightarrow \infty.$$

(b) Suppose  $L_1 \in \mathcal{S}(\gamma)$ ,  $\gamma > 0$ . Then also  $V_0 \in \mathcal{S}(\gamma)$ , and

$$\mathbb{P}(V_0 > x) \sim \frac{\mathbb{E}e^{\gamma V_0}}{\mathbb{E}e^{\gamma L_1}} \frac{1}{\gamma x} \mathbb{P}(L_1 > x), \quad x \rightarrow \infty.$$

In particular,  $\mathbb{P}(V_0 > x) = o(\mathbb{P}(L_1 > x))$  as  $x \rightarrow \infty$ .

*Proof.* (a) By (A.1) the Lévy measure  $\nu$  has representation

$$\nu(x, \infty) = c(x) \exp \left[ - \int_1^x \frac{1}{a(y)} dy \right], \quad x \geq 1, \quad (3.18)$$

for functions  $a, c : [1, \infty) \rightarrow \mathbb{R}_+$  with  $\lim_{x \rightarrow \infty} c(x) = c > 0$  and  $\lim_{x \rightarrow \infty} a(x) = 1/\gamma$ ,  $\lim_{x \rightarrow \infty} a'(x) = 0$ . Since we are only interested in the tail behavior we may assume without loss of generality that  $\nu$  is absolutely continuous and  $c(\cdot) \equiv c$ . Recall from (3.4) that  $\nu_V(dx) = x^{-1}\nu(x, \infty) dx$  and let  $\nu(dx) = \nu'(x) dx$ . Part (a) follows by an application of l'Hospital's rule, since

$$\frac{\nu_V(x, \infty)}{\nu(x, \infty)/(\gamma x)} \sim \frac{\nu(x, \infty)/x}{[\nu'(x)x + \nu(x, \infty)]/(\gamma x^2)} = \gamma \left[ \frac{1}{a(x)} + \frac{1}{x} \right]^{-1} \rightarrow 1, \quad x \rightarrow \infty.$$

(b) We first show that  $V_0 \in \mathcal{S}(\gamma)$ . By Theorem A.4 (i) it suffices to show that  $\nu_V(1, \cdot]/\nu(1, \infty) \in \mathcal{S}(\gamma)$ . Again, we can assume without loss of generality that  $\nu$  is absolutely continuous and has the representation (3.18) with constant  $c(\cdot) \equiv c$ . For simplicity, we further assume that  $c = 1$  and  $\nu(1, \infty) = 1$ ; the general case follows by a simple dilation. As in part (a) we use that  $\nu_V(dx) = x^{-1}\nu(x, \infty) dx$ . An application of l'Hospital's rule shows that

$$\frac{\nu_V(x-y, \infty)}{\nu_V(x, \infty)} \sim \frac{\nu(x-y, \infty) x}{\nu(x, \infty) (x-y)} \rightarrow e^{\gamma y}, \quad x \rightarrow \infty,$$

implying  $\nu_V(1, \cdot] \in \mathcal{L}(\gamma)$ . Denote by  $\nu_V^{2*}$  the convolution of  $\nu_V$  restricted to  $(1, \infty)$  with itself. Then for  $1 < y' < x/2$  we use the usual decomposition of the convolution integral

$$\begin{aligned} \frac{\nu_V^{2*}(dx)}{dx} &= \int_1^x \frac{\nu(u, \infty)}{u} \frac{\nu(x-u, \infty)}{x-u} du \\ &= 2 \int_1^{y'} \frac{\nu(u, \infty)}{u} \frac{\nu(x-u, \infty)}{x-u} du + \int_{y'}^{x-y'} \frac{\nu(u, \infty)}{u} \frac{\nu(x-u, \infty)}{x-u} du. \end{aligned} \quad (3.19)$$

In order to show that  $\nu_V(1, \cdot] \in \mathcal{S}(\gamma)$ , we calculate the limit ratio of the densities of  $\nu_V^{2*}$  and  $\nu_V$ . Observe that on every compact set  $\nu(x-u, \infty)/\nu(x, \infty)$  converges uniformly in  $u$  to  $\exp(\gamma u)$  as  $x \rightarrow \infty$ . For the first summand of (3.19) we thus obtain

$$\begin{aligned} \lim_{x \rightarrow \infty} 2 \int_1^{y'} \frac{x}{x-u} \frac{\nu(x-u, \infty)}{\nu(x, \infty)} \frac{\nu(u, \infty)}{u} du &= 2 \int_1^{y'} e^{\gamma u} \frac{\nu(u, \infty)}{u} du \\ &= 2 \int_1^{y'} e^{\gamma u} \nu_V(du) < \infty. \end{aligned} \quad (3.20)$$

For the second summand in (3.19) we estimate

$$\int_{y'}^{x-y'} \frac{x\nu(u, \infty)\nu(x-u, \infty)}{u(x-u)\nu(x, \infty)} du \leq \frac{x}{y'(x-y')} \int_{y'}^{x-y'} \frac{\nu(x-u, \infty)}{\nu(x, \infty)} \nu(u, \infty) du. \quad (3.21)$$

Furthermore, since

$$\nu(x, \infty) / \nu'(x) = a(x) \longrightarrow 1/\gamma, \quad x \rightarrow \infty,$$

there exist constants  $K, x_0 > 0$  such that  $\nu(x, \infty) \leq K\nu'(x)$  for  $x \geq x_0$ . We obtain for  $y' > x_0$

$$\frac{x}{y'(x-y')} \int_{y'}^{x-y'} \frac{\nu(x-u, \infty)\nu(u, \infty)}{\nu(x, \infty)} du \leq \frac{Kx}{y'(x-y')} \int_{y'}^{x-y'} \frac{\nu(x-u, \infty)}{\nu(x, \infty)} \nu(du). \quad (3.22)$$

Since  $\nu(1, \cdot] \in \mathcal{S}(\gamma)$ , the same decomposition as in (3.19) yields (for details see e. g. Pakes [32], Lemma 5.5)

$$\lim_{y' \rightarrow \infty} \lim_{x \rightarrow \infty} \int_{y'}^{x-y'} \frac{\nu(x-u, \infty)}{\nu(x, \infty)} \nu(du) = 0. \quad (3.23)$$

Furthermore,  $\lim_{y' \rightarrow \infty} \lim_{x \rightarrow \infty} x/[y'(x-y')] = \lim_{y' \rightarrow \infty} 1/y' = 0$ . By (3.19)-(3.23) we now obtain  $\nu_V^{2*}(dx) \sim (2 \int_1^\infty e^{\gamma u} \nu(du)) \nu_V(dx)$  for  $x \rightarrow \infty$ , showing that  $\nu_V(1, \cdot]$  and hence  $V_0$  are in  $\mathcal{S}(\gamma)$ . The assertion on the tail behavior now follows from (a) and Theorem A.4 (i).  $\square$

The following result is an analogon to Theorem 3.4 and describes the extremal behavior of  $V$  completely.

**Theorem 3.10 (Point process of exceedances of exponential models).**

Let  $V$  be a stationary version of the OU process given by (3.1) with  $L$  a positive compound Poisson process as in (3.5). Denote by  $(\Gamma_k)_{k \in \mathbb{N}}$  the jump times of the positive compound Poisson process  $L$  given by (3.5) and define  $I_k := \frac{1}{\lambda} [\Gamma_k, \Gamma_{k+1})$ ,  $k \in \mathbb{N}$ . Let  $\sum_{k=1}^\infty \varepsilon\{s_k, P_k\}$  be PRM( $\vartheta$ ) with  $\vartheta(dt \times dx) = dt \times e^{-x} dx$ .

(a) Assume  $L_1 \in \mathcal{S}(\gamma)$ ,  $\gamma > 0$ , with norming constants  $a_T > 0$ ,  $b_T \in \mathbb{R}$  such that

$$\lim_{T \rightarrow \infty} T\mathbb{P}(L_1 > a_T x + b_T) = \frac{\mathbb{E}e^{\gamma L_1}}{\mathbb{E}e^{\gamma V_0}} \exp(-x), \quad x \in \mathbb{R}. \quad (3.24)$$

Then

$$\sum_{k=0}^\infty \varepsilon \left\{ \frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1}(\sup_{s \in I_k} V_s - b_{\lambda T}) \right\} \xrightarrow{T \rightarrow \infty} \sum_{k=1}^\infty \varepsilon \{s_k, P_k\}. \quad (3.25)$$



(b) Assume  $V$  is the  $\Gamma(\mu, \gamma)$ -OU process. Let  $a_T > 0, b_T \in \mathbb{R}$  be the norming constants of a  $\Gamma(\mu + 1, \gamma)$  distributed r. v.  $W$ , such that

$$\lim_{T \rightarrow \infty} T\mathbb{P}(W > a_T x + b_T) = \mu^{-1} \exp(-x), \quad x \in \mathbb{R}. \quad (3.26)$$

Then

$$\sum_{k=0}^{\infty} \varepsilon \left\{ \frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1}(\sup_{s \in I_k} V_s - b_{\lambda T}) \right\} \xrightarrow{T \rightarrow \infty} \sum_{k=1}^{\infty} \varepsilon \{s_k, P_k\}. \quad (3.27)$$

The proof is divided into several steps. We shall utilize classical results for the extreme value theory of stationary discrete-time processes. As a discrete-time skeleton  $(V_{\Gamma_k/\lambda})_{k \in \mathbb{Z}}$  seems to be a good candidate. However,  $V_{\Gamma_k/\lambda} = \sum_{j \neq 0}^k e^{-(\Gamma_k - \Gamma_j)} \xi_j$ ,  $k \in \mathbb{N}$ , is not stationary. As we will show in Lemma 3.11 the process

$$\tilde{V}_k = \sum_{j=-\infty}^k e^{-(\Gamma_k - \Gamma_j)} \xi_j = V_{\Gamma_k/\lambda} + e^{-\Gamma_k} \xi_0, \quad k \in \mathbb{N}, \quad (3.28)$$

is stationary, where  $\Gamma_0 := 0$ . For increasing  $k$  the process  $e^{-\Gamma_k} \xi_0$  tends to 0. Thus it has no influence on the extremal behavior. We shall show that the point process behavior is the same for  $(V_{\Gamma_k/\lambda})_{k \in \mathbb{N}}$  and for  $(\tilde{V}_k)_{k \in \mathbb{N}}$ . For the proof we also need that the  $D$  and  $D'$  conditions (Definition B.1) hold for  $(\tilde{V}_k)_{k \in \mathbb{N}}$ . The highly technical Lemma B.2, where this is confirmed, is postponed to the Appendix.

**Lemma 3.11.** *Let  $V$  be a stationary version of the OU process given by (3.1) with  $L$  a positive compound Poisson process as in (3.5). Then  $(\tilde{V}_k)_{k \in \mathbb{Z}}$  as defined in (3.28) is stationary.*

*Proof.* Let  $h \in \mathbb{R}$  be fixed. Note that  $(\Gamma_{h+j} - \Gamma_h)_{j \in \mathbb{Z}} \stackrel{d}{=} (\Gamma_j)_{j \in \mathbb{Z}}$ . Then

$$\begin{aligned} \tilde{V}_{k+h} &= \sum_{j=-\infty}^{k+h} e^{-(\Gamma_{k+h} - \Gamma_j)} \xi_j = \sum_{j=-\infty}^{k+h} e^{-(\Gamma_{k+h} - \Gamma_h - (\Gamma_j - \Gamma_h))} \xi_j \stackrel{d}{=} \sum_{j=-\infty}^{k+h} e^{-(\Gamma_k - \Gamma_{j-h})} \xi_j \\ &= \sum_{j=-\infty}^k e^{-(\Gamma_k - \Gamma_j)} \xi_{j+h} \stackrel{d}{=} \sum_{j=-\infty}^k e^{-(\Gamma_k - \Gamma_j)} \xi_j = \tilde{V}_k. \end{aligned}$$

Similarly, for  $l \in \mathbb{N}$  we obtain  $(\tilde{V}_{k_1+h}, \dots, \tilde{V}_{k_l+h}) \stackrel{d}{=} (\tilde{V}_{k_1}, \dots, \tilde{V}_{k_l})$  for  $k_1, \dots, k_l \in \mathbb{N}$ .  $\square$

*Proof of Theorem 3.10.* Since  $V$  is decreasing between jumps, it follows that  $\sup_{s \in I_k} V_s = V_{\Gamma_k/\lambda}$ . Recall that  $\tilde{V}_k = V_{\Gamma_k/\lambda} + e^{-\Gamma_k} \xi_0 \stackrel{d}{=} V_0 + \xi_1$  and that  $(\tilde{V}_k)_{k \in \mathbb{N}}$

is stationary. We show first that the norming constants  $a_n > 0, b_n \in \mathbb{R}$  given by (3.24) and (3.26) satisfy

$$\lim_{n \rightarrow \infty} n\mathbb{P}(\tilde{V}_k > a_n x + b_n) = \mu^{-1} \exp(-x), \quad x \in \mathbb{R}. \quad (3.29)$$

To show this in case (a), observe that  $\mathbb{P}(V_0 > x) = o(\mathbb{P}(\xi_1 > x))$  for  $x \rightarrow \infty$  by Proposition 3.9 (b), so that Theorem A.4 (i,ii) yields

$$\mathbb{P}(\tilde{V}_k > x) \sim \mathbb{E}e^{\gamma V_0} \mathbb{P}(\xi_1 > x) \sim \mathbb{E}e^{\gamma V_0} [\mathbb{E}e^{\gamma L_1}]^{-1} \mu^{-1} \mathbb{P}(L_1 > x), \quad x \rightarrow \infty. \quad (3.30)$$

From this (3.29) follows immediately, and further we see that  $\tilde{V}_k \in \mathcal{S}(\gamma)$ .

In case (b),  $\tilde{V}_k$  is  $\Gamma(\mu + 1, \gamma)$  distributed as an independent sum of a  $\Gamma(\mu, \gamma)$  and an  $\text{Exp}(\gamma)$  r. v., and (3.29) is immediate. The norming constants of a  $\Gamma$  distribution can be found in Table 3.4.4 of Embrechts et al. [17].

Note that in both cases (a) and (b), we have  $\tilde{V}_k \in \mathcal{L}(\gamma)$ . Thus, by Lemma B.2 and Leadbetter et al. [29], Theorem 5.5.1,

$$\sum_{k=0}^{\infty} \varepsilon \left\{ k/(\mu n), a_n^{-1}(\tilde{V}_k - b_n) \right\} \xrightarrow{n \rightarrow \infty} \sum_{k=1}^{\infty} \varepsilon \{s_k, P_k\}. \quad (3.31)$$

Define point processes

$$\tilde{\kappa}_n := \sum_{k=0}^{\infty} \varepsilon \left\{ k/(\mu n), a_n^{-1}(\tilde{V}_k - b_n) \right\} \quad \text{and} \quad \kappa_n := \sum_{k=0}^{\infty} \varepsilon \left\{ k/(\mu n), a_n^{-1}(V_{\Gamma_{k/\lambda}} - b_n) \right\}.$$

For  $\epsilon > 0$  and  $I = [s, t) \times (x, \infty) \subseteq \mathbb{R}_+ \times \mathbb{R}$  define  $I_\epsilon := [s, t) \times (x, x + \epsilon]$ . Taking into account that  $V_{\Gamma_{k/\lambda}} \leq \tilde{V}_k$  we have for  $\delta \in (0, 1)$

$$\begin{aligned} & \mathbb{P}(\kappa_n(I) \neq \tilde{\kappa}_n(I)) \\ & \leq \mathbb{P}(\tilde{\kappa}_n(I_\epsilon) > 0) + \sum_{k \in [sn\mu, tn\mu)} \mathbb{P}(\tilde{V}_k > u_n + \epsilon a_n, V_{\Gamma_{k/\lambda}} \leq u_n) \\ & \leq \mathbb{P}(\tilde{\kappa}_n(I_\epsilon) > 0) + \sum_{k \in [0, n^\delta t\mu)} \mathbb{P}(\tilde{V}_k > u_n + \epsilon a_n) + \sum_{k \in [n^\delta t\mu, nt\mu)} \mathbb{P}(e^{-\Gamma_k} \xi_0 > \epsilon a_n). \end{aligned}$$

We shall show below that

$$\lim_{n \rightarrow \infty} \mathbb{P}(\kappa_n(I) \neq \tilde{\kappa}_n(I)) = 0. \quad (3.32)$$

Then by Rootzén [33], Lemma 3.3, the limit behavior of  $\tilde{\kappa}_n$  and  $\kappa_n$  is the same. Relation (3.27) then follows by transforming the time scale as in Hsing and Teugels [23] (for details see Fasen [18], Lemma 2.2.4).

To show (3.32), observe that by (3.31) we have

$$\lim_{n \rightarrow \infty} \mathbb{P}(\tilde{\kappa}_n(I_\epsilon) > 0) = 1 - \exp[(t-s)(\exp(-x) - \exp(-(x+\epsilon)))] \xrightarrow{\epsilon \downarrow 0} 0.$$

Furthermore, since  $\delta < 1$ ,

$$\lim_{n \rightarrow \infty} \sum_{k \in [0, n^\delta t \mu)} \mathbb{P}(\tilde{V}_k > u_n + \epsilon a_n) \leq \lim_{n \rightarrow \infty} n^\delta t \mu \mathbb{P}(\tilde{V}_k > a_n(x + \epsilon) + b_n) = 0.$$

Applying (B.3) we obtain

$$\begin{aligned} & \sum_{k \in [n^\delta t \mu, nt \mu)} \mathbb{P}(e^{-\Gamma_k} \xi_0 > \epsilon a_n) \\ & \leq \sum_{k \in [n^\delta t \mu, nt \mu)} \left( \mathbb{P}(e^{-\Gamma_k} \xi_0 > \epsilon a_n, \Gamma_k \geq \frac{k}{2\mu}) + \mathbb{P}(e^{-\Gamma_k} \xi_0 > \epsilon a_n, \Gamma_k < \frac{k}{2\mu}) \right) \\ & \leq \sum_{k \in [n^\delta t \mu, nt \mu)} \mathbb{P}(e^{-k/(2\mu)} \xi_0 > \epsilon a_n) + \sum_{k \in [n^\delta t \mu, nt \mu)} \frac{\tilde{K}}{k^3}. \end{aligned}$$

The last summand tends to 0 as  $n \rightarrow \infty$ , since  $\sum_{k=1}^{\infty} 1/k^3 < \infty$ . Moreover, there exists an  $n_0 \in \mathbb{N}$  such that  $a_n \geq 1/(2\gamma)$  and  $ke^{-k/(2\mu)} \leq 1/2$  for  $n, k \geq n_0$ . Then the first exponential moment of  $\gamma ke^{-k/(2\mu)} \xi_0$  exists, and for  $n^\delta t \mu \geq n_0$  we obtain

$$\begin{aligned} & \sum_{k \in [n^\delta t \mu, nt \mu)} \mathbb{P}(e^{-k/(2\mu)} \xi_0 > \epsilon a_n) \leq \sum_{k \in [n^\delta t \mu, nt \mu)} \mathbb{E}[\exp(\gamma ke^{-k/(2\mu)} \xi_0) e^{-k\gamma \epsilon a_n}] \\ & \leq \mathbb{E}[\exp(\gamma \xi_0 / 2)] \sum_{k=[n^\delta t \mu]}^{\infty} e^{-k\epsilon/2} \longrightarrow 0, \quad n \rightarrow \infty, \end{aligned}$$

since  $\sum_{k=1}^{\infty} e^{-k\epsilon/2} < \infty$ . This shows (3.32).  $\square$

Results (3.25) and (3.27) show that local extremes of such exponential models have no cluster behavior on high levels. The following two corollaries are immediate from Theorem 3.10.

**Corollary 3.12 (Point process of local maxima).**

*Let the assumptions of Theorem 3.10 hold. Denote by  $(j_k)_{k \in \mathbb{N}}$  the jump times of a Poisson process with intensity  $e^{-x}$  for  $x \in \mathbb{R}$ . Then,*

$$\sum_{k=1}^{\infty} \varepsilon \left\{ \frac{\Gamma_k}{\lambda T}, a_{\lambda T}^{-1}(\sup_{s \in I_k} V_s - b_{\lambda T}) \right\} (\cdot \times (x, \infty)) \xrightarrow{T \rightarrow \infty} \sum_{k=1}^{\infty} \varepsilon \{j_k\}.$$

**Corollary 3.13 (Running maxima).**

Let  $V$  be a stationary version of the OU process (3.1), where  $L$  is a positive, compound Poisson process as in (3.5). Define  $M(T) := \sup_{0 \leq t \leq T} V_t$  for  $T > 0$ .

(a) Assume  $L_1 \in \mathcal{S}(\gamma)$ ,  $\gamma > 0$ , with norming constants given by (3.24). Then

$$\lim_{T \rightarrow \infty} \mathbb{P}(a_{\lambda T}^{-1}(M(T) - b_{\lambda T}) \leq x) = \exp(-e^{-x}), \quad x \in \mathbb{R}.$$

(b) Assume  $V$  is the  $\Gamma(\mu, \gamma)$ -OU process with norming constants given by (3.26). Then

$$\lim_{T \rightarrow \infty} \mathbb{P}(a_{\lambda T}^{-1}(M(T) - b_{\lambda T}) \leq x) = \exp(-e^{-x}), \quad x \in \mathbb{R}. \quad (3.33)$$

For a subexponential OU process and  $h > 0$  fixed the r. v.  $M(h) = \sup_{0 \leq t \leq h} V_t$  is tail-equivalent to the increment of the Lévy process; cf. (3.9) and (3.12). In the class  $\mathcal{S}(\gamma)$ ,  $\gamma > 0$ , this is much more involved; see Braverman and Samorodnitsky [9]. Although the large jumps of the Lévy process determine the tail behavior, small jumps also have a non-negligible influence. For any  $h > 0$ , the tail of  $M(h)$  is of the same order of magnitude as the tail of the increment of the BDLP, but in general it is only possible to give upper and lower bounds on the asymptotic ratio of the two tails. Using Corollary 3.13 one can calculate this constant for the OU process explicitly.

**Corollary 3.14 (Extremal index function).**

Let  $V$  be a stationary version of the OU process given by (3.1), where  $L$  is a positive compound Poisson process as in (3.5). Define  $M(h) := \sup_{0 \leq t \leq h} V_t$  for  $h > 0$ .

(a) Let  $L_1 \in \mathcal{S}(\gamma)$ ,  $\gamma > 0$ . Then  $M(h) \in \mathcal{L}(\gamma)$  if and only if

$$\mathbb{P}(M(h) > x) \sim \lambda h \mathbb{E}e^{\gamma V_0} [\mathbb{E}e^{\gamma L_1}]^{-1} \mathbb{P}(L_1 > x), \quad x \rightarrow \infty. \quad (3.34)$$

In that case  $M(h) \in \mathcal{S}(\gamma)$  and  $\theta(\cdot) \equiv 1$ .

(b) Let  $V$  be the  $\Gamma(\mu, \gamma)$ -OU process with norming constants given by (3.26) and let  $W$  be a  $\Gamma(\mu + 1, \gamma)$  r. v. Then  $\theta(\cdot) \equiv 1$  and

$$\mathbb{P}(M(h) > x) \sim \lambda h \mu \mathbb{P}(W > x), \quad x \rightarrow \infty. \quad (3.35)$$

*Proof.* (a) First we assume  $M(h) \in \mathcal{L}(\gamma)$ . Let  $\tilde{a}_n > 0$ ,  $\tilde{b}_n \in \mathbb{R}$  and  $\tilde{u}_n = \tilde{a}_n x + \tilde{b}_n$  be constants such that

$$\lim_{n \rightarrow \infty} n \mathbb{P}(M(h) > \tilde{u}_n) = \exp(-x).$$

Denote by  $\widetilde{M}_k$  an i.i.d. sequence of copies of  $M(h)$ . Then we obtain from Lemma B.2 (b) and Leadbetter et al. [29], Theorem 3.5.1, for  $x \in \mathbb{R}$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P}(\widetilde{a}_n^{-1}(M(nh) - \widetilde{b}_n) \leq x) = \lim_{n \rightarrow \infty} \mathbb{P}(\widetilde{a}_n^{-1}(\max_{k=1, \dots, n} \widetilde{M}_k - \widetilde{b}_n) \leq x) = \exp(-e^{-x}),$$

showing in particular that  $\theta(h) = 1$ . On the other hand, by Corollary 3.13,

$$\lim_{n \rightarrow \infty} \mathbb{P}(a_{\lambda nh}^{-1}(M(nh) - b_{\lambda nh}) \leq x) = \exp(-e^{-x}), \quad x \in \mathbb{R}.$$

Then by the convergence to types theorem (see e.g. Theorem A 1.5 of Embrechts et al. [17]),  $\widetilde{a}_n/a_{\lambda nh} \xrightarrow{n \rightarrow \infty} 1$  and  $\widetilde{b}_n - b_{\lambda nh} \xrightarrow{n \rightarrow \infty} 0$ . Applying the convergence to types theorem a second time yields

$$\lim_{n \rightarrow \infty} \mathbb{P}(a_{\lambda nh}^{-1}(\max_{k=1, \dots, n} \widetilde{M}_k - b_{\lambda nh}) \leq x) = \exp(-e^{-x}), \quad x \in \mathbb{R}.$$

This implies by Leadbetter et al. [29], Theorem 1.5.1 that

$$\lim_{n \rightarrow \infty} n\mathbb{P}(M(h) > u_{\lambda nh}) = \exp(-x), \quad x \in \mathbb{R},$$

with  $u_{\lambda nh} = a_{\lambda nh}x + b_{\lambda nh}$ . By (3.29) also  $\lim_{n \rightarrow \infty} n\mathbb{P}(\widetilde{V}_k > u_{\lambda nh}) = \exp(-x)/(\lambda\mu h)$ . Hence  $\mathbb{P}(M(h) > x) \sim h\lambda\mu\mathbb{P}(\widetilde{V}_k > x)$  for  $x \rightarrow \infty$ , and (3.34) follows from (3.30).

Conversely, if (3.34) holds, then it is clear that  $L_1 \in \mathcal{S}(\gamma) \subseteq \mathcal{L}(\gamma)$  implies  $M(h) \in \mathcal{L}(\gamma)$  by tail-equivalence. By Lemma B.2 (b) follows  $\theta(h) = 1$ .

(b) We refer to Albin [1], Theorem 3, for (3.35). That  $\theta(h) = 1$  follows then from (3.26), (3.33) and (3.35).  $\square$

In both cases the extremal index function is equal to one, so that for any  $h > 0$  the sequence  $M_k = \sup_{(k-1)h \leq t \leq kh} V_t$  behaves like i.i.d. data. Hence such models cannot explain volatility clusters on high levels.

## 4 Extremal behavior of the COGARCh model

The volatility of the COGARCh(1,1) process as introduced in (1.6) is the (càglàg) solution to the SDE (1.6), which is given by

$$V_t = V_0 + \beta t - \eta \int_0^t V_s ds + \lambda e^\eta \sum_{0 < s < t} V_s (\Delta L_s)^2, \quad t \geq 0, \quad (4.1)$$

see Klüppelberg, Lindner and Maller [27, 28] for details. This process is a solution of the SDE

$$dV_{t+} = (\beta - \eta V_t)dt + \lambda e^\eta V_t d[L, L]_t^{(d)}, \quad (4.2)$$

where

$$[L, L]_t^{(d)} = \sum_{0 < s \leq t} (\Delta L_s)^2, \quad t \geq 0,$$

is the discrete part of the quadratic variation process of  $L$ . Comparing this with (1.5) we see that the COGARCH(1,1) can be interpreted as a generalized CIR model driven by the discrete part of the quadratic variation process of  $L$ . An essential feature of the COGARCH(1,1) model is that the same Lévy process drives the price process  $S$  and the volatility process  $V$ . An extension of the COGARCH(1,1) process to COGARCH(p,q) process with  $1 \leq p \leq q$  was recently obtained by Brockwell, Chadraa and Lindner [11]. There, (4.2) is replaced by a CARMA (continuous time ARMA) type stochastic differential equation, driven by  $[L, L]_t^{(d)}$ . We shall not go into further details, and by COGARCH we shall always mean the COGARCH(1,1) process.

Denote by  $\nu$  the Lévy measure of  $L$ . A stationary version of (4.1) exists if and only if

$$\int_{\mathbb{R}} \log(1 + \lambda e^\eta y^2) \nu(dy) < \eta. \quad (4.3)$$

With the auxiliary càdlàg process  $(X_t)_{t \geq 0}$  defined in (1.7), given for  $\eta > 0$ ,  $\lambda \geq 0$  by

$$X_t = \eta t - \sum_{0 < s \leq t} \log(1 + \lambda e^\eta (\Delta L_s)^2), \quad t \geq 0,$$

the stationary volatility process has representation

$$V_t = \left( \beta \int_0^t e^{X_s} ds + V_0 \right) e^{-X_t}, \quad t \geq 0, \quad (4.4)$$

with  $\beta > 0$  and  $V_0 \stackrel{d}{=} \beta \int_0^\infty e^{-X_t} dt$ , independent of  $L$ . The auxiliary process  $(X_t)_{t \geq 0}$  itself is a spectrally negative Lévy process of bounded variation with drift  $\gamma_X = \eta$ , no Gaussian component, and Lévy measure  $\nu_X$  given by

$$\nu_X [0, \infty) = 0, \quad \nu_X (-\infty, -x] = \nu \left( \{y \in \mathbb{R} : |y| \geq \sqrt{(e^x - 1)/(\lambda e^\eta)}\} \right) \quad x > 0.$$

We work with the Laplace transform  $\mathbb{E}e^{-sX_t} = e^{t\Psi(s)}$ , where the Laplace exponent is

$$\Psi(s) = -\eta s + \int_{\mathbb{R}} ((1 + \lambda e^\eta y^2)^s - 1) \nu(dy), \quad s \geq 0. \quad (4.5)$$

For fixed  $s \geq 0$ ,  $\mathbb{E}e^{-sX_t}$  exists (i. e. is finite) for one and hence all  $t > 0$ , if and only if the integral appearing in (4.5) is finite. This is equivalent to  $\mathbb{E}|L_1|^{2s} < \infty$ . Further, if there exists some  $s > 0$  such that  $\Psi(s) \leq 0$ , then (4.3) holds, and hence a stationary version of the volatility process exists.

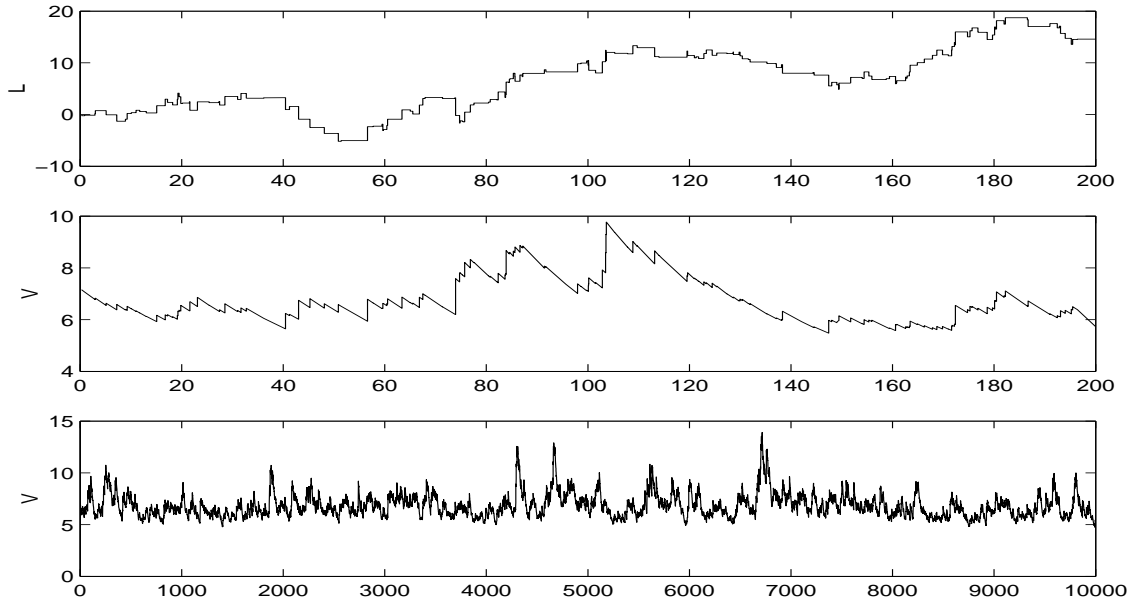


Figure 5: The first plot shows the sample path of a compound Poisson driving process  $(L_t)_{0 \leq t \leq 250}$  with rate 1 and normal jumps with mean 0 and variance 1 and the second plot the corresponding sample path of the COGARCH volatility process  $(V_t)_{0 \leq t \leq 250}$  driven by this Lévy process. The COGARCH parameters are  $\beta = 1$ ,  $\lambda = 0.04$  and  $\eta = 0.0619$ . Both plots indicate the micro-behavior of the model. The third plot gives  $(V_t)_{0 \leq t \leq 10000}$  indicating the macro-behavior of the model.

The qualitative extreme behavior of this volatility process can be seen in Figure 5, where the driving Lévy process is a compound Poisson process. As in the case of a Lévy driven OU process the volatility jumps upwards, whenever the driving Lévy process  $L$  jumps and decreases exponentially fast between two jumps.

It is instructive to observe that both, the OU process (3.1) and the right-continuous  $V_{t+}$  of the COGARCH volatility (4.4) are special cases of the generalized OU process

$$O_t = e^{-\xi t} \left( \int_0^t e^{\xi s} d\zeta_s + O_0 \right), \quad t \geq 0,$$

where  $(\xi_t, \zeta_t)_{t \geq 0}$  is a bivariate Lévy process. The stationarity conditions for  $(O_t)_{t \geq 0}$ , along with other properties, have been investigated by Lindner and Maller [30].

Returning to the COGARCH volatility, the next Theorem (cf. [28], Theorem 6) shows, that under weak conditions on the moments of  $L$ , the volatility process has Pareto like tails. Since we shall apply a similar argument in the proof of Theorem 4.3, we sketch the idea of the proof.

**Theorem 4.1 (Pareto tail behavior of COGARCH models).**

*Suppose there exists some  $\alpha > 0$  such that*

$$\mathbb{E}|L_1|^{2\alpha} \log^+ |L_1| < \infty \quad \text{and} \quad \Psi(\alpha) = 0. \quad (4.6)$$

*Let  $V$  be a stationary version of the volatility process given by (4.1). Then for some constant  $C > 0$  we have*

$$\lim_{x \rightarrow \infty} x^\alpha \mathbb{P}(V_0 > x) = C. \quad (4.7)$$

*Proof.* From (4.4) it is seen that the stationary volatility process  $V$  satisfies

$$V_t = e^{-X_t} V_0 + \beta \int_0^t e^{X_s - X_t} ds, \quad t > 0,$$

where  $V_0$  is independent of  $(e^{-X_t}, \beta \int_0^t e^{X_s - X_t} ds)$  for any  $t > 0$ . Thus the stationary solution  $V_0$  satisfies for every  $t > 0$  the distributional fix point equation

$$V_0 \stackrel{d}{=} A_t V_0 + B_t,$$

where  $V_0$  is independent of  $(A_t, B_t)$  and

$$A_t \stackrel{d}{=} e^{-X_t}, \quad B_t \stackrel{d}{=} \beta \int_0^t e^{-X_s} ds.$$

The result now follows from Theorem C.1, by choosing  $t$  such that  $(A_t, B_t)$  satisfies the assumptions. This is possible because of the structure of the process and condition (4.6); for details see Klüppelberg et al. [28], Theorem 6.  $\square$

The following remark gives a simple sufficient condition for (4.6) to hold.

**Remark 4.2.** Let  $D := \{d \in [0, \infty) : \mathbb{E}|L_1|^{2d} < \infty\}$  and  $d_0 := \sup D \in [0, \infty]$ . Suppose  $d_0 \notin D$ , or that there exists an  $s_0 > 0$  such that  $0 < \Psi(s_0) < \infty$ . Further suppose that  $(V_t)_{t \geq 0}$  is strictly stationary. Then (4.6) and hence (4.7) hold (cf. Klüppelberg et al. [28], Proposition 5.3).  $\square$

We aim at a precise asymptotic description of the COGARCH model above a high threshold like in Section 3. It is, however, clear from the definition of  $V$  that the influence of the spectrally negative Lévy process  $X$  is hard to analyze. In particular,



the influence of the small jumps of  $L$  to  $V$  needs special treatment. In this review paper we shall restrict ourselves again to the case of a compound Poisson driving process  $L$  as given in (3.5) by  $L_t = \sum_{j=1}^{N_t} \xi_j$  for  $t \geq 0$ , where  $\xi$  has support on  $\mathbb{R}$ .

In this case the auxiliary process  $X$  simplifies to

$$X_t = \eta t - \sum_{k=1}^{N_t} \log(1 + \lambda e^{\eta \xi_k^2}), \quad t \geq 0, \quad (4.8)$$

and the Laplace exponent becomes

$$\Psi(s) = -(\eta s + \mu) + \mu \mathbb{E}(1 + \lambda e^{\eta \xi_1^2})^s. \quad (4.9)$$

In the stationary volatility model we know that  $V_t \geq \beta/\eta$  a.s. and  $V$  jumps if and only if  $L$  jumps (cf. [28], Proposition 3.4 (a)). The jump sizes are positive and depend on the level of the process at that time. As shown in Proposition 3.4 (b,c) of [28],

$$V_{\Gamma_{k+}} - V_{\Gamma_k} = \lambda e^{\eta V_{\Gamma_k} \xi_k^2}, \quad k \in \mathbb{N}, \quad (4.10)$$

and the process decreases exponentially between jumps:

$$V_t = \frac{\beta}{\eta} + \left( V_{\Gamma_{k+}} - \frac{\beta}{\eta} \right) e^{-(t-\Gamma_k)\eta}, \quad t \in (\Gamma_k, \Gamma_{k+1}]. \quad (4.11)$$

In analogy to the OU process driven by a compound Poisson process of Example 3.1, the compound Poisson driven COGARCh process  $V$  achieves local suprema only at the right limits of its jump times (and at  $t = 0$ ). Hence it is no surprise that the discrete-time sequence  $(V_{\Gamma_{k+}})_{k \in \mathbb{N}}$  in combination with the deterministic behavior of  $V$  between jumps suffices to describe the extremal behavior of the continuous-time COGARCh process. Consequently, we investigate the discrete-time skeleton

$$\tilde{V}_k := V_{\Gamma_{k+}}, \quad k \in \mathbb{N}. \quad (4.12)$$

Using (4.10) and (4.11) we obtain

$$\tilde{V}_{k+1} = \tilde{V}_k (1 + \lambda e^{\eta \xi_{k+1}^2}) e^{-(\Gamma_{k+1} - \Gamma_k)\eta} + \frac{\beta}{\eta} (1 + \lambda e^{\eta \xi_{k+1}^2}) (1 - e^{-(\Gamma_{k+1} - \Gamma_k)\eta}),$$

and we see that  $(\tilde{V}_k)_{k \in \mathbb{N}}$  satisfies the stochastic recurrence equation

$$\tilde{V}_k = \tilde{A}_k \tilde{V}_{k-1} + \tilde{B}_k, \quad k \geq 2, \quad (4.13)$$

with  $\tilde{V}_1$  independent of  $(\tilde{A}_k, \tilde{B}_k)$  for any  $k \geq 2$ , where

$$\tilde{A}_k = (1 + \lambda e^\eta \xi_k^2) e^{-\eta(\Gamma_k - \Gamma_{k-1})}, \quad k \in \mathbb{N}, \quad (4.14)$$

$$\tilde{B}_k = \frac{\beta}{\eta} (1 + \lambda e^\eta \xi_k^2) (1 - e^{-\eta(\Gamma_k - \Gamma_{k-1})}), \quad k \in \mathbb{N}, \quad (4.15)$$

and  $((\tilde{A}_k, \tilde{B}_k))_{k \in \mathbb{N}}$  is an i.i.d. sequence. It is an interesting observation that by (4.14)

$$\log \prod_{j=1}^k \tilde{A}_j = \sum_{j=1}^k \log \tilde{A}_j = -\eta \Gamma_k + \sum_{j=1}^k \log(1 + \lambda e^\eta \xi_j^2) = -X_{\Gamma_k}.$$

On the other hand, by (4.15) and  $X_s - X_{\Gamma_k} = \log(1 + \lambda e^\eta \xi_k^2) + \eta(s - \Gamma_k)$  for  $s \in (\Gamma_k, \Gamma_{k+1})$ ,

$$\tilde{B}_k = \beta \int_{\Gamma_{k-1}}^{\Gamma_k} e^{X_s - X_{\Gamma_k}} ds.$$

Denote by  $(\tilde{A}, \tilde{B})$  a copy of  $(\tilde{A}_1, \tilde{B}_1)$ , independent of  $L$  and  $\tilde{V}_1$ . Then it follows that

$$\tilde{A} \stackrel{d}{=} e^{-X_{\Gamma_1}} \quad \text{and} \quad \tilde{B} \stackrel{d}{=} \beta e^{-X_{\Gamma_1}} \int_0^{\Gamma_1} e^{X_s} ds, \quad k \in \mathbb{N}. \quad (4.16)$$

Moreover,

$$\tilde{V}_k = \tilde{V}_1 \prod_{j=2}^k \tilde{A}_j + \sum_{i=2}^k \tilde{B}_i \prod_{j=i+1}^k \tilde{A}_j = e^{-X_{\Gamma_k}} \left[ \tilde{V}_1 + \sum_{i=2}^k \tilde{B}_i e^{X_{\Gamma_i}} \right], \quad k \geq 2.$$

We are now ready to present the analogue of Theorem 4.1 for the sequence  $(\tilde{V}_k)_{k \in \mathbb{N}}$ . As can be seen from (4.8), the process  $(X_{\Gamma_k})_{k \in \mathbb{N}}$  is a random walk with increments

$$X_{\Gamma_k} - X_{\Gamma_{k-1}} = \eta(\Gamma_k - \Gamma_{k-1}) - \log(1 + \lambda e^\eta \xi_k^2), \quad k \in \mathbb{N}.$$

**Theorem 4.3 (Pareto tail behavior of  $\tilde{V}$ ).**

*Suppose there exists some  $\alpha > 0$  such that*

$$\mathbb{E}|L_1|^{2\alpha} \log^+ |L_1| < \infty \quad \text{and} \quad \Psi(\alpha) = 0. \quad (4.17)$$

*Then a stationary solution  $(\tilde{V}_k)_{k \in \mathbb{N}}$  of (4.13) exists. Its marginal stationary distribution, denoted by  $\tilde{V}_0$ , is the unique solution of the random fix point equation*

$$\tilde{V}_0 \stackrel{d}{=} \tilde{A} \tilde{V}_0 + \tilde{B},$$

where  $(\tilde{A}, \tilde{B})$  is given by (4.16) and independent of  $L$  and  $\tilde{V}_0$ . Furthermore,

$$\mathbb{P}(\tilde{V}_0 > x) \sim \tilde{C}x^{-\alpha}, \quad x \rightarrow \infty,$$

where

$$\tilde{C} = \frac{\mathbb{E} \left[ (\tilde{A}\tilde{V}_0 + \tilde{B})^\alpha - (\tilde{A}\tilde{V}_0)^\alpha \right]}{\alpha \mathbb{E}|\tilde{A}|^\alpha \log^+ |\tilde{A}|} > 0. \quad (4.18)$$

*Proof.* We shall show that conditions (i)-(iv) of Theorem C.1 are satisfied: by definition,  $\log \tilde{A} \stackrel{d}{=} -\eta\Gamma_1 + \log(1 + \lambda e^\eta \xi_1^2)$ , where  $\Gamma_1$  is exponentially distributed. Consequently, (i) follows.

To show (ii) note that by the independence of  $\Gamma_1$  and  $\xi_1$ , for  $\alpha > 0$  we have by (4.9)

$$\begin{aligned} \mathbb{E}|\tilde{A}|^\alpha &= \mathbb{E}e^{-\alpha\eta\Gamma_1} \mathbb{E}(1 + \lambda e^\eta \xi_1^2)^\alpha \\ &= \frac{\mu}{\mu + \alpha\eta} \frac{\mu + \alpha\eta + \Psi(\alpha)}{\mu} \\ &= 1 + \frac{1}{\mu + \alpha\eta} \Psi(\alpha) = 1, \end{aligned}$$

by the second assumption in (4.17).

In order to prove (iii) note that

$$\mathbb{E}|\tilde{A}|^\alpha \log^+ |\tilde{A}| \leq \mathbb{E}|1 + \lambda e^\eta \xi_1^2|^\alpha \log^+ (1 + \lambda e^\eta \xi_1^2) < \infty,$$

if and only if the first assumption in (4.17) holds, see Sato [35], Theorem 25.3. Finally, (iv) follows from

$$\mathbb{E}|\tilde{B}|^\alpha \leq (\beta/\eta)^\alpha \mathbb{E}|1 + \lambda e^\eta \xi_1^2|^\alpha < \infty.$$

That the constant  $\tilde{C}$  is indeed strictly positive follows from the fact that  $\tilde{A}$ ,  $\tilde{B}$  and  $\tilde{V}_0$  are strictly positive, almost surely.  $\square$

**Remark 4.4.**

(i)  $X_{\Gamma_k}$  tends almost surely to  $\infty$  if and only if  $\mathbb{E}X_{\Gamma_1} > 0$  or, equivalently,  $\mu \mathbb{E} \log(1 + \lambda e^\eta \xi_1^2) < \eta$ . Notice that the stationarity condition (4.3) is for this model equivalent to  $\mathbb{E}X_{\Gamma_1} > 0$ .

(ii) In a sense it is remarkable that the tail of the stationary r. v. of the continuous-time model  $V_0$  and of the discrete-time skeleton  $\tilde{V}_0$  are so similar. As the discrete-time skeleton considers only local suprema of the process, one expects  $\tilde{V}_0$  to be

stochastically larger. As the Pareto index  $\alpha$  is the same for both models, any difference can only appear in the constants  $C$  and  $\tilde{C}$ . Brockwell, Chadraa and Lindner [11] have established a precise relationship between the distributions of  $V_0$  and  $\tilde{V}_0$ , showing that

$$\left(V_0 - \frac{\beta}{\eta}\right) \stackrel{d}{=} e^{-\eta\Gamma} \left(\tilde{V}_0 - \frac{\beta}{\eta}\right),$$

where  $\Gamma \stackrel{d}{=} \Gamma_1$  is exponentially distributed with parameter  $\mu$  and independent of  $\tilde{V}_0$ . Using the classical result of Breiman [10], it then follows that

$$C = \mathbb{E} \left( e^{-\eta\Gamma} \right)^\alpha \tilde{C} = \frac{\mu}{\mu + \alpha\eta} \tilde{C} = \frac{1}{\mathbb{E}(1 + \lambda e^{\eta\xi_1^2})^\alpha} \tilde{C}, \quad (4.19)$$

where the last equation follows from (4.9).  $\square$

The extremal behavior of solutions to stochastic recurrence equations is studied in de Haan et al. [22]. Their results can be applied to the stationary discrete-time skeleton of the volatility process  $(\tilde{V}_k)_{k \in \mathbb{N}}$  as defined in (4.12).

**Theorem 4.5 (Extremal behavior of the COGARCH model).**

Let  $V$  be a stationary version of the volatility process given by (4.1) and define  $M(T) := \sup_{0 \leq t \leq T} V_t$  for  $T > 0$ . Suppose there exists some  $\alpha > 0$  such that

$$\mathbb{E}|L_1|^{2\alpha} \log^+ |L_1| < \infty \quad \text{and} \quad \Psi(\alpha) = 0.$$

Let  $\tilde{C}$  be the constant in (4.18) and define  $a_T := (\mu T)^{1/\alpha}$  for  $T > 0$ . Then

$$\lim_{T \rightarrow \infty} \mathbb{P}(a_T^{-1} M(T) \leq x) = \exp(-\tilde{C}\theta x^{-\alpha}), \quad x > 0,$$

where for  $X_t^+ = \max\{0, X_t\}$

$$\theta = 1 - \mathbb{E} \left[ \sup_{t \geq \Gamma_1} \{e^{-\alpha X_t^+}\} \right] \in (0, 1).$$

Denote by  $(\Gamma_k)_{k \in \mathbb{N}}$  the jump times of the compound Poisson process  $L$  given by (3.5) and define  $I_k := (\Gamma_k, \Gamma_{k+1}]$  for  $k \in \mathbb{N}$ . Let  $(j_k)_{k \in \mathbb{N}}$  be the jump times of a Poisson process with intensity  $\tilde{C}\theta x^{-\alpha}$  for  $x > 0$ . Then,

$$\sum_{k=1}^{\infty} \varepsilon \left\{ \frac{\Gamma_k}{T}, a_T^{-1} \sup_{s \in I_k} V_s \right\} (\cdot \times (x, \infty)) \xrightarrow{T \rightarrow \infty} \sum_{k=1}^{\infty} \zeta_k \varepsilon \{j_k\}, \quad (4.20)$$

where  $(\zeta_k)_{k \in \mathbb{N}}$  are i. i. d. discrete r. v. s, independent of  $(j_k)_{k \in \mathbb{N}}$ , with probability distribution

$$\pi_k = \mathbb{P}(\zeta_1 = k) = (\theta_k - \theta_{k+1})/\theta, \quad k \in \mathbb{N}.$$

Moreover,

$$\begin{aligned}\theta_k &:= \mathbb{E}[\exp(\alpha \min\{T_{k-1}, 0\}) - \exp(\alpha \min\{T_k, 0\})] \\ &= \int_0^1 \mathbb{P}\left(\text{card}\left\{j \in \mathbb{N} : e^{-\alpha X_{\Gamma_j}} > y\right\} = k-1\right) dy,\end{aligned}$$

where  $\infty = T_0 \geq T_1 \geq \dots$  are the ordered values of the sequence  $(-X_{\Gamma_k}^+)_{k \in \mathbb{N}}$ . Finally,  $\theta = \theta_1$ .

*Proof.* Since  $\sup_{s \in I_k} V_s = \tilde{V}_k$ , Theorem 4.3 and de Haan et al. [22], Theorem 2.1, show that

$$\sum_{k=1}^{\infty} \varepsilon \left\{ \frac{k}{\mu n}, a_n^{-1} \tilde{V}_k \right\} (\cdot \times (x, \infty)) \xrightarrow{T \rightarrow \infty} \sum_{k=1}^{\infty} \zeta_k \varepsilon \{j_k\}$$

and that  $(\tilde{V}_k)_{k \in \mathbb{N}}$  has extremal index  $\theta \in (0, 1)$ , given by

$$\begin{aligned}\theta &= \alpha \int_1^{\infty} \mathbb{P}\left(\bigvee_{j=1}^{\infty} \prod_{k=1}^j \tilde{A}_k \leq y^{-1}\right) y^{-\alpha-1} dy \\ &= \alpha \int_1^{\infty} \mathbb{P}\left(\bigvee_{j=1}^{\infty} \exp(-X_{\Gamma_j}) \leq y^{-1}\right) y^{-\alpha-1} dy \\ &= \int_0^1 \mathbb{P}\left(\sup_{t \geq \Gamma_1} \{e^{-\alpha X_t}\} \leq z\right) dz \\ &= 1 - \mathbb{E}\left[\min\left\{1, \sup_{t \geq \Gamma_1} \{e^{-\alpha X_t}\}\right\}\right].\end{aligned}$$

For the first expression for  $\theta_k$ , see de Haan et al. [22], and the second expression follows by a similar calculation as above. By an application of Hsing and Teugels [23] (see also Fasen [18], Lemma 2.2.4) we transform the time scale, such that (4.20) holds. Then we obtain

$$\begin{aligned}\lim_{T \rightarrow \infty} \mathbb{P}(a_T^{-1} M(T) \leq x) &= \lim_{T \rightarrow \infty} \mathbb{P}\left(\sum_{k=1}^{\infty} \varepsilon \left\{ \frac{\Gamma_k}{T}, a_T^{-1} \sup_{s \in I_k} V_s \right\} ((0, 1) \times (x, \infty)) = 0\right) \\ &= \mathbb{P}\left(\sum_{k=1}^{\infty} \zeta_k \varepsilon \{j_k\}((0, 1)) = 0\right) = \exp(-\tilde{C}\theta x^{-\alpha}).\end{aligned}$$

□

By the Poisson result (4.20) we observe clusters in local extremes of the continuous-time process. So the COGARCH is a suitable model for heavy tailed volatility models with clusters on high levels.

## 5 Conclusion

In this paper we have investigated the extremal behavior of the most popular continuous-time volatility models. We have concentrated on models with tails ranging from exponential to regularly varying; i. e. tails as they are found in empirical volatility. The quantities derived for such models include

- the tail of the stationary volatility  $V_0$  and the relation to the tail of the distribution governing the extreme behavior,
- the asymptotic distribution of the running maxima, i. e. their MDA and the norming constants,
- the cluster behavior of the model on high levels.

We found interesting similarities in the extremal behavior of certain models, which was quite unexpected.

Recall the GCIR model of Example 2.3, where the tail of the stationary distribution  $F$  of  $V_0$  is compared to the tail of  $H$ , the d.f. describing the extreme behavior. Example 2.3 (2) belongs to  $\mathcal{S} \cap \text{MDA}(\Lambda)$ , it has stationary distribution with a semi-heavy Weibull like tail. Relation (3.11) is mimicked by the fact that (2.6) can be rewritten to

$$\bar{F}(x) \sim \frac{a(x)}{x} \bar{H}(x) \quad x \rightarrow \infty.$$

Moreover, as the norming constants in the GCIR examples are calculated based on the d.f.  $H$ , analogously, by the above Corollary 3.6, for the OU process in  $\text{MDA}(\Lambda)$ , the norming constants are derived from  $L_1$  and not from the stationary distribution of the process  $V$ .

Analogous results hold for Example 2.3 (3), which belongs to  $\mathcal{S} \cap \text{MDA}(\Phi_\alpha)$  for some  $\alpha > 0$ . Here the tails of  $F$  and  $H$  are both regularly varying of the same index; this corresponds to (3.8).

Also in the case, where  $V_0$  is gamma distributed, the behavior of the running maxima of the GCIR model in Example 2.3 (1) and of the  $\Gamma$ -OU process as given in (3.26) and (3.33), respectively, are identical.

This means also that, if the stationary distribution of a GCIR model coincides with the stationary distribution of a subexponential OU model, then also the norming constants and the behavior of the running maxima coincide. The role of  $M(h)$  in  $\mathcal{S}(\gamma)$ ,  $\gamma \geq 0$ , corresponds for the GCIR models to the d.f.  $H$ ; the influence of the driving Brownian motion plays no role whatsoever for the extreme behavior.

Concerning volatility clusters, no model in  $\text{MDA}(\Lambda)$  presented in this paper can model such clusters on high levels. And here there is a profound difference between the GCIR models with regularly varying tails and the regularly varying OU models. Whereas all GCIR models fail to model high level volatility clusters, regularly varying OU models have the potential to model them.

The COGARCH model resembles the GCIR models only in the sense that heavy tails occur, although the driving process can be very light tailed; the difference being that the COGARCH model always has heavy tails. There is no obvious relationship between the tail behavior of the stationary r. v.  $V_0$  and  $L_1$ ; the heavy tails occur by the very intrinsic dependence structure of the model.

With respect to volatility clusters, only regularly varying OU processes and COGARCH processes exhibit volatility clusters on high levels, which can be described quite precisely by the distribution of the cluster sizes; see Corollary 3.5 and Theorem 4.5.

In this paper we have refrained from discussing another important stylized fact of empirical volatility: it exhibits often long memory in the sense that the autocovariance function decreases very slowly. This phenomenon can have various reasons, as for instance discussed in Mikosch and Stărică [31]. Certainly an important issue is here a possible non-stationarity of the data. On the other hand, long range dependence is an important fact, which should not be completely ignored. All models presented in this paper have an exponentially decreasing covariance functions, which only exhibit some visual long memory, when the process is close to non-stationarity.

For diffusion models like the GCIR models, a remedy, which introduces long range dependence in such models, is to replace the driving Brownian motion by a fractional Brownian motion. This generates a new class of stationary long memory models. Such models have been suggested and analyzed in [12, 13].

For the OU process the exponentially decreasing covariance function is due to the exponential kernel function; see (3.3). The often observed long-range dependence effect in the empirical volatility cannot be modelled this way. There are two ways to introduce long memory into such models. The first one is to replace the exponential kernel function by a hyperbolic kernel function of the form  $f(x) \sim |x|^{-\beta}$  for large  $x$  and for some  $\beta \in (0.5, 1)$ . This introduces long memory into the model, which can be modelled by regularly varying Lévy driven MA processes. The second method has been suggested by Barndorff-Nielsen and Shephard [3]: a superposition of several OU processes (supOU processes) can create long memory. Regularly varying supOU processes exhibit also volatility clusters; see Fasen [19].

## Acknowledgement

The first author acknowledges gratefully the financial support by the Deutsche Forschungsgemeinschaft through the graduate program “Angewandte Algorithmische Mathematik” at the Munich University of Technology. The third author takes pleasure in thanking the Deutsche Forschungsgemeinschaft for a research grant, and the Department of Statistics at Colorado State University for its pleasant hospitality.

## Appendix

In this Appendix we summarize some definitions and concepts used throughout the paper.

### A Basic notation and definition

For details and further references see Embrechts et al. [17].

**Definition A.1.** A positive measurable function  $u : \mathbb{R} \rightarrow \mathbb{R}_+$  is regularly varying with index  $\alpha$ , denoted by  $u \in \mathcal{R}_\alpha$  for  $\alpha \in \mathbb{R}$ , if

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

The function  $u$  is said to be slowly varying if  $\alpha = 0$ , and rapidly varying, denoted by  $u \in \mathcal{R}_{-\infty}$ , if the above limit is 0 for  $x > 1$  and  $\infty$  for  $0 < x < 1$ .

**Definition A.2.** A d.f.  $F$  belongs to the class  $\mathcal{L}(\gamma)$ ,  $\gamma \geq 0$ , if for every  $y \in \mathbb{R}$ ,

$$\lim_{x \rightarrow \infty} \bar{F}(x - y)/\bar{F}(x) = e^{\gamma y}.$$

The class  $\mathcal{L}(\gamma)$  is related to the class  $\mathcal{R}_{-\gamma}$  by the fact that

$$F \in \mathcal{L}(\gamma) \quad \text{if and only if} \quad \bar{F} \circ \log \in \mathcal{R}_{-\gamma}.$$

Thus the convergence of  $\bar{F}(x - y)/\bar{F}(x)$  in Definition A.2 is uniform on compact  $y$ -intervals. For an excellent monograph on regular variation we refer to Bingham, Goldie and Teugels [7].

Applying Karamata’s representation for regularly varying functions to the class  $\mathcal{L}(\gamma)$  we obtain for  $F \in \mathcal{L}(\gamma)$ ,  $\gamma \geq 0$ , the representation

$$\bar{F}(x) = c(x) \exp \left[ - \int_0^x \frac{1}{a(y)} dy \right], \quad x > 0, \tag{A.1}$$



where  $a, c : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  and  $\lim_{x \rightarrow \infty} c(x) = c > 0$  and  $a$  is absolutely continuous with  $\lim_{x \rightarrow \infty} a(x) = 1/\gamma$  and  $\lim_{x \rightarrow \infty} a'(x) = 0$ .

**Definition A.3 (Convolution equivalent distributions).**

Let  $\gamma \geq 0$  and  $X$  have d.f.  $F$ . We say that  $F$  or  $X$  belongs to the class  $\mathcal{S}(\gamma)$ , if the following properties hold.

- (i)  $F \in \mathcal{L}(\gamma)$ ,
- (ii)  $\lim_{x \rightarrow \infty} \frac{\overline{F^{*2}}(x)}{\overline{F}(x)} = 2\hat{f}(\gamma) < \infty$ ,

where  $\hat{f}(\gamma) = \mathbb{E}e^{\gamma X}$  is the moment generating function of  $X$  at  $\gamma$ . The class  $\mathcal{S} := \mathcal{S}(0)$  is called the class of subexponential distributions.

**Theorem A.4.**

- (i) Let  $F$  be infinitely divisible with Lévy measure  $\nu$  and  $\gamma \geq 0$ . Then

$$F \in \mathcal{S}(\gamma) \iff \nu(1, \cdot] / \nu(1, \infty) \in \mathcal{S}(\gamma) \iff \lim_{x \rightarrow \infty} \overline{F}(x) / \nu(x, \infty) = \hat{f}(\gamma).$$

- (ii) Suppose  $F \in \mathcal{S}(\gamma)$ ,  $\lim_{x \rightarrow \infty} \overline{F}_i(x) / \overline{F}(x) = q_i \geq 0$  and  $\hat{f}_i(\gamma) < \infty$  for  $i = 1, 2$ . Then

$$\lim_{x \rightarrow \infty} \frac{\overline{F_1 * F_2}(x)}{\overline{F}(x)} = q_1 \hat{f}_2(\gamma) + q_2 \hat{f}_1(\gamma).$$

If  $q_i > 0$  for some  $i \in \{1, 2\}$ , then also  $F_i, F_1 * F_2 \in \mathcal{S}(\gamma)$ .

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

$$\overline{G}(x) \sim \mu \hat{f}(\gamma) \overline{F}(x), \quad x \rightarrow \infty.$$

The following is the fundamental theorem in extreme value theory.

**Theorem A.5 (Fisher-Tippett Theorem).**

Let  $(X_n)_{n \in \mathbb{N}}$  be an i.i.d. sequence with d.f.  $F$  and denote  $M_n = \max_{k=1, \dots, n} X_k$ . Suppose we can find sequences of real numbers  $a_n > 0, b_n \in \mathbb{R}$  such that

$$\lim_{n \rightarrow \infty} \mathbb{P}(a_n^{-1}(M_n - b_n) \leq x) = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G(x), \quad x \in \mathbb{R},$$

for some non-degenerate d.f.  $G$  (we say  $F$  is in the maximum domain of attraction of  $G$  and write  $F \in \text{MDA}(G)$ ). Then there are  $a > 0, b \in \mathbb{R}$  such that  $x \mapsto G(ax + b)$  is one of the following three extreme value d.f. s:

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

We summarize some well-known facts related to domains of attraction.

**Proposition A.6.**

(a) *The following Poisson characterization holds:  $F \in \text{MDA}(G)$  if and only if  $a_n > 0, b_n \in \mathbb{R}$  exist such that*

$$\lim_{n \rightarrow \infty} n\bar{F}(a_n x + b_n) = -\log G(x), \quad x > 0.$$

(b) *If  $F \in \mathcal{L}(\gamma)$  for  $\gamma > 0$ , then  $F \in \text{MDA}(\Lambda)$  with  $a_n \rightarrow 1/\gamma$  as  $n \rightarrow \infty$  and  $e^{b_n} \in \mathcal{R}_{1/\gamma}$ .*

(c) *If  $F \in \mathcal{S} \cap \text{MDA}(\Lambda)$ , then  $b_n \rightarrow \infty, a_n \rightarrow \infty$  and  $b_n/a_n \rightarrow \infty$  as  $n \rightarrow \infty$ .*

(d) *If  $F \in \text{MDA}(\Phi_\alpha) = \mathcal{R}_{-\alpha}$  for  $\alpha > 0$ , then  $b_n = 0, a_n \in \mathcal{R}_{1/\alpha}$  and  $a_n \rightarrow \infty$  as  $n \rightarrow \infty$ .*

The following concept has proved useful in comparing tails.

**Definition A.7 (Tail-equivalence).**

*Two d. f. s  $F$  and  $G$  (or two measures  $\mu$  and  $\nu$ ) are called tail-equivalent if both have support unbounded to the right and there exists some  $c > 0$  such that*

$$\lim_{x \rightarrow \infty} \bar{F}(x)/\bar{G}(x) = c \quad \text{or} \quad \lim_{x \rightarrow \infty} \nu(x, \infty)/\mu(x, \infty) = c.$$

Important in the context of our paper is that all the following classes are closed with respect to tail-equivalence:  $\text{MDA}(G)$  for  $G \in \{\Phi_\alpha, \alpha > 0, \Lambda\}$ ,  $\mathcal{R}_{-\alpha}$  for  $\alpha \in [0, \infty)$ ,  $\mathcal{L}(\gamma)$  for  $\gamma \geq 0$ ,  $\mathcal{S}(\gamma)$  for  $\gamma \geq 0$ . Moreover, for two tail-equivalent d. f. s in some  $\text{MDA}(G)$  one can choose the same norming constants.

**Definition A.8 (Poisson random measure).**

*Let  $(A, \mathcal{A}, \vartheta)$  be a measurable space, where  $\vartheta$  is  $\sigma$ -finite, and  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space. A Poisson random measure  $N$  with mean measure  $\vartheta$ , denoted by  $\text{PRM}(\vartheta)$ , is a collection of r. v. s  $(N(A))_{A \in \mathcal{A}}$ , where  $N(A) : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{N}_0, \mathcal{B}(\mathbb{N}_0))$ , with  $N(\emptyset) = 0$ , such that:*

(a) Given any sequence  $(A_n)_{n \in \mathbb{N}}$  of mutually disjoint sets in  $\mathcal{A}$ :

$$N\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} N(A_n) \quad a.s.$$

(b)  $N(A)$  is Poisson distributed with mean  $\vartheta(A)$  for every  $A \in \mathcal{A}$ .

(c) For mutually disjoint sets  $A_1, \dots, A_n \in \mathcal{A}$ ,  $n \in \mathbb{N}$ , the r. v. s  $N(A_1), \dots, N(A_n)$  are independent.

**Definition A.9 (Extremal index).**

Let  $X = (X_n)_{n \in \mathbb{Z}}$  be a strictly stationary sequence and  $\theta \geq 0$ . If for every  $\tau > 0$  there exists a sequence  $u_n(\tau)$  with

$$\lim_{n \rightarrow \infty} n\mathbb{P}(X_1 > u_n(\tau)) = \tau \quad \text{and} \quad \lim_{n \rightarrow \infty} \mathbb{P}\left(\max_{k=1, \dots, n} X_k \leq u_n(\tau)\right) = e^{-\theta\tau},$$

then  $\theta$  is called the extremal index of  $X$  and has value in  $[0, 1]$ .

## B The conditions $D_r(\mathbf{u}_n)$ and $D'(\mathbf{u}_n)$

Classical results for the extremal behavior of stationary sequences are based on two conditions: the first one is a specific type of asymptotic dependence, and the second is an anti-clustering condition.

**Definition B.1.** Let  $X = (X_n)_{n \in \mathbb{N}}$  be a strictly stationary sequence, such that for  $m = 1, \dots, r$ ,  $r \in \mathbb{N}$ , the sequences of constants  $(u_n^{(m)})_{n \in \mathbb{N}}$  and  $(u_n)_{n \in \mathbb{N}}$  satisfy  $\lim_{n \rightarrow \infty} n\bar{F}(u_n^{(m)}) = \tau^{(m)}$  and  $\lim_{n \rightarrow \infty} n\bar{F}(u_n) = \tau$ .

(a) For any integers  $p, q$  and  $n$  let

$$1 \leq i_1 < \dots < i_p < j_1 < \dots < j_q \leq n$$

such that  $j_1 - i_p \geq l$  and  $\mathbf{v}_n = (v_n^{(1)}, \dots, v_n^{(p)})$ ,  $\mathbf{w}_n = (w_n^{(1)}, \dots, w_n^{(q)})$  with  $v_n^{(l)}, w_n^{(s)} \in \{u_n^{(1)}, \dots, u_n^{(r)}\}$ . Write  $I = \{i_1, \dots, i_p\}$ ,  $J = \{j_1, \dots, j_q\}$ ,  $\mathbf{X}_I = (X_{i_1}, \dots, X_{i_p})$  and  $\mathbf{X}_J = (X_{j_1}, \dots, X_{j_q})$ . If for each choice of indices of  $I, J$ ,

$$|\mathbb{P}(\mathbf{X}_I \leq \mathbf{v}_n, \mathbf{X}_J \leq \mathbf{w}_n) - \mathbb{P}(\mathbf{X}_I \leq \mathbf{v}_n)\mathbb{P}(\mathbf{X}_J \leq \mathbf{w}_n)| \leq \alpha_{n,l},$$

where  $\alpha_{n,l_n} \rightarrow 0$  as  $n \rightarrow \infty$  for some sequence  $l_n = o(n)$ , then  $X$  satisfies the condition  $D_r(\mathbf{u}_n)$ .

(b)  $X$  satisfies the condition  $D'(u_n)$ , if

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} n \sum_{j=2}^{\lfloor n/k \rfloor} \mathbb{P}(X_1 > u_n, X_j > u_n) = 0.$$

We show that  $(\tilde{V}_k)_{k \in \mathbb{Z}}$  as defined in (3.28) satisfies the  $\mathbf{D}_r(\mathbf{u}_n)$  and  $D'(u_n)$  conditions. The result is an analogon for discrete-time MA processes given in Rootzén [33], Lemma 3.2.

**Lemma B.2.** *Let  $V$  be a stationary version of the OU process given by (3.6) with  $L$  a positive, compound Poisson process as in (3.5).*

(a) *Assume  $\tilde{V}_k = V_{\Gamma_k/\lambda} + e^{-\Gamma_k} \xi_0 \in \mathcal{L}(\gamma)$ ,  $\gamma > 0$ , such that for  $a_n > 0$ ,  $b_n \in \mathbb{R}$  and  $u_n = a_n x + b_n$ ,*

$$\lim_{n \rightarrow \infty} n \mathbb{P}(\tilde{V}_k > u_n) = e^{-x}, \quad x \in \mathbb{R}. \quad (\text{B.1})$$

*For  $r \in \mathbb{N}$  and  $\mathbf{x} = (x_1, \dots, x_r)$  let  $\mathbf{u}_n = (a_n x_1 + b_n, \dots, a_n x_r + b_n)$ . Then  $(\tilde{V}_k)_{k \in \mathbb{N}}$  satisfies the  $\mathbf{D}_r(\mathbf{u}_n)$  and  $D'(u_n)$  conditions.*

(b) *Let  $L_1$  be in  $\mathcal{S}(\gamma)$ ,  $\gamma > 0$ . Define for  $h > 0$  fixed  $M_k := \sup_{(k-1)h \leq t \leq kh} V_t$  for  $k \in \mathbb{N}$ . Suppose  $M_1 \in \mathcal{L}(\gamma)$  such that for  $a_n > 0$ ,  $b_n \in \mathbb{R}$  and  $u_n = a_n x + b_n$ ,*

$$\lim_{n \rightarrow \infty} n \mathbb{P}(M_k > u_n) = e^{-x}, \quad x \in \mathbb{R}.$$

*For  $r \in \mathbb{N}$  and  $\mathbf{x} = (x_1, \dots, x_r)$  let  $\mathbf{u}_n = (a_n x_1 + b_n, \dots, a_n x_r + b_n)$ . Then  $(M_k)_{k \in \mathbb{N}}$  satisfies the  $\mathbf{D}_r(\mathbf{u}_n)$  and  $D'(u_n)$  conditions.*

*Proof.* (a) To show the  $\mathbf{D}_r(\mathbf{u}_n)$  condition, let  $u_n^{(m)} = a_n x_m + b_n$ ,  $x_m \in \mathbb{R}$ ,  $m = 1, \dots, r$ . Let  $\mathbf{v}_n = (v_n^{(1)}, \dots, v_n^{(p)})$ ,  $\mathbf{w}_n = (w_n^{(1)}, \dots, w_n^{(q)})$  with  $v_n^{(l)}, w_n^{(s)} \in \{u_n^{(1)}, \dots, u_n^{(r)}\}$  and  $1 \leq i_1 < \dots < i_p < j_1 < \dots < j_q \leq n$ . Define  $V_k^n := \sum_{j=k-n}^k e^{-(\Gamma_k - \Gamma_j)} \xi_j$ ,  $\mathbf{V}_I := (\tilde{V}_{i_1}, \dots, \tilde{V}_{i_p})$ ,  $\mathbf{V}_I^n := (V_{i_1}^n, \dots, V_{i_p}^n)$  and, similarly,  $\mathbf{V}_J := (\tilde{V}_{j_1}, \dots, \tilde{V}_{j_q})$ ,  $\mathbf{V}_J^n := (V_{j_1}^n, \dots, V_{j_q}^n)$ . Then  $(V_k^n)_{k \in \mathbb{Z}}$  is stationary and  $\mathbf{V}_I^{[n\delta]}$  is independent of  $\mathbf{V}_J^{[n\delta]}$  for  $j_1 - i_p > [n\delta]$ ,  $\delta > 0$ . Since  $\mathbb{P}(\xi_k < 0) = 0$ , we obtain  $\mathbf{V}_I^n \leq \mathbf{V}_I$ . Here,  $\mathbf{x} = (x_1, \dots, x_p) \leq \mathbf{y} = (y_1, \dots, y_p)$  means that  $x_i \leq y_i$  for all  $i = 1, \dots, p$ . It now

follows that for any  $\epsilon > 0$ ,

$$\begin{aligned}
\mathbb{P}(\mathbf{V}_I \leq \mathbf{v}_n, \mathbf{V}_J \leq \mathbf{w}_n) &\leq \mathbb{P}\left(\mathbf{V}_I^{[n\delta]} \leq \mathbf{v}_n, \mathbf{V}_J^{[n\delta]} \leq \mathbf{w}_n\right) \\
&= \mathbb{P}\left(\mathbf{V}_I^{[n\delta]} \leq \mathbf{v}_n\right) \mathbb{P}\left(\mathbf{V}_J^{[n\delta]} \leq \mathbf{w}_n\right) \\
&\leq \mathbb{P}\left(\mathbf{V}_I \leq \mathbf{v}_n + \epsilon(a_n, \dots, a_n)\right) \mathbb{P}\left(\mathbf{V}_J \leq \mathbf{w}_n + \epsilon(a_n, \dots, a_n)\right) \\
&\quad + n\mathbb{P}\left(|\tilde{V}_1 - V_1^{[n\delta]}| > \epsilon a_n\right) \\
&\leq \mathbb{P}\left(\mathbf{V}_I \leq \mathbf{v}_n\right) \mathbb{P}\left(\mathbf{V}_J \leq \mathbf{w}_n\right) + n\mathbb{P}\left(|\tilde{V}_1 - V_1^{[n\delta]}| > \epsilon a_n\right) \\
&\quad + \sum_{m=1}^r n\mathbb{P}\left(u_n^{(m)} \leq \tilde{V}_1 \leq u_n^{(m)} + \epsilon a_n\right).
\end{aligned}$$

Similarly we can find a lower bound, such that for  $j_1 - i_p > [n\delta]$ ,

$$\begin{aligned}
\alpha_{n, [n\delta]} &:= |\mathbb{P}(\mathbf{V}_I \leq \mathbf{v}_n, \mathbf{V}_J \leq \mathbf{w}_n) - \mathbb{P}(\mathbf{V}_I \leq \mathbf{v}_n) \mathbb{P}(\mathbf{V}_J \leq \mathbf{w}_n)| \\
&\leq n\mathbb{P}\left(|\tilde{V}_1 - V_1^{[n\delta]}| > \epsilon a_n\right) + \sum_{m=1}^r n\mathbb{P}\left(u_n^{(m)} - \epsilon a_n \leq \tilde{V}_1 \leq u_n^{(m)} + \epsilon a_n\right) \\
&=: \tilde{\alpha}_{n, [n\delta], \epsilon}.
\end{aligned} \tag{B.2}$$

Let  $X_i := \Gamma_i - \Gamma_{i-1} - 1/\mu$ ,  $i \in \mathbb{N}$ . Then  $(X_i)_{i \in \mathbb{N}}$  is a centered i.i.d. sequence such that  $\sum_{i=1}^n X_i = \Gamma_n - n/\mu$ . It follows that there exists a constant  $K > 0$ , such that for every  $n \in \mathbb{N}$ ,

$$\begin{aligned}
\mathbb{E}(\Gamma_n - n/\mu)^6 &= n\mathbb{E}X_1^6 + \binom{n}{3} \binom{6}{2} \binom{4}{2} (\mathbb{E}X_1^2)^3 \\
&\quad + \binom{n}{2} \binom{6}{3} (\mathbb{E}X_1^3)^2 + \binom{n}{2} \binom{6}{2} (\mathbb{E}X_1^2)(\mathbb{E}X_1^4) \\
&\leq n^3 K.
\end{aligned}$$

Hence, by Markov's inequality there is a constant  $\tilde{K} > 0$  such that for all  $n \in \mathbb{N}$ ,

$$\mathbb{P}(\Gamma_n < n/(2\mu)) \leq \mathbb{P}\left(|\Gamma_n - n/\mu| > n/(2\mu)\right) \leq (2\mu/n)^6 \mathbb{E}(\Gamma_n - n/\mu)^6 \leq \tilde{K}/n^3. \tag{B.3}$$

Thus we obtain for  $n \in \mathbb{N}$ ,

$$\begin{aligned}
n\mathbb{P}\left(|\tilde{V}_0 - V_0^{[n\delta]}| > \epsilon a_n\right) &= n\mathbb{P}\left(e^{-\Gamma_{[n\delta]}} \sum_{j=-\infty}^{-[n\delta]-1} e^{(\Gamma_{[n\delta]} + \Gamma_j)} \xi_j > \epsilon a_n\right) \\
&\leq n\mathbb{P}\left(e^{-[n\delta]/(2\mu)} \tilde{V}_1 > \epsilon a_n\right) + \frac{\tilde{K}}{n^2}.
\end{aligned} \tag{B.4}$$

Note that the first exponential moment of  $\gamma\beta e^{-\lfloor\delta n\rfloor/(2\mu)}\tilde{V}_1$  exists for  $\beta e^{-\lfloor\delta n\rfloor/(2\mu)} < 1$ . Choose  $\beta_n = 2/(\epsilon(1-\epsilon))\log n$ . There exists  $n_0 = n_0(\delta, \epsilon) \in \mathbb{N}$  such that  $\beta_n e^{-\lfloor n\delta\rfloor/(2\mu)} < (1-\epsilon)$  and  $a_n \geq (1-\epsilon)/\gamma$  for  $n \geq n_0$ . The first term on the right hand side of (B.4) is by Markov's inequality for  $n \geq n_0$  bounded above by

$$n\mathbb{E} \exp \left[ \beta_n \gamma \left( e^{-\lfloor\delta n\rfloor/(2\mu)} \tilde{V}_1 \right) \right] e^{-\beta_n \epsilon \gamma a_n} \leq n\mathbb{E} \exp \left[ (1-\epsilon)\gamma \tilde{V}_1 \right] e^{-2\log n}, \quad (\text{B.5})$$

which converges to 0 as  $n \rightarrow \infty$ . Together with (B.1), (B.2) and (B.4) this gives

$$\lim_{n \rightarrow \infty} \tilde{\alpha}_{n, \lfloor n\delta \rfloor, \epsilon} = \sum_{m=1}^r [e^{-(x_m - \epsilon)} - e^{-(x_m + \epsilon)}],$$

so that

$$\lim_{n \rightarrow \infty} \alpha_{n, \lfloor n\delta \rfloor} \leq \lim_{\epsilon \downarrow 0} \lim_{n \rightarrow \infty} \tilde{\alpha}_{n, \lfloor n\delta \rfloor, \epsilon} = 0,$$

which implies the  $\mathbf{D}_r(\mathbf{u}_n)$  condition by Lemma 3.2.1 in Leadbetter et al. [29].

To show the  $D'(u_n)$  condition, let  $\epsilon > 0$ . Then there exists an  $x_0 > 0$  such that  $\mathbb{P}(\Gamma_2 - \Gamma_1 < x_0) = \epsilon$ . Since  $(\Gamma_{i+1} - \Gamma_i)_{i \in \mathbb{N}}$  is a positive i.i.d. sequence, it follows that

$$\mathbb{P}(\Gamma_j - \Gamma_1 < x_0) \leq \mathbb{P}(\Gamma_j - \Gamma_{j-1} < x_0, \dots, \Gamma_2 - \Gamma_1 < x_0) = \epsilon^{j-1}, \quad j \geq 2.$$

Now choose  $\beta$  such that  $1/2 < \beta < (1 + e^{-x_0})^{-1}$  and  $\delta > 0$  such that  $1 + \delta < 2\beta$ . Then, for any  $k, n \in \mathbb{N}$ ,

$$\sum_{j=2}^{\lfloor n/k \rfloor} \mathbb{P}(\tilde{V}_1 > u_n, \tilde{V}_j > u_n) = \sum_{j=2}^{\lfloor n^\delta \rfloor} \mathbb{P}(\tilde{V}_1 > u_n, \tilde{V}_j > u_n) + \sum_{j=\lfloor n^\delta \rfloor + 1}^{\lfloor n/k \rfloor} \mathbb{P}(\tilde{V}_1 > u_n, \tilde{V}_j > u_n). \quad (\text{B.6})$$

We first show that the first summand, when multiplied by  $n$ , tends to 0 as  $n \rightarrow \infty$ . Note that by the independence of  $\tilde{V}_1$  and  $\Gamma_j - \Gamma_1$ ,

$$\mathbb{P}(\tilde{V}_1 > u_n, \tilde{V}_j > u_n) \leq \epsilon^{j-1} \mathbb{P}(\tilde{V}_1 > u_n) + \mathbb{P}(\tilde{V}_1 + \tilde{V}_j > 2u_n, \Gamma_j - \Gamma_1 \geq x_0). \quad (\text{B.7})$$

Let  $\tilde{V}'_1$  be an independent copy of  $\tilde{V}_1$ . Then

$$\begin{aligned} & \mathbb{P} \left( \tilde{V}_1 + \tilde{V}_j > 2u_n, \Gamma_j - \Gamma_1 \geq x_0 \right) \\ &= \mathbb{P} \left( (1 + e^{-(\Gamma_j - \Gamma_1)}) \tilde{V}_1 + \sum_{k=2}^j e^{-(\Gamma_j - \Gamma_k)} \xi_k > 2u_n, \Gamma_j - \Gamma_1 \geq x_0 \right) \\ &\leq \mathbb{P} \left( (1 + e^{-x_0}) \tilde{V}_1 + \sum_{k=2}^j e^{-(\Gamma_j - \Gamma_k)} \xi_k > 2u_n \right) \\ &\leq \mathbb{P} \left( (1 + e^{-x_0}) \tilde{V}_1 + \tilde{V}'_1 > 2u_n \right). \end{aligned} \quad (\text{B.8})$$

As  $\beta(1 + e^{-x_0}) < 1$ , the first exponential moment of  $\beta\gamma((1 + e^{-x_0})\tilde{V}_1 + \tilde{V}'_1)$  exists, and by Markov's inequality the last expression can be bounded above by

$$\mathbb{E} \exp \left[ \beta\gamma((1 + e^{-x_0})\tilde{V}_1 + \tilde{V}'_1) \right] e^{-2\beta\gamma u_n}. \quad (\text{B.9})$$

Recall that  $u_n = u_n(x)$ . Since  $(n \mapsto e^{-2\beta\gamma u_n}) \in \mathcal{R}_{-2\beta}$  for fixed  $x$ , it follows that  $(n \mapsto n^{1+\delta}e^{-2\beta\gamma u_n}) \in \mathcal{R}_{(1+\delta)-2\beta}$ . We then obtain by (B.1), (B.7)-(B.9) and Bingham et al. [7], Proposition 1.5.1, that

$$\begin{aligned} & n \sum_{j=2}^{\lfloor n^\delta \rfloor} \mathbb{P}(\tilde{V}_1 > u_n, \tilde{V}_j > u_n) \\ & \leq n \mathbb{P}(\tilde{V}_1 > u_n) \sum_{j=2}^{\lfloor n^\delta \rfloor} \epsilon^{j-1} + \mathbb{E}[\beta\gamma((1 + e^{-x_0})\tilde{V}_1 + \tilde{V}'_1)] n^{1+\delta} e^{-2\beta\gamma u_n} \\ & \rightarrow e^{-x} \frac{\epsilon}{1 - \epsilon}, \quad n \rightarrow \infty. \end{aligned} \quad (\text{B.10})$$

For the second term in (B.6), using the independence of  $V_j^{\lfloor n^\delta \rfloor}$  and  $V_1^{\lfloor n^\delta \rfloor}$  for  $j > \lfloor n^\delta \rfloor$ , we obtain

$$\begin{aligned} & n \sum_{j=\lfloor n^\delta \rfloor+1}^{\lfloor n/k \rfloor} \mathbb{P}(\tilde{V}_1 > u_n, \tilde{V}_j > u_n) \\ & \leq n \sum_{j=\lfloor n^\delta \rfloor+1}^{\lfloor n/k \rfloor} \mathbb{P}(V_1^{\lfloor n^\delta \rfloor} > u_n - \epsilon a_n, V_j^{\lfloor n^\delta \rfloor} > u_n - \epsilon a_n) + 2n^2 \mathbb{P}(|\tilde{V}_1 - V_1^{\lfloor n^\delta \rfloor}| > \epsilon a_n) \\ & = n \sum_{j=\lfloor n^\delta \rfloor+1}^{\lfloor n/k \rfloor} \mathbb{P}(V_1^{\lfloor n^\delta \rfloor} > u_n - \epsilon a_n) \mathbb{P}(V_j^{\lfloor n^\delta \rfloor} > u_n - \epsilon a_n) + 2n^2 \mathbb{P}(|\tilde{V}_1 - V_1^{\lfloor n^\delta \rfloor}| > \epsilon a_n) \\ & \leq (n^2/k) \mathbb{P}(\tilde{V}_1 > u_n - \epsilon a_n)^2 + 2n^2 \mathbb{P}(|\tilde{V}_1 - V_1^{\lfloor n^\delta \rfloor}| > \epsilon a_n). \end{aligned} \quad (\text{B.11})$$

Analogously to (B.4) and (B.5), with  $\beta_n = 3/(\epsilon(1 - \epsilon)) \log n$ , we have  $n^2 \mathbb{P}(|\tilde{V}_1 - V_1^{\lfloor n^\delta \rfloor}| > \epsilon a_n) \rightarrow 0$  as  $n \rightarrow \infty$ . Using (B.1), we also have

$$\lim_{n \rightarrow \infty} (n^2/k) \mathbb{P}(\tilde{V}_1 > u_n - \epsilon a_n)^2 = \exp(-2(x - \epsilon))/k, \quad (\text{B.12})$$

which converges to 0 as  $k \rightarrow \infty$ . Then by (B.6), (B.10)-(B.12), and letting  $\epsilon \downarrow 0$ , the  $D'(u_n)$  condition holds.

(b) To prove condition  $\mathbf{D}_r(\mathbf{u}_n)$ , we replace  $V_k^n$  in (a) by

$$M_k^n := \sup_{(k-1)h \leq t \leq kh} \int_{t-nh}^t e^{-\lambda(t-s)} dL_{\lambda s}.$$

We then obtain an analogue result to (B.2). Further, since

$$\begin{aligned} |M_k - M_k^n| &\leq \sup_{(k-1)h \leq t \leq kh} \int_{-\infty}^{t-nh} e^{-\lambda(t-s)} dL_{\lambda s} \\ &\stackrel{d}{=} e^{-\lambda nh} \sup_{(k-1)h \leq t \leq kh} \int_{-\infty}^t e^{-\lambda(t-s)} dL_{\lambda s} = e^{-\lambda nh} M_k, \end{aligned}$$

we obtain for any  $\delta > 0$  that

$$n\mathbb{P}(|M_k - M_k^{\lfloor n\delta \rfloor}| > \epsilon a_n) \leq n\mathbb{P}(e^{-\lambda \lfloor n\delta \rfloor h} M_k > \epsilon a_n).$$

Since the first exponential moment of  $\beta\gamma e^{-\lambda \lfloor n\delta \rfloor h} M_k$  exists for  $\beta e^{-\lambda \lfloor n\delta \rfloor h} < 1$ , similar reasoning as in (B.4) and (B.5) shows that  $\lim_{n \rightarrow \infty} n\mathbb{P}(|M_k - M_k^{\lfloor n\delta \rfloor}| > \epsilon a_n) = 0$ . As in the proof of (a) we then conclude that the  $\mathbf{D}_r(\mathbf{u}_n)$  condition holds.

For the proof of the  $D'(u_n)$  condition we use

$$\begin{aligned} M_k &\leq \int_{-\infty}^{\infty} \sup_{(k-1)h \leq t \leq kh} e^{-\lambda(t-s)} \mathbf{1}_{(-\infty, t]}(s) dL_{\lambda s} \\ &= L_{\lambda kh} - L_{\lambda(k-1)h} + \int_{-\infty}^{(k-1)h} e^{-\lambda((k-1)h-s)} dL_{\lambda s} =: \bar{V}_k. \end{aligned}$$

Let  $j \geq 3$ . Then we have the upper bound

$$\begin{aligned} \bar{V}_j &= L_{\lambda jh} - L_{\lambda(j-1)h} + e^{-\lambda(j-1)h} \int_{-\infty}^0 e^{\lambda s} dL_{\lambda s} \\ &\quad + e^{-\lambda(j-2)h} \int_0^h e^{-\lambda(h-s)} dL_{\lambda s} + \int_h^{(j-1)h} e^{-\lambda((j-1)h-s)} dL_{\lambda s} \\ &\leq L_{\lambda jh} - L_{\lambda(j-1)h} + e^{-\lambda(j-2)h} \int_{-\infty}^0 e^{\lambda s} dL_{\lambda s} \\ &\quad + e^{-\lambda(j-2)h} L_{\lambda h} + \int_h^{(j-1)h} e^{-\lambda((j-1)h-s)} dL_{\lambda s} \\ &\leq L_{\lambda jh} - L_{\lambda(j-1)h} + e^{-\lambda h} \bar{V}_1 + \int_h^{(j-1)h} e^{-\lambda((j-1)h-s)} dL_{\lambda s}. \end{aligned}$$

Let  $\bar{V}'_1$  be an independent copy of  $\bar{V}_1$ . Then

$$\begin{aligned} n \sum_{j=3}^{\lfloor n\delta \rfloor} \mathbb{P}(M_1 > u_n, M_j > u_n) &\leq n \sum_{j=3}^{\lfloor n\delta \rfloor} \mathbb{P}(\bar{V}_1 > u_n, \bar{V}_j > u_n) \\ &\leq n \sum_{j=3}^{\lfloor n\delta \rfloor} \mathbb{P}(\bar{V}_1 + \bar{V}_j > 2u_n) \\ &\leq n^{1+\delta} \mathbb{P}((1 + e^{-\lambda h})\bar{V}_1 + \bar{V}'_1 > 2u_n). \end{aligned}$$



The tail of  $\bar{V}_1$  behaves by Proposition 3.9 (b) and Theorem A.4 (ii) like

$$\mathbb{P}(\bar{V}_1 > x) = \mathbb{P}(L_{\lambda h} + V_0 > x) \sim \mathbb{E}e^{\gamma V_0} \mathbb{P}(L_{\lambda h} > x), \quad x \rightarrow \infty,$$

so that  $\bar{V}_1 \in \mathcal{S}(\gamma)$ . An analogue result to (B.8) and (B.9) gives

$$\lim_{n \rightarrow \infty} n^{1+\delta} \mathbb{P}((1 + e^{-\lambda h})\bar{V}_1 + \bar{V}'_1 > 2u_n) = 0,$$

and arguing similarly as in (B.11) and (B.12), we obtain

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} n \sum_{j=3}^{\lfloor n/k \rfloor} \mathbb{P}(M_1 > u_n, M_j > u_n) = 0.$$

It remains to show that

$$\lim_{n \rightarrow \infty} n \mathbb{P}(M_1 > u_n, M_2 > u_n) = 0. \quad (\text{B.13})$$

Note that

$$\begin{aligned} \mathbb{P}(M_1 > u_n) &= \mathbb{P}(M_1 > u_n, N_{\lambda h} > 0) + \mathbb{P}(M_1 > u_n, N_{\lambda h} = 0) \\ &\geq \mathbb{P}(V_{\Gamma_1/\lambda} > u_n, N_{\lambda h} > 0) \\ &\geq \mathbb{P}(\xi_1 > u_n) \mathbb{P}(N_{\lambda h} > 0), \end{aligned}$$

so that

$$n \mathbb{P}(M_1 > u_n, M_2 > u_n) \leq \frac{n \mathbb{P}(M_1 > u_n)}{\mathbb{P}(N_{\lambda h} > 0)} \frac{\mathbb{P}(M_1 > u_n, M_2 > u_n)}{\mathbb{P}(\xi_1 > u_n)}. \quad (\text{B.14})$$

Furthermore, we have the upper bound

$$\begin{aligned} \mathbb{P}(M_1 > u_n, M_2 > u_n) &\leq \mathbb{P}(\bar{V}_1 + \bar{V}_2 > 2u_n) \\ &\leq \mathbb{P}\left(\frac{L_{2\lambda h} - L_{\lambda h}}{2} + \frac{1 + e^{-\lambda h}}{2} \int_{-\infty}^0 e^{\lambda s} dL_{\lambda s} + \frac{1}{2} \int_0^h (1 + e^{-\lambda h + \lambda s}) dL_{\lambda s} > u_n\right). \end{aligned}$$

The three summands are independent, and we shall show that for each of them the probability to be greater than  $u_n$  is of order  $o(\mathbb{P}(\xi_1 > u_n))$  for  $n \rightarrow \infty$ , so that by Theorem A.4 (ii)

$$\lim_{n \rightarrow \infty} \mathbb{P}(M_1 > u_n, M_2 > u_n) / \mathbb{P}(\xi_1 > u_n) = 0. \quad (\text{B.15})$$

Equation (B.13) and hence condition  $D'(u_n)$  then follow from (B.2), (B.14) and (B.15).

The rapidly varying tails and Theorem A.4 (i) give

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}(L_{2\lambda h} - L_{\lambda h} > 2x)}{\mathbb{P}(\xi_1 > x)} = \lim_{x \rightarrow \infty} \mu \frac{\mathbb{P}(L_{\lambda h} > x)}{\nu(x, \infty)} \frac{\mathbb{P}(L_{\lambda h} > 2x)}{\mathbb{P}(L_{\lambda h} > x)} = 0,$$

which is the assertion for the first summand. Further, also by the rapidly varying tails, Proposition 3.9 (b) and Theorem A.4 (i),

$$\frac{\mathbb{P}\left(\left(1 + e^{-\lambda h}\right) \int_{-\infty}^0 e^{\lambda s} dL_{\lambda s} > 2x\right)}{\mathbb{P}(\xi_1 > x)} = \frac{\mathbb{P}\left(\left(1 + e^{-\lambda h}\right)V_0 > 2x\right)}{\mathbb{P}(V_0 > x)} \frac{\mu \mathbb{P}(V_0 > x)}{\nu(x, \infty)} \xrightarrow{x \rightarrow \infty} 0.$$

For the last summand we use that

$$X := \int_0^h (1 + e^{-\lambda h} e^{\lambda s}) dL_{\lambda s} \stackrel{d}{=} \sum_{i=1}^{N_{\lambda h}} (1 + e^{-\lambda h} e^{\lambda h U_i}) \xi_i \stackrel{d}{=} \sum_{i=1}^{N_{\lambda h}} (1 + e^{-\lambda h U_i}) \xi_i,$$

where  $(U_i)_{i \in \mathbb{N}}$ ,  $U$  are i. i. d. uniform on  $(0, 1)$  and independent of  $L$  (cf. e. g. Sato [35], Proposition 3.4).

From Theorem A.4 (iii) then follows

$$\frac{\mathbb{P}\left(\int_0^h (1 + e^{-\lambda s}) dL_{\lambda s} > 2x\right)}{\mathbb{P}(\xi_1 > x)} \sim \mu \lambda h \mathbb{E} e^{\gamma X} \frac{\mathbb{P}(\xi_1 (1 + e^{-\lambda h U})/2 > x)}{\mathbb{P}(\xi_1 > x)} \xrightarrow{x \rightarrow \infty} 0. \quad \square$$

## C Stationary solution of a random recurrence equation

The following result is the central result for proving stationarity and the tail behavior of a stochastic process defined by a random recurrence equation; it goes back to seminal work by Kesten [24] and Vervaat [36].

**Theorem C.1 (Goldie [21], Theorem 4.1, Lemma 2.2).**

Let  $(Y_k)_{k \in \mathbb{N}}$  be a stochastic process defined by  $Y_k = A_k Y_{k-1} + B_k$ , where  $((A_k, B_k))_{k \in \mathbb{N}}$ ,  $(A, B)$  are i. i. d. sequences. Assume that the following conditions are satisfied for some  $\alpha > 0$ :

- (i) The law of  $\log |A|$ , given  $|A| \neq 0$ , is not concentrated on a lattice  $-\infty \cap r\mathbb{Z}$  for any  $r > 0$ .
- (ii)  $\mathbb{E}|A|^\alpha = 1$ .
- (iii)  $\mathbb{E}|A|^\alpha \log^+ |A| < \infty$ .

(iv)  $\mathbb{E}|B|^\alpha < \infty$ .

Then the equation  $Y_\infty \stackrel{d}{=} AY_\infty + B$ , where  $Y_\infty$  is independent of  $(A, B)$ , has the solution unique in distribution

$$Y_\infty \stackrel{d}{=} \sum_{m=1}^{\infty} B_m \prod_{k=1}^m A_k.$$

The process  $(Y_k)_{k \in \mathbb{N}}$  with  $Y_0 \stackrel{d}{=} Y_\infty$  is stationary and has tails

$$\mathbb{P}(Y_\infty > x) \sim \frac{\mathbb{E} [((AY_\infty + B)^+)^{\alpha} - ((AY_\infty)^+)^{\alpha}]}{\alpha \mathbb{E}|A|^\alpha \log^+ |A|} x^{-\alpha}, \quad x \rightarrow \infty.$$

## References

- [1] ALBIN, J. M. P. On extremes of infinitely divisible Ornstein-Uhlenbeck processes. Preprint, available at <http://www.math.chalmers.se/~palbin/>.
- [2] BARNDORFF-NIELSEN, O. E. (1998). Processes of normal inverse Gaussian type. *Finance Stoch.* **2**, 41–68.
- [3] BARNDORFF-NIELSEN, O. E. AND SHEPHARD, N. (2001). Modelling by Lévy processes for financial econometrics. In: O. E. Barndorff-Nielsen, T. Mikosch, and S. I. Resnick (Eds.), *Lévy Processes: Theory and Applications*, pp. 283–318. Birkhäuser, Boston.
- [4] BARNDORFF-NIELSEN, O. E. AND SHEPHARD, N. (2001). Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics (with discussion). *J. Roy. Statist. Soc. Ser. B* **63** (2), 167–241.
- [5] BARNDORFF-NIELSEN, O. E. AND SHEPHARD, N. (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *J. Roy. Statist. Soc. Ser. B* **64**, 253–280.
- [6] BERTOIN, J. (1996). *Lévy Processes*. Cambridge University Press, Cambridge.
- [7] BINGHAM, N. H., GOLDIE, C. M., AND TEUGELS, J. L. (1987). *Regular Variation*. Cambridge University Press, Cambridge.
- [8] BORKOVEC, M. AND KLÜPPELBERG, C. (1998). Extremal behavior of diffusions models in finance. *Extremes* **1** (1), 47–80.

- [9] BRAVERMAN, M. AND SAMORODNITSKY, G. (1995). Functionals of infinitely divisible stochastic processes with exponential tails. *Stochastic Process. Appl.* **56** (2), 207–231.
- [10] BREIMAN, L. (1965). On some limit theorems similar to the arc-sine law. *Theory Probab. Appl.* **10**, 323–331.
- [11] BROCKWELL, P. J., CHADRAA, E., AND LINDNER, A. M. (2004). Continuous time GARCH processes of higher order. Preprint, available at <http://www.ma.tum.de/stat/>.
- [12] BUCHMANN, B. AND KLÜPPELBERG, C. (2004). Fractional integral equation and state space transforms. Preprint, available at <http://www.ma.tum.de/stat/>.
- [13] BUCHMANN, B. AND KLÜPPELBERG, C. (2004). Maxima of stochastic processes driven by fractional Brownian motion. Preprint, available at <http://www.ma.tum.de/stat/>.
- [14] CLINE, D. B. H. (1987). Convolution of distributions with exponential and subexponential tails. *J. Austral. Math. Soc. Ser. A* **43** (3), 347–365.
- [15] CONT, R. AND TANKOV, P. (2004). *Financial Modelling with Jump Processes*. Chapman & Hall, Boca Raton.
- [16] DROST, F. AND WERKER, B. J. M. (1996). Closing the GARCH gap: continuous time GARCH modelling. *J. Econometrics* **74**, 31–57.
- [17] EMBRECHTS, P., KLÜPPELBERG, C., AND MIKOSCH, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.
- [18] FASEN, V. (2004). *Extremes of Lévy Driven MA Processes with Applications in Finance*. Ph.D. thesis, Munich University of Technology.
- [19] FASEN, V. (2005). Extremes of regularly varying mixed moving average processes. Preprint, available at <http://www.ma.tum.de/stat/>.
- [20] FASEN, V. (2005). Extremes of subexponential Lévy driven moving average processes. Preprint, available at <http://www.ma.tum.de/stat/>.
- [21] GOLDIE, C. M. (1991). Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* **1** (1), 126–166.

- [22] HAAN, L. D., RESNICK, S. I., ROOTZÉN, H., AND VRIES, C. G. (1989). Extremal behavior of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Process. Appl.* **32**, 213–224.
- [23] HSING, T. AND TEUGELS, J. L. (1989). Extremal properties of shot noise processes. *Adv. Appl. Probability* **21**, 513–525.
- [24] KESTEN, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131**, 207–248.
- [25] KLÜPPELBERG, C. (1989). Subexponential distributions and characterizations of related classes. *Probab. Theory Relat. Fields* **82**, 259–269.
- [26] KLÜPPELBERG, C. (2004). Risk management with extreme value theory. In: B. Finkenstädt and H. Rootzén (Eds.), *Extreme Values in Finance, Telecommunication and the Environment*, pp. 101–168. Chapman & Hall/CRC, Boca Raton.
- [27] KLÜPPELBERG, C., LINDNER, A., AND MALLER, R. (2004). A continuous time GARCH process driven by a Lévy process: stationarity and second order behaviour. *J. Appl. Probab.* **41** (3), 601–622.
- [28] KLÜPPELBERG, C., LINDNER, A., AND MALLER, R. (2004). Continuous time volatility modelling: COGARCH versus Ornstein-Uhlenbeck models. In: *Invited paper for the Bachelier Colloquium 2005*.
- [29] LEADBETTER, M. R., LINDGREN, G., AND ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York.
- [30] LINDNER, A. AND MALLER, R. (2004). Lévy integrals and the stationarity of generalised Ornstein-Uhlenbeck processes. Preprint, available at <http://www.ma.tum.de/stat/>.
- [31] MIKOSCH, T. AND STĂRICĂ, C. (2000). Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. *Ann. Statist.* **28**, 1427–1451.
- [32] PAKES, A. G. (2004). Convolution equivalence and infinite divisibility. *J. Appl. Probab.* **41** (2), 407–424.

- [33] ROOTZÉN, H. (1986). Extreme value theory for moving average processes. *Ann. Probab.* **14** (2), 612–652.
- [34] ROSINSKI, J. AND SAMORODNITSKY, G. (1993). Distributions of subadditive functionals of sample paths of infinitely divisible processes. *Ann. Probab.* **21** (2), 996–1014.
- [35] SATO, K. I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge.
- [36] VERVAAT, W. (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. Appl. Probability* **11**, 750–783.