Technische Universität München

Zentrum Mathematik

A Fractionally Integrated COGARCH(1,1) Model

Masterarbeit

von

German Straub

Themenstellerin:	Prof. Dr. Klüppelberg
Betreuer:	Dr. Stephan Haug
Abgabetermin:	26. Juni 2014

Hiermit erkläre ich, dass ich die Masterarbeit selbstständig angefertigt und nur die angegebenen Quellen verwendet habe.

Garching, den 25. Juni 2014

Acknowledgments

I would like to thank Prof. Claudia Klüppelberg and Dr. Stephan Haug for their helpful advice, support and supervision throughout the work on this thesis.
Special thanks go to my parents, who enabled me to study.

Abstract

When modelling financial time series, the main difficulty consists in finding a model that captures the so-called stylized facts. These are statistical regularities, such as leptocurticity, volatility clustering or strong autocorrelations for absolute and squared returns, which are common to most financial series. The most popular way to take such characteristics into account is formed by models of generalized autoregressive conditional heteroscedasticity (GARCH). These models, however, cannot explain the empirically often observed strong dependence in volatility. In 1996, Baillie, Bollerslev and Mikkelsen therefore introduced the fractionally integrated GARCH (FIGARCH) model. While the existence of a strictly stationary solution is ensured, its ability to model long range dependence in volatility is controversial.

In the literature there exist serveral approaches to define a continuous-time analogue to the discrete GARCH process. The continuous-time GARCH (COGARCH) model of Klüppelberg, Lindner and Maller stands out as it directly generalizes the essential features of its discrete time analogue. In this thesis we present two approaches to incorporate long range dependence into the volatility process of the COGARCH(1,1). The first one is based on Molchan-Golosov fractional Lévy processes (FLP), while for the second we use a modification of the Mandelbrot-van-Ness FLP.

Contents

Notation

Elements of $\mathbb{R}^m$ are marked by bold letters. The inner product of $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^m$ is denoted by $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ and $|\boldsymbol{x}| := \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$ stands for the Euclidean norm. Further, $\mathcal{B}(\mathbb{R}^m)$ is the Borel $\sigma-$algebra generated by the open sets of $\mathbb{R}^m$, while $M_m(\mathbb{R})$ denotes the set of $m \times m-$matrices with entries in $\mathbb{R}$. We write $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, while $\mathbb{R}_0 := \mathbb{R} \setminus \{0\}$. Besides, let $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ be two sequences, then by $a_n \sim b_n$ we mean that $\frac{a_n}{b_n} \longrightarrow 1$ for $n \longrightarrow \infty$.

A filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \geq 0})$ is said to be *complete* if the σ-algebra $\mathcal{F}_0$ contains all (P-)null sets. If the filtration $(\mathcal{F}_t)_{t \geq 0}$ additionally is right-continuous, i.e. for all $t \geq 0$ we have $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$, then it fulfills the *usual conditions*.
Let the index set T be either given by $T = \mathbb{R}$ or $T = \mathbb{Z}$. Then a stochastic process $X = (X_t)_{t \in T}$ is said to be *strictly stationary*, if for $t_1, \ldots, t_n, h \in T$ the random vectors $(X_{t_1}, \ldots, X_{t_n})$ and $(X_{t_1+h}, \ldots, X_{t_n+h})$ have the same joint distributions. Moreover, X is called *weakly stationary*, if its first two moments exist and are constant and its autocovariance function (ACVF) $\gamma_X(t_i, t_j) := \mathrm{Cov}(X_{t_i}, X_{t_j})$ depends only on the distance $h := t_i - t_j$. We then write $\gamma_X(h) := \gamma(t_i, t_j)$. A weakly stationary process X such that $\mathrm{Cov}(X_{t_i}, X_{t_j}) = 0, i \neq j$, is referred to as *white noise.*
Finally, a stochastic process $X = (X_t)_{t \geq 0}$ has *càdlàg* paths, if the sample paths $t \mapsto X_t(\omega)$ are almost surely (a.s.) right-continuous with left limits. Jumps are denoted by $\Delta X_t := X_t - X_{t-}$ where $X_{t-} := \lim_{s \uparrow t} X_s$.

Chapter 1

Introduction

From the econometric point of view, modeling time series with financial background, such as asset returns, is especially challenging. The difficulty basically consists in finding a model, which captures the so-called *stylized facts*, i.e. statistical regularities shared by many financial series. These facts include leptokurticity (i.e heavy-tailed marginal distributions), volatility clustering (i.e. periods of high and low volatility are observable) and strong autocorrelation for the squared and absolute values, while the series itself displays only small autocorrelation. Clearly, linear time series models, as for example the class of *autoregressive moving-average* (ARMA) models, are insufficient in this respect, as they are based on the second-order structure of the underlying process.
The last two of the above mentioned stylized facts suggest models, which allow for a non-constant conditional variance. This concept is usually referred to as *conditional heteroscedasticity* and led to the development of volatility models, where the series of observations X_n is modeled as

$$X_n = \sigma_n \varepsilon_n, \qquad n \in \mathbb{Z}, \tag{1.1}$$

where $(\varepsilon_n)_{n\in\mathbb{Z}}$ denotes a sequence of i.i.d. random variables with zero mean and unit variance, and $(\sigma_n)_{n\in\mathbb{Z}}$ is a deterministic function of $X_m, \varepsilon_m, m \leq n-1$. This class of models should not be confused with discrete *stochastic volatility models*, where the sequence $(\sigma_n)_{n\in\mathbb{Z}}$ usually is assumed to be independent of the noise $(\varepsilon_n)_{n\in\mathbb{Z}}$, see Shephard [2008].
The random variable σ_n in (1.1) is referred to as *volatility*, but in fact its square represents the conditional variance of X_n, i.e.

$$\mathrm{Var}(X_n | X_m, m \leq n-1) = \sigma_n^2, \qquad n \in \mathbb{Z}. \tag{1.2}$$

The basis for the development of such volatility models was laid by Engle [1982] with introducing the class of *autoregressive conditionally heteroscedastic* (ARCH) models, which represents the most popular and most widely used class of nonlinear time series models. It is characterized by the conditional variance σ_n^2 in (1.1) being parametrized as a linear function of finitely many $X_m^2, m \leq n-1$. In applications it turned out that lags of high orders of X_m^2 needed to be included into the parametrization of σ_n^2 to sufficiently model the observed dependence structure. This inconvenience was removed when Bollerslev [1986] introduced generalized ARCH (GARCH) models, where the squared volatility σ_n^2 is given as linear function of both its own past $\sigma_m^2, m \leq n-1$ and $X_m^2, m \leq n-1$.
An important characteristic of GARCH models is given by its short memory behavior.

That means, the autocorrelation of its squares decays at a fast, namely exponential rate. As a result, GARCH models are not able to capture the empirically often observed persistence in volatility (cf. Baillie [1996]).
In the case of linear time series models, long-memory (understood in the sense of a non-summable ACVF) was successfully incorporated by the introduction of *fractionally integrated* ARMA, shortly ARFIMA models, see Granger and Joyeux [1980] and Hosking [1981]. However, turning to the ARCH framework things get more involved. More precisely, as we summarize below, for the broad class of ARCH(∞) processes the conditions for weak stationarity directly rule out the possibility for long memory. However, dropping the requirement of a finite second moment, Baillie et al. [1996] defined a fractionally integrated GARCH (FIGARCH) model, which was meant to generalize the GARCH model in the same way as it did the ARFIMA in the ARMA case. The existence of a strictly stationary solution was proved by Douc et al. [2008] for the FIGARCH$(0, d, 0)$ model. We show that their idea can be generalized by proving the existence of the FIGARCH$(1, d, 0)$ and FIGARCH$(0, d, 1)$.

Especially in the last decades interest in continuous time models has increased enormously. This development can be attributed on the one hand to the very successful application of continuous time models in finance, starting with the introduction of the famous Black-Scholes-Merton model. On the other hand, it is the result of the widespread availability of high-frequency data, which is characterized by irregular spacing in time. To avoid a loss of information, which the application of a discrete time model to high-frequency data would result in, continuous time parameter models can be utilized.

Regarding GARCH there exist several approaches to approximate the discrete time model by a continuous time one, see for example Lindner [2009]. Yet, in many cases these continuous time models lack essential features of the discrete GARCH. An example would be the approach of diffusion approximation taken by Nelson [1990]. This basically consists in embedding the discrete GARCH$(1, 1)$ into a continuous time process by piecewise constant interpolation between grid points and letting the mesh converge to zero. However, quite counterintuitively, the limiting process $(G_t)_{t\geq 0}$ is given by a diffusion equation of the following type

$$\begin{aligned} dG_t &= \sigma_t \, dW_t^{(1)}, \\ d\sigma_t^2 &= (\omega - \theta\sigma_t^2)\, dt + \lambda\sigma_t^2 \, dW_t^{(2)}, \qquad t \geq 0, \end{aligned}$$

where $W^{(1)}, W^{(2)}$ denote two independent Brownian motions and ω, θ and λ are constants. Consequently, the limiting process has lost an important feature of the discrete GARCH process, namely the *feedback mechanism*. This is based on the fact that only one driving source of randomness exists, such that an innovation in the observed process directly leads to an innovation in the volatility process.

In contrast, Klüppelberg et al. [2004] introduced a continuous time analogue to the discrete GARCH$(1, 1)$, called COGARCH$(1, 1)$, which is driven by a single Lévy process, such that the described feedback mechanism is preserved. Moreover it directly generalizes many essential features of its discrete analogue, such as volatility clustering, heavy tails and uncorrelated increments. Brockwell et al. [2006] generalized the COGARCH$(1, 1)$ by

introducing the COGARCH(p, q) model.
The analogy between discrete GARCH and COGARCH also applies to the memory properties, meaning that the squared volatility of the COGARCH exhibits short memory, too. This thesis deals with the resultant question of whether the COGARCH model can be modified (probably in analogy to the FIGARCH) to allow for long range dependence.

Chapter 2 on the one hand provides the *mathematical tools* we need to define a *fractionally integrated* COGARCH model. More precisely, after a review of Lévy processes and integration with respect to Lévy processes we introduce fractional Lévy processes (FLP), which are generalizations of fractional Brownian motions. On the other hand, we recall discrete- and continuous-time ARMA, shortly CARMA, processes. In doing so we are especially interested in the way long range dependence is incorporated into these models. This leads us to fractionally integrated ARMA (ARFIMA) and their continuous-time counterpart, namely FICARMA processes.
In Chapter 3 we firstly give a detailed review of ARCH(∞) processes, introducing FIGARCH models as a particular subclass. Subsequently we turn to the continuous-time setting and introduce COGARCH processes. Finally, we propose two fractionally integrated COGARCH models and give the corresponding simulation results.

Chapter 2

Preliminaries

2.1 Lévy processes

This section is devoted to the introduction of Lévy processes as well as to the presentation of two important results, the first one concerning the representation of Lévy processes, namely the so called *Lévy-Itô decomposition*, the second, the *Lévy-Khintchine formula*, giving a formula for the Fourier transform. Further we intend to give some insight into the so-called *Lévy measure*. The presented results can be found in Protter [2004], Applebaum [2004] and Sato [1999].

We start with the definition of Lévy processes.

Definition 2.1 (Lévy processes) *A (m-dimensional) stochastic process $(\boldsymbol{L}_t)_{t\geq 0}$ is called Lévy process if it is adapted and fulfills the following conditions.*

(i) It starts in zero, i.e. $L_0 = 0$ a.s.,

(ii) it has independent increments, i.e. for all $0 \leq s < t$ the increment $\boldsymbol{L}_t - \boldsymbol{L}_s$ is independent of $\mathcal{F}_s$,

(iii) its increments are stationary, i.e. for all $0 \leq s < t$ we have $\boldsymbol{L}_t - \boldsymbol{L}_s \stackrel{\mathscr{D}}{=} \boldsymbol{L}_{t-s}$ and

(iv) and, finally, it is stochastically continuous, i.e. for all $t > 0$ is holds that $|\boldsymbol{L}_t - \boldsymbol{L}_s| \stackrel{P}{\to} 0$ for $s \to t$.

Many well-known stochastic processes turn out to be Lévy processes. Examples are the Brownian motion, Poisson processes or compound Poisson processes. Before continuing, note that for any Lévy process there exists a càdlàg modification, meaning, we can find a stochastic process $(\tilde{\boldsymbol{L}}_t)_{t\geq 0}$ with has right-continuous paths and left limits and satisfies

$$P(\boldsymbol{L}_t = \tilde{\boldsymbol{L}}_t) = 1, \qquad \text{for all } t \geq 0. \tag{2.1}$$

The union of finitely many null sets is again a null set, consequently (2.1) implies that $(\boldsymbol{L}_t)_{t\geq 0}$ and $(\tilde{\boldsymbol{L}}_t)_{t\geq 0}$ have the same finite dimensional distributions. In this sense we lose nothing when considering instead of $(\boldsymbol{L}_t)_{t\geq 0}$ its càdlàg modification. Hence, in the following we will assume that the Lévy process under consideration has right-continuous paths with left limits.

A key result in the theory of Lévy processes is that their sample paths decompose into continuous parts and jump parts. For the representation of the jump parts we will make use of a particular Poisson random measure (see Definition A.2) associated to the Lévy process.

Consider a Lévy process $(\boldsymbol{L}_t)_{t\geq 0}$. For $B \in \mathcal{B}(\mathbb{R}_0^m)$ we set

$$\nu_L(B) := E\left[\sum_{s\leq 1} 1_{\{\Delta \boldsymbol{L}_s \in B\}}\right]. \tag{2.2}$$

Recalling that $\boldsymbol{L}$ has independent increments and càdlàg sample paths, we conclude that ν_L defines a σ-finite measure on $\mathcal{B}(\mathbb{R}_0^m)$. Obviously, for $B \in \mathcal{B}(\mathbb{R}_0^m)$ ν_L basically counts the expected number of jumps of $\boldsymbol{L}$ in a unit interval with height lying in B. A well known result states that

$$\int_{\mathbb{R}_0^m} \min(|\boldsymbol{x}|^2, 1)\, \nu_L(d\boldsymbol{x}) < \infty.$$

We now get the following result.

Proposition 2.2 (Jump measure) *Let $(\boldsymbol{L}_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}^m$ and set $\Sigma := (0,\infty) \times \mathbb{R}_0^m$. Furthermore, we define for $S \in \mathcal{B}(\Sigma)$*

$$\mu_L(S,\omega) := \#\left\{t > 0 : (t, \Delta \boldsymbol{L}_t(\omega)) \in S)\right\}.$$

Then the collection $\{\mu_L(S) : S \in \mathcal{B}(\Sigma)\}$ is a Poisson random measure on Σ with intensity measure given by $\lambda \otimes \nu_L$, where λ denotes the Lebesgue measure.

Proof. See Theorem 19.2 of Sato [1999]. □

Definition 2.3 *The measures μ_L and ν_L from Proposition 2.2 are usually called* jump- *and* Lévy measure *respectively.*

We are now ready to state the above mentioned decomposition result.

Theorem 2.4 (Lévy-Itô-Decomposition) *Let $(\boldsymbol{L}_t)_{t\geq 0}$ be a Lévy process. Then there exists a nonnegative-definite and symmetric matrix $A_L \in M_m(\mathbb{R})$ such that $\boldsymbol{L}$ decomposes into*

$$\boldsymbol{L}_t = t\boldsymbol{\gamma}_L + \boldsymbol{B}_t + \int_{[0,t]\times\{|\boldsymbol{x}|>1\}} \boldsymbol{x}\mu_L(ds, d\boldsymbol{x}) + \int_{[0,t]\times\{|\boldsymbol{x}|\leq 1\}} \boldsymbol{x}\Big(\mu_L(ds, d\boldsymbol{x}) - ds\, \nu_L(d\boldsymbol{x})\Big), \qquad t \geq 0, \tag{2.3}$$

where the constant $\boldsymbol{\gamma}_L \in \mathbb{R}^m$ satisfies

$$\boldsymbol{\gamma}_L = E[\boldsymbol{L}_1] - \int_{|\boldsymbol{x}|>1} \boldsymbol{x}\, \nu_L(d\boldsymbol{x}), \tag{2.4}$$

$(\boldsymbol{B}_t)_{t\geq 0}$ denotes an $m-$dimensional Brownian motion with covariance matrix A_L. Furthermore, the four summands are independent.

Proof. See [Applebaum, 2004, Theorem 2.4.16]. □

Remark 2.5 *Obviously, the first two summands in* (2.3) *represent the continuous whereas the last two represent the jump part of the Lévy process. Furthermore, it is noticeable, that when describing the jump part one has to distinguish between "large" and "small" jumps, as the latter ones are in general not summable. The idea is to compensate them - which basically means substracting the mean. As a consequence the compensated sum of jumps*

$$\int_{[0,t]\times\{|\boldsymbol{x}|\leq 1\}} \boldsymbol{x}\Big(\mu_L(ds,d\boldsymbol{x}) - ds\ \nu_L(d\boldsymbol{x})\Big) := \lim_{\varepsilon\downarrow 0}\int_{[0,t]\times\{\varepsilon<|\boldsymbol{x}|\leq 1\}} \boldsymbol{x}\Big(\mu_L(ds,d\boldsymbol{x}) - ds\ \nu_L(d\boldsymbol{x})\Big)$$

exists as limit in $L^2(\Omega)$. *In particular, in the case where the Lévy measure satisfies*

$$\int_{|\boldsymbol{x}|\leq 1} \boldsymbol{x}\ \nu_L(d\boldsymbol{x}) < \infty,$$

i.e. where the "small" jumps are summable, the decomposition (2.3) *simplifies to*

$$\boldsymbol{L}_t = t\boldsymbol{\gamma} + B_t + \int_{[0,t]\times\mathbb{R}_0^m} \boldsymbol{x}\mu_L(ds,d\boldsymbol{x}) \qquad t\geq 0, \tag{2.5}$$

with constant $\boldsymbol{\gamma}$ *given by*

$$\boldsymbol{\gamma} = E[\boldsymbol{L}_1] - \int_{\mathbb{R}_0^m} \boldsymbol{x}\ \nu_L(d\boldsymbol{x}).$$

An important by-product of the Lévy-Itô decomposition is the so-called *Lévy-Khintchine* formula. It is basically a result concerning the Fourier transform of a Lévy process.

Theorem 2.6 (Lévy-Khintchine formula) *Let* $(\boldsymbol{L}_t)_{t\geq 0}$ *be a Lévy process with Lévy measure* ν_L. *Then there exists a nonnegative-definite symmetric matrix* $A_L \in M_m(\mathbb{R})$ *such that for all* $t\geq 0$ *and* $\boldsymbol{u}\in\mathbb{R}^m$

$$E\Big[e^{i\langle \boldsymbol{u},\boldsymbol{L}_t\rangle}\Big] = e^{t\psi(\boldsymbol{u})}.$$

The function $\psi:\mathbb{R}^m\to\mathbb{R}$ *is called Lévy symbol and satisfies*

$$\psi(\boldsymbol{u}) = i\langle\boldsymbol{\gamma}_L,\boldsymbol{u}\rangle - \frac{1}{2}\langle\boldsymbol{u},A_L\boldsymbol{u}\rangle + \int_{\mathbb{R}_0^m}\Big[e^{i\langle\boldsymbol{u},\boldsymbol{x}\rangle} - 1 - i\langle\boldsymbol{u},\boldsymbol{x}\rangle 1_{\{|\boldsymbol{x}|\leq 1\}}\Big]\nu_L(d\boldsymbol{x}), \tag{2.6}$$

where $\boldsymbol{\gamma}_L$ *is given by* (2.4). *Moreover, given* $\boldsymbol{\gamma}_L, A_L$ *and* ν_L *the corresponding Lévy process is unique in distribution.*

Proof. See [Applebaum, 2004, Corollary 2.4.20]. □

Remark 2.7 *1. Observe that the constant $\boldsymbol{\gamma}_L$ actually depends on the chosen representation. This means, instead of $\boldsymbol{x} \mapsto 1_{\{|\boldsymbol{x}|\leq 1\}}$ other so-called cut-off functions denoted by $\boldsymbol{x} \mapsto c(\boldsymbol{x})$ may be used in* (2.6). *Possible choices can be found in [Sato, 1999, Remark 8.4]. In fact, the crucial thing is that*

$$\boldsymbol{x} \mapsto e^{i\langle \boldsymbol{u},\boldsymbol{x}\rangle} - 1 - i\langle \boldsymbol{u}, \boldsymbol{x}\rangle c(\boldsymbol{x})$$

is ν_L-integrable.

2. *We also would like to remark, that the above Lévy-Khintchine formula establishes a link between Lévy processes and* infinite divisible distributions[1], *see [Applebaum, 2004, Theorem 1.2.14].*

Usually, $\boldsymbol{\gamma}_L$ and A_L are referred to as *drift* and *diffusion* coefficient, respectively. The triplet $(\boldsymbol{\gamma}_L, A_L, \nu_L)$ is called *characteristic triplet* of $(\boldsymbol{L}_t)_{t\geq 0}$. Many properties of a Lévy process a determined by this triplet. We present some of them.

Proposition 2.8 *Let $(L_t)_{t\geq 0}$ be a Lévy process with characteristic triplet $(\boldsymbol{\gamma}_L, A_L, \nu_L)$.*

(i) The sample paths $t \mapsto \boldsymbol{L}_t(\omega)$ are a.s. continuous if and only if $\nu_L = 0$.

(ii) $(L_t)_{t\geq 0}$ is a.s. of finite variation if and only if $A_L = 0$ and $\int_{|\boldsymbol{x}|\leq 1} |\boldsymbol{x}|\ \nu_L(d\boldsymbol{x}) < \infty$.

(iii) Let $p \geq 0$. Then $E[|\boldsymbol{L}|^p] < \infty$ if and only if $\int_{|\boldsymbol{x}|>1} |\boldsymbol{x}|^p\ \nu_L(d\boldsymbol{x}) < \infty$.

So far, we defined Lévy processes only on the positive real line $[0, \infty)$. Note that these can be easily extended to the whole line $\mathbb{R}$. Consider therefore a Lévy process $\boldsymbol{L}$ and assume $\boldsymbol{L}'$ to be an independent copy, i.e. $\boldsymbol{L}'$ is an independent Lévy process and satisfies $\boldsymbol{L} \overset{\mathscr{D}}{=} \boldsymbol{L}'$. Then by setting

$$\tilde{\boldsymbol{L}}_t = \begin{cases} \boldsymbol{L}_t, & t \geq 0 \\ -\boldsymbol{L}'_{-t-}, & t < 0, \end{cases} \tag{2.7}$$

we define a two-sided Lévy process $\tilde{\boldsymbol{L}} = (\tilde{\boldsymbol{L}}_t)_{t\in\mathbb{R}}$.
The next section deals with integration with respect to Lévy processes. Especially we are interested in the question under which conditions *Lévy integrals* exist in $L^2(\Omega)$.

2.2 Integration with respect to Lévy processes

In the following we restrict ourselves to univariate Lévy processes L and denote the corresponding diffusion component by $\sigma_L^2 := A_L$ such that the characteristic triplet turns into $(\gamma_L, \sigma_L^2, \nu_L)$. We are now concerned with the definition of integrals of the form

$$\int_{\mathbb{R}} f(s)\ dL_s, \tag{2.8}$$

[1] Recall that a distribution F is said to be infinitely divisible, if for all $n \geq 1$ there exist i.i.d. random variables $X_1, \ldots, X_n$ such that their sum has distribution F.

where $f : \mathbb{R} \to \mathbb{R}$ is a nonrandom function and $L = (L_t)_{t\in\mathbb{R}}$ denotes a two-sided Lévy process.

A very general approach was developed by Rajput and Rosinski [1989], where stochastic integration of deterministic functions is defined with respect to infinitely divisible, independently scattered (IDIS) random measures. This, in particular, includes integrals of the form (2.8), as any Lévy process L defines an IDIS random measure Λ on $\mathcal{B}(\mathbb{R})$ by setting

$$\Lambda([a,b]) := L_b - L_a, \qquad a \leq b.$$

Following this concept we first define (2.8) for step functions f_n,

$$f_n(s) = \sum_{k=1}^{n} a_k 1_{(s_{k-1},s_k)}(s), \tag{2.9}$$

where $s_0, a_i, s_i \in \mathbb{R}$ for $i = 1, \ldots, n$ such that $-\infty < s_0 \leq s_1 \leq \ldots \leq s_n < \infty$, by setting

$$\int_{\mathbb{R}} f_n(s)\, dL_s := \sum_{k=1}^{n} a_k (L_{s_k} - L_{s_{k-1}}). \tag{2.10}$$

Now, the stochastic integral of a measurable function f can be defined using a limit argument.

Definition 2.9 *Let L denote a Lévy process and $f : \mathbb{R} \to \mathbb{R}$ be measurable. Assume there exists a sequence of measurable step functions $f_n : \mathbb{R} \to \mathbb{R}$, which are defined as in* (2.9), *such that*

(i) $f_n \to f$ *a.e. for* $n \to \infty$ *and*

(ii) $\left(\int_{\mathbb{R}} f_n(s)\, dL_s\right)_{n\in\mathbb{N}}$ *converges in probability.*

*Then f is said to be L-*integrable *and its integral with respect to L is defined by*

$$\int_{\mathbb{R}} f(s)\, dL_s := \lim_{n\to\infty} \int_{\mathbb{R}} f_n(s)\, dL_s, \tag{2.11}$$

where the limit is taken in probability.

In particular, if it exists, (2.8) is independent of the approximating sequence. The following result characterizes the integrability of a measurable function f in terms of the characteristic triplet of L.

Theorem 2.10 (Characterization of integrable functions) *Let $q > 0$ and denote by L a Lévy process with characteristics $(\gamma_L, \sigma_L^2, \nu_L)$. Then a measurable function $f : \mathbb{R} \to \mathbb{R}$ is L-integrable in the sense of Definition 2.9, if and only if the following conditions hold*

(i) $\int_{\mathbb{R}} \left| f(s) \left(\gamma_L + \int_{\mathbb{R}_0} x [1_{\{|f(s)x|\leq 1\}} - 1_{\{|x|\leq 1\}}]\, \nu_L(dx) \right) \right| ds < \infty,$

(ii) $\int_{\mathbb{R}} \int_{\mathbb{R}_0} \min(1, |f(s)x|^2)\, \nu_L(dx)\, ds < \infty,$

(iii) $\sigma_L^2 \int_{\mathbb{R}} |f(s)|^2 \, ds < \infty$.

Furthermore, if f is L-integrable and $E[|L_t|^q] < \infty, t \in \mathbb{R}$, then for $0 \leq p \leq q$ it holds that

$$E\Big[\big| \int_{\mathbb{R}} f(s) \, dL_s \big|^p \Big] < \infty$$

if and only if

$$\int_{\mathbb{R}} \int_{\mathbb{R}_0} |f(s)x|^p 1_{\{|f(s)x|>1\}} \, \nu_L(dx) ds < \infty. \tag{2.12}$$

Proof. Theorems 2.7 and 3.3 of Rajput and Rosinski [1989].

Corollary 2.11 *Let L be a Lévy process with characteristics $(\gamma_L, \sigma_L^2, \nu_L)$ and consider a measurable function $f : \mathbb{R} \to \mathbb{R}$.*

(i) Let $E[L_1] \neq 0$. Then it holds that

$$f \text{ is } L\text{-integrable and } E\Big[\Big(\int_{\mathbb{R}} f(s) \, dL_s \Big)^2 \Big] < \infty, \tag{2.13}$$

if and only if

$$f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \text{ and } E[L_1^2] < \infty. \tag{2.14}$$

(ii) Let $E[L_1] = 0$. Then the condition

$$f \in L^2(\mathbb{R}) \text{ and } E[L_1^2] < \infty \tag{2.15}$$

is necessary and sufficient for (2.13)*.*

Proof. Let us assume that (2.14) is satisfied. Clearly condition *(iii)* of Theorem 2.10 holds. For condition (i) we make use of $\gamma_L = E[L_1] - \int_{|x|>1} x \, \nu_L(dx)$ and find that

$$\begin{aligned}
&\int_{\mathbb{R}} \Big| f(s) \Big(\gamma_L + \int_{\mathbb{R}_0} x \left[1_{\{|f(s)x| \leq 1\}} - 1_{\{|x| \leq 1\}} \right] \, \nu_L(dx) \Big) \Big| \, ds \\
&= \int_{\mathbb{R}} \Big| f(s) \Big(E[L_1] + \int_{\mathbb{R}_0} x \left[1_{\{|f(s)x| \leq 1\}} - 1 \right] \, \nu_L(dx) \Big) \Big| \, ds \\
&\leq \big| E\big[L_1\big] \big| \int_{\mathbb{R}} |f(s)| \, ds + \int_{\mathbb{R}} \int_{\mathbb{R}_0} |f(s)x| 1_{\{|f(s)x|>1\}} \, \nu_L(dx) \, ds \\
&\leq \big| E\big[L_1\big] \big| \, ||f||_{L^1(\mathbb{R})} + \int_{\mathbb{R}} \int_{\mathbb{R}_0} (f(s)x)^2 \, \nu_L(dx) \, ds \\
&= \big| E\big[L_1\big] \big| \, ||f||_{L^1(\mathbb{R})} + ||f||^2_{L^2(\mathbb{R})} \int_{\mathbb{R}_0} x^2 \, \nu_L(dx),
\end{aligned}$$

which is finite due to our assumption. Obviously, we can draw the same conclusion in the case where $E[L_1] = 0$ and (2.15) holds.

Observe that the double integrals of Theorem 2.10 (*ii*) as well as that of condition (2.12) (for $p = 2$) are bounded, as

$$\int_{\mathbb{R}}\int_{\mathbb{R}_0} |f(s)x|^2 \, \nu_L(dx)\, ds \le ||f||^2_{L^2(\mathbb{R})} \int_{\mathbb{R}_0} |x|^2 \, \nu_L(dx) < \infty,$$

such that both (2.14) and (2.15) are sufficient for (2.13).

To show the necessity of the conditions we assume that (2.13) is satisfied. Now, the condition (*ii*) of Theorem 2.10 and (2.12) imply

$$\begin{aligned}
\infty &> \int_{\mathbb{R}}\int_{\mathbb{R}_0} \min(1, |f(s)x|^2) \, \nu_L(dx)\, ds + \int_{\mathbb{R}}\int_{\mathbb{R}_0} |f(s)x|^2 1_{\{|f(s)x|>1\}} \, \nu_L(dx) ds \\
&\ge \int_{\mathbb{R}}\int_{\mathbb{R}_0} |f(s)x|^2 1_{\{|f(s)x|\le 1\}} \, \nu_L(dx)\, ds + \int_{\mathbb{R}}\int_{\mathbb{R}_0} |f(s)x|^2 1_{\{|f(s)x|>1\}} \, \nu_L(dx) ds \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}_0} |f(s)x|^2 \, \nu_L(dx) ds \\
&= ||f||^2_{L^2(\mathbb{R})} \int_{\mathbb{R}_0} x^2 \, \nu_L(dx),
\end{aligned} \tag{2.16}$$

which means (assuming that $f \neq 0$ as well as $\nu_L \neq 0$) that $f \in L^2(\mathbb{R})$ and $E[L_1^2] < \infty$. Besides, again using $\gamma_L = E[L_1] - \int_{|x|>1} x \, \nu_L(dx)$ we have from condition (*i*) of Theorem 2.10

$$\begin{aligned}
\infty &> \int_{\mathbb{R}} \left| f(s) \left(\gamma_L + \int_{\mathbb{R}_0} x \left[1_{\{|f(s)x|\le 1\}} - 1_{\{|x|\le 1\}} \right] \, \nu_L(dx) \right) \right| \, ds \\
&= \int_{\mathbb{R}} \left| f(s) \, \left(E[L_1] - \int_{\mathbb{R}_0} x 1_{\{|f(s)x|>1\}} \, \nu_L(dx) \right) \right| \, ds \\
&\ge |E[L_1]| \int_{\mathbb{R}} |f(s)| \, ds - \int_{\mathbb{R}}\int_{\mathbb{R}_0} (f(s)x)^2 \nu_L(dx) ds \\
&\ge |E[L_1]| \, ||f||_{L^1(\mathbb{R})} - ||f||^2_{L^2(\mathbb{R})} \int_{\mathbb{R}_0} x^2 \, \nu_L(dx).
\end{aligned}$$

From (2.16) we know that $||f||^2_{L^2(\mathbb{R})} \int_{\mathbb{R}_0} x^2 \, \nu_L(dx) < \infty$, such that f additionally satisfies $f \in L^1(\mathbb{R})$. □

Having established the well-definedness of (2.9), we obtain the following useful result on its distribution.

Proposition 2.12 *Let L be a Lévy process with characteristic exponent given by*

$$\psi_L(u) = iu\gamma_L - \frac{1}{2}\sigma_L^2 u^2 + \int_{\mathbb{R}_0} \left(e^{iux} - 1 - iux 1_{\{|x|\le 1\}} \right) \, \nu_L(dx), \qquad u \in \mathbb{R}. \tag{2.17}$$

If $f : \mathbb{R} \to \mathbb{R}$ is L-integrable, then (2.8) *is infinitely divisible and its characteristic functions satisfies*

$$E[e^{iu \int_{\mathbb{R}} f(s) \, dL_s}] = e^{\int_{\mathbb{R}} \psi_L(uf(s)) \, ds}, \qquad u \in \mathbb{R}. \tag{2.18}$$

Proof. Proposition 2.6 of Rajput and Rosinski [1989].

In the following we will see that integration with respect to Lévy processes may also be defined in $L^2(\Omega)$-sense. Before considering the general case we restrict ourselves to driving Lévy processes L without Brownian motion component satisfying $E[L_1] = 0$ and $E[L_1^2] < \infty$, such that the corresponding Lévy-Itô-decomposition turns into

$$L_t = \int_{[0,t]} \int_{\mathbb{R}_0} x \; (\mu_L(dx, ds) - \; \nu_L(dx)ds).$$

Theorem 2.13 (Existence in $L^2(\Omega)$-sense) *Let $f \in L^2(\mathbb{R})$ and denote by L a square integrable, zero mean Lévy process without Brownian motion component. Then the stochastic integral* (2.8) *exists as $L^2(\Omega)$-limit of approximating step functions and does not depend on the approximating sequence.*

Proof. Proposition 2.1 of Marquardt [2006]. □

In the case where L also has a Brownian motion component σ_L^2, we can apply the Lévy-Itô-decomposition to represent L as

$$L_t = \tilde{L}_t + \sigma_L^2 B_t, \qquad t \in \mathbb{R},$$

where the characteristic triplet of $\tilde{L}$ is given by $(\gamma_L, 0, \nu_L)$ and B denotes a standard Brownian motion. If $f \in L^2(\mathbb{R})$ the Itô-integral is well-defined (in the $L^2(\Omega)$-sense), which shows that the above theorem also holds in this case. We summarize this result in the subsequent corollary.

Corollary 2.14 *Let $f \in L^2(\mathbb{R})$ and denote by L a square integrable, zero mean Lévy process. Then the stochastic integral* (2.8) *exists as $L^2(\Omega)$-limit of approximating step functions and does not depend on the approximating sequence.*

Finally, the following result shows that the condition of a centered driving Lévy process may be relaxed if the kernel function f is both integrable and square integrable.

Corollary 2.15 *Let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and denote by L a square integrable Lévy process. Then the stochastic integral* (2.8) *exists as $L^2(\Omega)$-limit of approximating step functions and does not depend on the approximating sequence.*

Proof. Firstly, we denote by $\mathcal{H}$ the space $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ eqipped with the norm $\|\cdot\|_{\mathcal{H}} = \|\cdot\|_{L^1(\mathbb{R})} + \|\cdot\|_{L^2(\mathbb{R})}$. Observe that $\mathcal{H}$ defines a Banach space. Further, the space of simple functions is dense in $\mathcal{H}$. Consequently, let $(f_n)_{n\in\mathbb{N}}$ be a sequence of simple functions satisfying $f_n \to f$ in $\mathcal{H}$. Now, using the decomposition

$$L_t = t \; E[L_1] + \tilde{L}_t, \qquad t \in \mathbb{R},$$

where $\tilde{L}_t = L_t - E[L_t]$, we obtain for $m, n \in \mathbb{N}$

$$\left\| \int_{\mathbb{R}} (f_n(s) - f_m(s)) \; dL_s \right\|_{L^2(\Omega)} \leq \left\| \int_{\mathbb{R}} (f_n(s) - f_m(s)) d\tilde{L}_s \right\|_{L^2(\Omega)} + E[L_1] \int_{\mathbb{R}} (f_n(s) - f_m(s)) \; ds. \tag{2.19}$$

The terms on the right-hand side of (2.19) tend to zero for $m, n \to \infty$ (such that $\int f_n \, dL$ constitutes a Cauchy-sequence in $L^2(\Omega)$) due to the previous corollary and the observation that $f_n \longrightarrow f$ in $\mathcal{H}$ and therefore in $L^1(\mathbb{R})$. Hence,

$$\int_{\mathbb{R}} f(s) \, dL_s = E[L_1] \int_{\mathbb{R}} f(s) \, ds + \int_{\mathbb{R}} f(s) \, d\tilde{L}_s$$

exists as limit in the $L^2(\Omega)$-sense. □

Remark 2.16 *Notice that the previous corollary also applies to the case where $f \in L^2(\mathbb{R})$ is compactly supported. As we will see later, this allows for Molchan-Golosov FLPs with non-centered driving Lévy process to be defined in $L^2(\Omega)$-sense (cf. Fink [2013]).*

Remark 2.17 *We saw that for $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and a square integrable Lévy process L the integral $\int_{\mathbb{R}} f(s) \, dL_s$ exists both as limit in probability and in $L^2(\Omega)$. Clearly, the limits agree, as $L^2(\Omega)$-convergence implies convergence in probability and the limit in the latter case is unique.*

We conclude this section by calculating the first two moments of (2.8).

Proposition 2.18 *Consider a function $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and let L be a square integrable Lévy process. Then it holds that*

$$E\Big[\int_{\mathbb{R}} f(s) \, dL_s\Big] = E[L_1] \int_{\mathbb{R}} f(s) \, ds,$$

and

$$E\Big[\Big(\int_{\mathbb{R}} f(s) \, dL_s\Big)^2\Big] = \mathrm{Var}(L_1) \int_{\mathbb{R}} f^2(s) \, ds. \tag{2.20}$$

Observe that in the case where the driving Lévy process has zero mean and unit variance the equation (2.20) shows that $f \mapsto \int f \, dL$ is a linear isometry between $L^2(\mathbb{R})$ and $L^2(\Omega)$, more precisely

$$\Big\| \int_{\mathbb{R}} f(s) \, dL_s \Big\|_{L^2(\Omega)} = \|f\|_{L^2(\mathbb{R})}.$$

We now turn to the proof of the above proposition.

Proof. The fact that the integral is defined as $L^2(\Omega)$-limit of integrals of approximating simple functions $f_n = \sum_{k=1}^n a_k^{(n)} (L_{s_k} - L_{s_{k-1}})$ implies that also $f_n \longrightarrow f$ in $L^1(\Omega)$.

Consequently,

$$\begin{aligned}
E\Big[\int_{\mathbb{R}} f(s)\, dL_s\Big] &= \lim_{n\to\infty} E\Big[\int_{\mathbb{R}} f_n(s)\, dL_s\Big] \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} a_k^{(n)} E[L_{s_k} - L_{s_{k-1}}] \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} a_k^{(n)} (s_k - s_{k-1}) E[L_1] \\
&= E[L_1] \int_{\mathbb{R}} f(s)\, ds.
\end{aligned}$$

Analogously, using $f_n \longrightarrow f$ in $L^2(\Omega)$, we find that

$$\begin{aligned}
E\Big[\Big(\int_{\mathbb{R}} f(s)\, dL_s\Big)^2\Big] &= \lim_{n\to\infty} E\Big[\Big(\int_{\mathbb{R}} f_n(s)\, dL_s\Big)^2\Big] \\
&= \lim_{n\to\infty} \sum_{k=1,\, j=1}^{n} a_k^{(n)} a_j^{(n)} E[\big(L_{s_k} - L_{s_{k-1}}\big)\big(L_{s_j} - L_{s_{j-1}}\big)] \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} (a_k^{(n)})^2 (s_k - s_{k-1}) \mathrm{Var}(L_1) \\
&= \mathrm{Var}(L_1) \int_{\mathbb{R}} f^2(s)\, ds.
\end{aligned}$$

□

2.3 Fractional Calculus

In this section we will give a brief overview over fractional integrals and derivatives. Details on the concept of fractional calculus can be found in Samko et al. [1993]. We start by introducing fractional integrals.

Definition 2.19 *Let $\alpha \in (0,1)$ and consider a finite interval (a,b). Then the integrals*

$$(I_{b-}^{\alpha} f)(x) = \frac{1}{\Gamma(\alpha)} \int_x^b f(t)(t-x)^{\alpha-1} dt, \qquad s \in (a,b), \tag{2.21}$$

and

$$(I_{a+}^{\alpha} f)(x) = \frac{1}{\Gamma(\alpha)} \int_a^x f(t)(x-t)^{\alpha-1} dt, \qquad s \in (a,b), \tag{2.22}$$

where Γ denotes the Gamma-function, are called right- and left-sided fractional Riemann-Liouville integrals of order α, if they exist almost everywhere.

Remark 2.20 *As can be found in Zähle [1998], the integral* (2.21) *and* (2.22) *exist almost everywhere if $f \in L^1([a,b])$.*

We now turn to fractional derivatives, which may be introduced as inverse operation. Let us therefore define the following set of functions,

$$I^{\alpha}_{b-}(L^1) = \{f \in L^1([a,b]) : \text{ It exists } \phi \in L^1([a,b]) \text{ with } f = (I^{\alpha}_{b-}\phi)\},$$

and let $I^{\alpha}_{a+}(L^1)$ be defined analogously.

Definition 2.21 *Let $\alpha \in (0,1)$ and $f \in I^{\alpha}_{b-}(L^1)$ $\big(f \in I^{\alpha}_{a+}(L^1)\big)$. Then the unique function $\phi \in L^1([a,b])$ satisfying $f = I^{\alpha}_{b-}(\phi)$ $\big(f = I^{\alpha}_{a+}(\phi)\big)$ is called right-(left-)sided fractional Riemann-Liouville derivative of f of order α and is almost everywhere given by*

$$(\mathcal{D}^{\alpha}_{b-})(x) = -\frac{1}{\Gamma(1-\alpha)}\frac{d}{dx}\int_x^b f(t)(t-x)^{-\alpha}dt \tag{2.23}$$

$$\left((\mathcal{D}^{\alpha}_{b-})(x) = -\frac{1}{\Gamma(1-\alpha)}\frac{d}{dx}\int_x^b f(t)(t-x)^{-\alpha}dt\right). \tag{2.24}$$

Remark 2.22 1. *The convergence of the integrals* (2.23) *and* (2.24) *at the singularity $t = x$ holds pointwise for almost every x.*

2. *Fractional Riemann-Liouville integrals and derivatives can analogously be defined on the whole real line, i.e. for $a = -\infty$ and $b = \infty$. In this case, we write*

$$(I^{\alpha}_{-}f)(x) = \frac{1}{\Gamma(\alpha)}\int_x^{\infty} f(t)(t-x)^{\alpha-1}dt, \qquad s \in \mathbb{R}, \tag{2.25}$$

and

$$(I^{\alpha}_{+}f)(x) = \frac{1}{\Gamma(\alpha)}\int_{-\infty}^{x} f(t)(x-t)^{\alpha-1}dt, \qquad s \in \mathbb{R}.$$

3. *As it is common in the literature, we write*

$$I^{-\alpha}_{b-} = \mathcal{D}^{\alpha}_{b-} \text{ and } I^{-\alpha}_{a+} = \mathcal{D}^{\alpha}_{a+}.$$

2.4 Fractional Lévy processes

Having seen how stochastic integrals with respect to Lévy processes can be defined, we are now prepared to introduce the class of fractional Lévy processes (FLPs), which represents a generalization of the class of fractional Brownian motions (FBMs). Hence, we start by introducing the latter.

Definition 2.23 (Fractional Brownian motion) *A two-sided Gaussian process $(B^d_t)_{t\in\mathbb{R}}$ is called* fractional Brownian motion *(FBM) with fractional difference parameter*[2] $d \in (-0.5, 0.5)$ *if it has zero mean and covariance structure defined by*

$$\mathrm{Cov}(B^d_t, B^d_s) = \frac{1}{2}\left(|t|^{2d+1} + |s|^{2d+1} - |t-s|^{2d+1}\right). \tag{2.26}$$

[2] Usually, the FBM is defined using the so-called *Hurst* parameter $H \in (0,1)$. In this case, $H = d + \frac{1}{2}$.

In contrast to the ordinary Brownian motion B (which is obtained by setting $d = 0$), the FBM B^d has dependent increments for $d \neq 0$. This feature makes it attractive for applications, where a dependence structure in continuous time is to be modeled. The following result shows that increments of FBMs may even exhibit long memory for particular values of d.

Proposition 2.24 *The increments of a FBM B^d as defined in Definition 2.23 are positively correlated and exhibit long memory if $d \in (0, 0.5)$. In contrast, if $d \in (-0.5, 0)$, then the increments are negatively correlated.*
Furthermore, denote by ρ_d^h the covariance function of the increments of length h, i.e.

$$\rho_d^h(s) = \mathrm{Cov}(B^d_{t+s+h} - B^d_{t+s}, B^d_{t+h} - B^d_t), \qquad t \in \mathbb{R}, s, h > 0.$$

Then we have

$$\frac{\rho_d^h(s)}{s^{2d-1}} \longrightarrow d(2d+1)h^2, \qquad \textit{as } s \longrightarrow \infty.$$

Proof. Let $t \in \mathbb{R}$ and $h, s > 0$. Then it holds

$$\begin{aligned} \mathrm{Cov}(B^d_{t+s+h} - B^d_{t+s}, B^d_{t+h} - B^d_t) &= \mathrm{Cov}(B^d_{t+s+h}, B^d_{t+h}) - \mathrm{Cov}(B^d_{t+s+h}, B^d_t) \\ &\quad - \mathrm{Cov}(B^d_{t+s}, B^d_{t+h}) + \mathrm{Cov}(B^d_{t+s}, B^d_t) \\ &= \frac{1}{2}\left[(s+h)^{2d+1} + (s-h)^{2d+1} - 2s^{2d+1}\right] \qquad (2.27) \\ &= \frac{1}{2}s^{2d+1}\left[\left(1+\frac{h}{s}\right)^{2d+1} + \left(1-\frac{h}{s}\right)^{2d+1} - 2\right]. \end{aligned}$$

Note that by applying Taylor expansion we can rewrite

$$(1+x)^\alpha + (1-x)^\alpha = 2 + \alpha(\alpha-1)x^2 + \mathcal{O}(x^4).$$

Consequently, the covariance (2.27) is positive as long as $2d(2d+1) > 0$, i.e. $d > 0$. Moreover, concerning the asymptotics of (2.27) we find

$$\begin{aligned} \mathrm{Cov}(B^d_{t+s+h} - B^d_{t+s}, B^d_{t+h} - B^d_t) &= \frac{1}{2}s^{2d+1}\left[2d(2d+1)\frac{h^2}{s^2} + \mathcal{O}\left(\frac{1}{s^4}\right)\right] \\ &= d(2d+1)h^2 \; s^{2d-1} + \mathcal{O}(s^{2d-3}). \end{aligned}$$

□

As pointed out by Tikanmäki and Mishura [2011], a FBM B^d may also be represented as integral transformation in the sense of

$$B^d_t = \int_{\mathbb{R}} f(t,s)\, dB_s, \qquad t \in \mathbb{R},$$

where B denotes an ordinary Brownian motion. Possible kernels $f(t,s)$ for such a representation are given in the following.

Definition 2.25 *(i) The so-called* Molchan-Golosov *(MG) kernel* f_d^{MG} *is defined for each* $t > 0$ *in the case* $d \in (0, 0.5)$ *by*

$$f_d^{MG}(t,s) = c_d \int_s^t (u-s)^{d-1} \left(\frac{u}{s}\right)^d \, du \; 1_{[0,t]}(s), \qquad s \in \mathbb{R},$$

and for $d \in (-0.5, 0]$ *by*

$$f_d^{MG}(t,s) = c_d \left(d \left(\frac{t}{s}\right)^d (t-s)^d - s^{-d} \int_s^t (u-s)^d u^{d-1} \, du \right) 1_{[0,t]}(s), \qquad s \in \mathbb{R},$$

where the constant c_d *satisfies*

$$c_d = d \left(\frac{(2d+1)\Gamma(1-d)}{\Gamma(1+d)\Gamma(1-2d)} \right)^{\frac{1}{2}}. \tag{2.28}$$

(ii) The Mandelbrot-van-Ness *(MvN) kernel is defined for* $t \in \mathbb{R}$ *by*

$$f_d^{MvN}(t,s) = \tilde{c}_d \left((t-s)_+^d - (-s)_+^d \right) \; 1_{(-\infty,t]}(s), \qquad s \in \mathbb{R}, \tag{2.29}$$

with constant

$$\tilde{c}_d = \frac{1}{\Gamma(d+1)}. \tag{2.30}$$

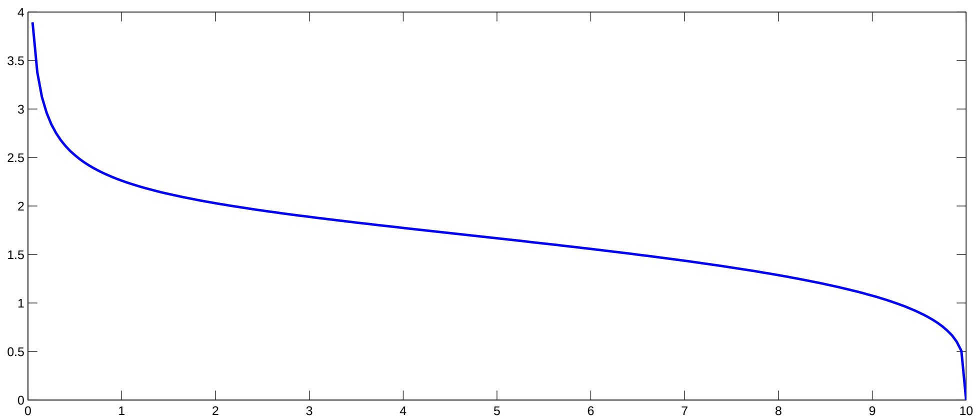

Figure 2.1: Molchan-Golosov kernel $f_d^{MG}(10, \cdot)$ for positive integration parameter $d = 0.25$.

Remark 2.26 *Note that the above defined kernels may also be represented using fractional integrals. To put it more concretely, Fink [2013] uses the following representation for the Molchan-Golosov kernel,*

$$f_d^{MG}(t,s) = c_d \Gamma(d+1) s^{-d} I_{t-}^d \left((\cdot)^d \, 1_{[0,t)}(\cdot) \right)(s) \; 1_{[0,t)}(s), \qquad t > 0, \tag{2.31}$$

where the constant c_d is given by (2.28) *and I^d_{t-} denotes the Riemann-Liouville fractional integral, see* (2.21)*. Moreover, Marquardt [2006] shows that the Mandelbrot-van-Ness kernel can be rewritten as*

$$f_d^{MvN}(t,s) = I_-^d(1_{(0,t]}(\cdot))(s), \qquad t \in \mathbb{R}, \tag{2.32}$$

with I^d_- being defined as in (2.25)*.*

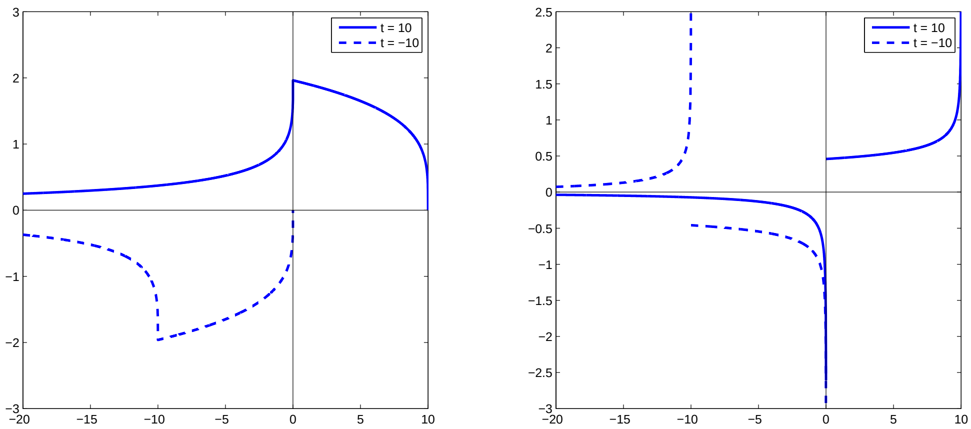

Figure 2.2: Mandelbrot-van-Nes kernel $f_d^{MvN}(t,\cdot)$ for positive integration parameter $d = 0.25$ (left) and negative integration parameter $d = -0.25$ (right) with $t = 10$ (solid line) and $t = -10$ (dashed line).

An obvious difference between these two kernels is that the MG-kernel has a finite support, while the MvN-kernel is infinitely supported. Moreover, we get the following important distinction.

Proposition 2.27 *Consider the kernels in Definition 2.25. Then the following holds.*

(i) If $d \in (-0.5, 0.5)$ then $f_d^{MG}(t,\cdot) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ for all $t > 0$.

(ii) Let $t \in \mathbb{R}$. Assume that $d \in (0, 0.5)$, then $f_d^{MvN}(t,\cdot) \in L^2(\mathbb{R})$ but $f_d^{MvN}(t,\cdot) \notin L^1(\mathbb{R})$. Moreover, if $d \in (-0.5, 0)$ then $f_d^{MvN}(t,\cdot) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$.

Proof. For part (i) see [Jost, 2007, Remark 3.3]. For the proof of the second part see [Engelke and Woerner, 2013, Proposition 2]. □

By means of the MG- and MvN-kernel the FBM can now be represented as integral transformation.

Proposition 2.28 *Let $d \in (-0.5, 0.5)$ and denote by $(B_t^d)_{t\in\mathbb{R}}$ the FBM according to Definition 2.23. Then it holds*

(i)

$$(B_t^d)_{t\geq 0} \stackrel{\mathscr{D}}{=} \left(\int_{\mathbb{R}} f^{MG}(t,s)\, dB_s \right)_{t\geq 0},$$

and

(ii)

$$(B_t^d)_{t\in\mathbb{R}} \stackrel{\mathscr{D}}{=} \left(\int_{\mathbb{R}} c_d^* f^{MvN}(t,s)\, dB_s \right)_{t\in\mathbb{R}}$$

where the constant is given by $c_d^* = [\Gamma(2d+2)\sin(\pi(d+0.5))]^{\frac{1}{2}}$

Remark 2.29 *We directly could have defined the constant in* (2.30) *as*

$$\frac{\left[\Gamma(2d+2)\sin(\pi(d+0.5))\right]^{\frac{1}{2}}}{\Gamma(1+d)},$$

as it is for example done by Tikanmäki and Mishura [2011]. In this case the constant c_d^* would drop out in the above proposition. In defining (2.30) *we followed Marquardt [2006], as we will refer to her paper several times.*

Notice that being a Gaussian process the distribution of a FBM is uniquely determined by its covariance structure (2.26). Hence, it is desirable to drop the assumption of gaussianity to allow for a more flexible distribution while keeping the second order structure. A natural approach to achieve this is to substitute the Brownian motion B in the integral representations of Proposition 2.28 by a Lévy process, leading to the concept of fractional Lévy processes.

2.4.1 Definition by Mandelbrot-van-Ness

We start with the Mandelbrot-van-Ness representation of fractional Lévy processes, which was originally introduced by Benassi et al. [2004] and further studied by Marquardt [2006]. The following definition is due to Tikanmäki and Mishura [2011].

Definition 2.30 (Mandelbrot-van-Ness FLP) *Let L denote a Lévy process without Brownian motion component that satisfies $E[L_1] = 0$ and $E[L_1^2] < \infty$. Furthermore, let f_d^{MvN} be given as in* (2.29)*. Then the stochastic process defined for $d \in (-0.5, 0.5)$ by*

$$L_t^d = \int_{\mathbb{R}} f_d^{MvN}(t,s)\, dL_s, \qquad t \in \mathbb{R}, \tag{2.33}$$

is called Mandelbrot-van-Ness fractional Lévy process (MvN-FLP).

Remark 2.31

1. *As we have seen in Corollary 2.11 and Theorem 2.13 the stochastic integral* (2.33) *can be understood as limit in probability of elementary integrals or in $L^2(\Omega)-$sense.*

2. *As $f_d^{MvN} \notin L^1(\mathbb{R})$ for $d \in (0, 0.5)$, Corollary 2.11 implies that the MvN-FLP cannot be defined for a driving Lévy process with nonzero expectation as long as $d \in (0, 0.5)$.*

The subsequent results shows that MvN-FLP is indeed a generalization of FBM in the sense that it shares the covariance structure (2.26). Futhermore, we give some other useful properties.

Proposition 2.32 *Let $d \in (-0.5, 0.5)$ and denote by L^d a MvN-FLP with driving Lévy process L. Then the following holds.*

(i) The increments of L^d are stationary,

(ii) the covariance structure of L^d coincides with (2.26) *up to a constant, i.e. for $s, t \in \mathbb{R}$ it holds*

$$\operatorname{Cov}(L^d(s), L^d(s)) = \frac{E[L_1^2]}{2\Gamma(2d+2)\sin(\pi(d+0.5))}\left(|t|^{2d+1} + |s|^{2d+1} - |t-s|^{2d+1}\right),$$

(iii) and L^d is not adapted to the filtration generated by its driving Lévy process L.

Proof. The proof of the first two parts can be found for the case $d \in (0, 0.5)$ in [Marquardt, 2006, Theorem 4.4], but can be exercised analogously for the case $d \in (-0.5, 0)$. Finally, observe that for $t < 0$ the kernel $f_d^{MvN}(t, \cdot)$ has support $(-\infty, 0]$, such that

$$\sigma(L_s, s \in (-\infty, 0]) \subseteq \sigma(L^d(s), s \leq t) \nsubseteq \sigma(L_s, s \in (-\infty, t]).$$

Consequently, L^d cannot be adapted to the filtration generated by L. □

Remark 2.33 1. *Marquardt [2006] points out that a MvN-FLP $L^d = (L_t^d)_{t\in\mathbb{R}}$ for $d \in (0, 0.5)$ has a continuous modification. More specifically, for all $t \in \mathbb{R}$ L_t^d is almost surely equal to an improper Riemann integral, see [Marquardt, 2006, Theorem 3.7].*

2. *Further it is shown by Marquardt [2006] that a MvN-FLP must not necessarily have the semimartingale property. A complete characterization of MvN-FLPs being semimartingales is given in the long memory case, i.e. $d \in (0, 0.5)$, by Bender et al. [2012].*

Finally, notice that the equation (2.32) implies that for $d \in (0, 0.5)$

$$\int_{\mathbb{R}} I_-^d(1_{(0,t]}(\cdot))(s)\, dL_s = L_t^d =: \int_{\mathbb{R}} 1_{(0,t]}(s)\, dL_s^d, \qquad t \in \mathbb{R}.$$

Marquardt [2006] shows that this equation holds in a more general setup, too. More precisely, define H to be the completion of $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ with respect to the norm

$$||g||_H = \left[E[L_1^2]\int_{\mathbb{R}} \left(I_-^d g\right)^2(u)\, du\right]^{\frac{1}{2}}.$$

Then the following theorem holds.

Theorem 2.34 (Integration with respect to MvN-FLP) *Denote by $L^d = (L^d_t)_{t\in\mathbb{R}}$ a MvN-FLP and let $g \in H$. Then the integral*

$$\int_{\mathbb{R}} g(s)\, dL^d_s$$

exists as limit in $L^2(\Omega)$. Further it holds that

$$\int_{\mathbb{R}} g(s)\, dL^d_s = \int_{\mathbb{R}} (I^d_- g)(s)\, dL_s.$$

Proof. See [Marquardt, 2006, Theorem 5.3]. □

For more theory on MvN-FLP (at least in the long memory case $d \in (0, 0.5)$) see Marquardt [2006]. We now turn the representation using the MG-kernel.

2.4.2 Definition by Molchan-Golosov

Fractional Lévy processes of Molchan-Golosov type were introduced in Tikanmäki and Mishura [2011] by integrating the Molchan-Golosov kernel f^{MG}_d (see (2.25)) with respect to a zero mean Lévy process without Brownian motion component. This approach was further developed by Fink [2013] by considering the multivariate case as well as allowing for non-centered driving Lévy processes, which may have a Brownian motion component. In the following we restrict ourselves to the case where the driving Lévy process has no Brownian motion component.

Definition 2.35 (Molchan-Golosov - FLP) *Let L denote a Lévy process without Brownian motion component that satisfies $E[L_1^2] < \infty$. Furthermore, let f^{MG}_d be given as in* (2.25)*. Then the stochastic process defined for $d \in (-0.5, 0.5)$ by*

$$L^d_t = \int_{\mathbb{R}} f^{MG}_d(t, s)\, dL_s, \qquad t \geq 0, \tag{2.34}$$

is called Molchan-Golosov fractional Lévy process (FLP-MG).

Remark 2.36
1. *As in the case of MvN-FLPs one can understand the integral* (2.34) *as limit in probability of elementary integrals or in $L^2(\Omega)-$sense (see Corollary 2.11 and Theorem 2.13).*

2. *In contrast to MvN-FLPs, the MG-representation allows for non-centered driving Lévy processes. This is due to the MG-kernel $f^{MG}_d(t, \cdot)$, which is both integrable and square integrable, see section 2.2. Bender and Marquardt [2009], for example, make use of so-called* fractional subordinators, *i.e. MG-FLPs driven by subordinators.*

In the following we summarize some useful or important properties of MG-FLP.

Proposition 2.37 *For $d \in (-0.5, 0.5)$ let $L^d = (L^d_t)_{t\geq 0}$ be a MG-FLP with driving Lévy process $L = (L_t)_{t\geq 0}$. Then the following holds.*

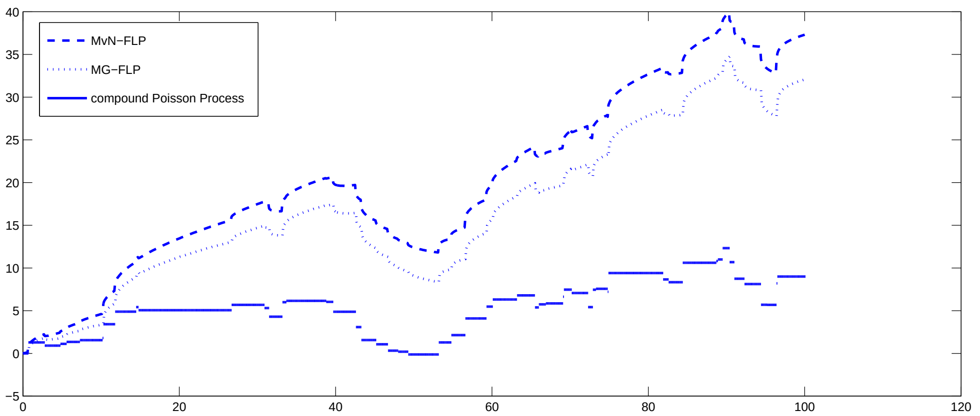

Figure 2.3: Simulated Molchan-Golosov (dotted) and Mandelbrot-van-Ness FLPs (dashed) with fractional integration parameter $d = 0.25$ driven by a compound Poisson process (solid) with rate 0.5 and standard normally distributed jump sizes

(i) The second order structure coincides with (2.26) *up to a constant, i.e. for* $t, s > 0$ *it holds*

$$\mathrm{Cov}(L_t^d, L_s^d) = \frac{\mathrm{Var}(L_1)}{2}\left(|t|^{2d+1} + |s|^{2d+1} - |t-s|^{2d+1}\right),$$

(ii) L^d *is adapted to the filtration generated by its driving Lévy process,*

(iii) and L^d *has for* $d \in (0, 0.5)$ *a continuous modification.*

Proof. The proof of (i) can be found in [Fink, 2013, Proposition 3.1]. Parts $(ii), (iii)$ are due to [Tikanmäki and Mishura, 2011, Proposition 3.11 and Proposition 3.7]. □

Remark 2.38 *An important drawback to MG-FLPs is that in general their increments are not stationary. More precisely, in [Tikanmäki and Mishura, 2011, Proposition 3.11] it is shown, that the increments of a MG-FLP driven by a compound Poisson process* L *with Lévy measure* $\nu_L := \frac{1}{2}(\delta_1 + \delta_{-1})$*, where* $\delta_.$ *denotes the Dirac measure, are non-stationary.*

2.5 Discrete ARMA models

The next two sections are devoted to discrete and continuous ARMA processes. We will see, how the memory of these models can be increased by means of *fractional integration.* Throughout this section, we assume as given a probability space $(\Omega, \mathcal{F}, P)$. Following Beran et al. [2013], we start by introducing the simplest class of time series models.

Definition 2.39 (MA(∞)) *Let* $(\varepsilon_n)_{n\in\mathbb{Z}}$ *be a white noise process. An* infinite moving-average *(MA(∞)) process* $(X_n)_{n\in\mathbb{Z}}$ *is defined by*

$$X_n = \sum_{j=0}^{\infty} a_j \varepsilon_{n-j}, \qquad n \in \mathbb{Z}, \tag{2.35}$$

with so-called moving-average coefficients $a_j \in \mathbb{R}, j \in \mathbb{N}_0$.

Remark 2.40 *In the literature processes with representation* (2.35) *are said to be* causal.

We immediately get the following result concerning the existence of MA(∞) processes.

Theorem 2.41 *Consider a MA(∞) process $(X_n)_{n\in\mathbb{Z}}$ as defined in* (2.35) *with square-summable moving-average coefficients, i.e $\sum_{j=0}^{\infty} a_j^2 < \infty$. Then for all $n \in \mathbb{Z}$*

$$X_n = \sum_{j=0}^{\infty} a_j \varepsilon_{n-j}$$

exists as limit in $L^2(\Omega)-$sense.

Proof. Recall that the space $L^2(\Omega)$ is complete. Using $\sum_{j=0}^{\infty} a_j^2 < \infty$ it is straightforward to see that for all $n \in \mathbb{Z}$ th process $(X_n^m)_{m\in\mathbb{N}_0}$ given by $X_n^m := \sum_{j=0}^{m} a_j \varepsilon_{n-j}, m \in \mathbb{N}_0$, defines a Cauchy sequence in $L^2(\Omega)$. □

Consider a MA(∞) process $(X_n)_{n\in\mathbb{Z}}$ with square summable moving-average coefficients a_j. Then its ACVF satisfies

$$\begin{aligned} \gamma_X(k) &= \mathrm{Cov}(X_{n+k}, X_n) \\ &= \sigma_\varepsilon^2 \sum_{j=0}^{\infty} a_j a_{j+k}, \qquad k \in \mathbb{N}. \end{aligned} \tag{2.36}$$

In particular, we see that $(X_n)_{n\in\mathbb{Z}}$ is weakly stationary.

Definition 2.42 *Let $X = (X_n)_{n\in\mathbb{Z}}$ be a weakly stationary MA(∞) process with ACVF $\gamma_X(\cdot)$ and denote by $C \neq 0$ a constant. Then we say that*

(i) X has long memory if

$$\sum_{k=1}^{\infty} |\gamma_X(k)| = \infty.$$

(ii) X has short memory if

$$\sum_{k=1}^{\infty} |\gamma_X(k)| < \infty.$$

In view of (2.36), it is clear that the second order structure of X is determined by the rate of decay of the MA-coefficients $a_j, j \in \mathbb{N}_0$. The subsequent proposition describes this relationship in more detail and shows that we could have defined the properties of long and short memory by means of the rate of decay of the MA-coefficients, too. This is for example done in [Beran et al., 2013, section 2.1.1.3].

Proposition 2.43 *Let $X = (X_n)_{n\in\mathbb{Z}}$ be a weakly stationary MA(∞) process as given by* (2.35).

(i) If the MA-coefficients satisfy for $C \neq 0$

$$a_j \sim C j^{d-1} \text{ for } d \in (0, 0.5),$$

then

$$\gamma_X(k) \sim C\, k^{2d-1}.$$

In particular, X has long memory.

(ii) If the MA-coefficients satisfy

$$\sum_{j=0}^{\infty} |a_j| < \infty \text{ and } \sum_{j=0}^{\infty} a_j \neq 0,$$

then

$$|\gamma_X(k)| \leq C\, \rho^k, \qquad \rho \in (0,1),$$

where $C > 0$. In particular, X has short memory.

Proof. See [Beran et al., 2013, Lemmas 4.13-4.15]. □

We now take a closer look at a fundamental example of MA(∞) processes.

2.5.1 ARMA(p, q) processes

The popularity of the ARMA class of models is due to its simplicity and flexibility. It was originally introduced by Box and Jenkins [1970].

Definition 2.44 (ARMA(p, q) processes) *Let $p, q \in \mathbb{N}_0$ and $(\varepsilon_n)_{n\in\mathbb{Z}}$ be a white noise process. A stochastic process $X = (X_n)_{n\in\mathbb{Z}}$ is called* auto-regressive moving-average process of order p and q, *shortly ARMA(p, q), if it is the causal stationary solution of the difference equation*

$$\phi(L)X_n = \theta(L)\varepsilon_n, \qquad n \in \mathbb{Z}, \tag{2.37}$$

where ϕ, θ given by

$$\begin{aligned} \phi(z) &= 1 - \sum_{j=1}^{p} \phi_j z^j, \\ \theta(z) &= \sum_{j=1}^{q} \theta_j z^j, \qquad z \in \mathbb{C}, \end{aligned} \tag{2.38}$$

are polynomials in the backshift *or* lag operator L *defined by $LX_n = X_{n-1}$.*

The next result deals with the existence of ARMA processes.

Theorem 2.45 *Assume that $X = (X_n)_{n\in\mathbb{Z}}$ satisfies the difference equation (2.37), where the lag polynomials ϕ, θ have no common roots. Then X is an ARMA(p, q) process if and only if $\phi(z) \neq 0$ for all $z \in \mathbb{C}, |z| \leq 1$.*

Proof. See [Brockwell and Davis, 1991, Theorem 3.1.1]. □

Remark 2.46 *An ARMA(p,q) process X is said to be* invertible*, if there exists an absolute summable sequence of coefficients $(\tilde{\psi}_j)_{j\in\mathbb{N}_0}$, $\sum_{j=0}^{\infty}|\tilde{\psi}_j|<\infty$, such that*

$$\varepsilon_n = \sum_{j=0}^{\infty}\tilde{\psi}_j X_{n-j}, \qquad n\in\mathbb{Z}.$$

In the situation of Theorem 2.45, this is equivalent to $\theta(z)\neq 0$ for all $z\in\mathbb{C}, |z|\leq 1$, see [Brockwell and Davis, 1991, Theorem 3.1.2].

Using the classification in Definition 2.42 we can now characterize the memory of ARMA processes. This is done in the following remark.

Remark 2.47 (Memory of ARMA processes) *Let $X=(X_n)_{n\in\mathbb{Z}}$ be an ARMA(p,q) process as defined in* (2.37)*. Then its MA(∞) representation is given by*

$$X_n = \psi(L)\varepsilon_n = \sum_{j=0}^{\infty}\psi_j\varepsilon_{n-j} \qquad n\in\mathbb{Z}, \tag{2.39}$$

with lag coefficients $\psi_j, j\in\mathbb{N}_0$, determined by

$$\psi(z) = \frac{\theta(z)}{\phi(z)}, \qquad |z|\leq 1. \tag{2.40}$$

It is well-known, that these lag coefficients decay exponentially. We conclude using Proposition 2.43 that ARMA processes have short memory in the sense of Definition 2.42. In particular, its ACVF is bounded exponentially.

Before we end this section, we would like to remark that Box and Jenkins [1970] extended their ARMA definition to include *integrated* processes, leading to the notion of ARIMA(p,d,q) processes.

Definition 2.48 *Let $p,q,d\in\mathbb{N}_0$. Then a process $X=(X_n)_{n\in\mathbb{Z}}$ is called integrated ARMA(p,q) with integration parameter d, shortly ARIMA(p,d,q), process, if*

$$Y_n := (1-L)^d X_n, \qquad n\in\mathbb{Z},$$

defines an ARMA(p,q) process.

Let X denote an ARIMA(p,d,q) process with $d\geq 1$. Then it clearly satisfies the difference equation

$$\phi(L)(1-L)^d X_n = \theta(L)\varepsilon_n, \qquad n\in\mathbb{Z},$$

where $\phi,\theta,(\varepsilon_n)_{n\in\mathbb{Z}}$ are defined as in Definition 2.44 and ϕ,θ have no common zeroes. Now, the autoregressive polynomial is given by

$$z\mapsto \phi(z)(1-z)^d, \qquad z\in\mathbb{C},$$

and obviously has a unit root. Theorem 2.45 therefore implies that X is not weakly stationary. For this reason ARIMA(p,d,q) processes with $d\geq 1$ are also called *unit root non-stationary.*

2.5.2 ARFIMA(p, d, q) processes

The short memory property of ARMA(p, q) processes as mentioned in Remark 2.47 raises the question whether it is possible to modify these processes to capture long range dependence. Granger and Joyeux [1980] and Hosking [1981] independently addressed this issue. Considering the ARIMA(p, d, q) model (see Definition 2.48), they proposed to allow for noninteger values of d, leading to the class of *fractionally integrated* ARMA processes. Before we give the definition, we introduce an important linear time-invariant filter. The following definition stems from [Brockwell and Davis, 1991, section 13.2]

Definition 2.49 *Let $d > -1$. Then the operator $(1 - L)^d$, which can be rewritten by means of a binomial expansion as*

$$(1 - L)^d = \sum_{j=0}^{\infty} \pi_j^d L^j \tag{2.41}$$

with coefficients satisfying

$$\pi_j^d = \frac{\Gamma(j - d)}{\Gamma(j + 1)\Gamma(-d)}, \qquad j \in \mathbb{N}_0, \tag{2.42}$$

and where Γ denotes the Gamma-function,

$$\Gamma(x) = \begin{cases} \int_0^\infty t^{x-1} e^{-t} dt, & x > 0, \\ \infty, & x = 0, \\ x^{-1}\Gamma(1 + x), & x < 0, \end{cases}$$

is called fractional difference operator*.*

Remark 2.50 *According to [Brockwell and Davis, 1991, section 13.2] the coefficients π_j^d can be rewritten as*

$$\pi_0^d = 1,$$
$$\pi_j^d = \prod_{0<k\leq j} \frac{k - 1 - d}{k}, \qquad j \geq 1.$$

In particular, if $d \in (0, 1)$ then $\pi_j^d < 0$ for all $j \geq 1$, while for $d \in (-1, 0)$ it holds that $\pi_j^d > 0, j \geq 0$.

The relevance of the filter $(1 - L)^d$ is due to the asymptotic behavior of the coefficients (2.42) in its series expansion. These decay by a slow hyperbolic rate, as shown in the following proposition.

Proposition 2.51 *Let $d > -1$ and consider the fractional difference operator* (2.41)*. Then the coefficients* (2.42) *satisfy*

$$\pi_j^d \sim \frac{1}{\Gamma(d)} j^{-1-d}, \qquad \text{for } j \longrightarrow \infty. \tag{2.43}$$

Proof. See [Brockwell and Davis, 1991, Remark 4 in section 13.2]. □

We are now prepared to state the definition of the ARFIMA model.

Definition 2.52 (ARFIMA(p, d, q)) *Let $d \in (-1, 0.5)$. Then the weakly stationary and causal solution of the difference equation*

$$\phi(L)(1-L)^d X_n = \theta(L)\varepsilon_n, \qquad n \in \mathbb{Z}, \tag{2.44}$$

where the lag polynomials ϕ and θ as well as the sequence $(\varepsilon_n)_{n\in\mathbb{Z}}$ are defined as in Definition 2.44, is called fractionally integrated ARMA(p, q) process with fractional difference parameter d*, or shortly ARFIMA(p, d, q).*

Remark 2.53 (Admissible values of d) *Note that the range of admissible values for the fractional difference parameter d is limited in the above definition to $(-1, 0.5)$. According to [Beran et al., 2013, section 2.1.1.4] this way the well-definedness of the filter $(1-L)^{-d} = \sum_{j=0}^{\infty} \pi_j^{-d} L^j$ is ensured. To see this, recall that (2.43) implies*

$$\pi_j^{-d} \sim \frac{1}{\Gamma(d)} j^{d-1} \text{ as } j \longrightarrow \infty.$$

Applying [Brockwell and Davis, 1991, Theorem 4.10.1], $(1-L)^{-d}$ thus defines a valid filter in the $L^2(\Omega)$-sense, i.e. when being applied to a weakly stationary sequence, say $\varepsilon = (\varepsilon_n)_{n\in\mathbb{Z}}$, then the random variables

$$Y_n := (1-L)^{-d}\varepsilon_n, \qquad n \in \mathbb{Z},$$

are well-defined as limit in $L^2(\Omega)$ and the process Y is again weakly stationary.
In the literature, however, the ARFIMA model is usually defined for values of d in $(-0.5, 0.5)$, as in this case by the same argumentation also the "inverse" filter $(1-L)^d$ is well defined, such that the difference equation (2.44) does not only admit a causal stationary solution but also an invertible one (cf. [Brockwell and Davis, 1991, Remark 7, section 13.2]).

The subsequent result characterizes the existence of ARFIMA processes.

Theorem 2.54 *Let $d \in (-1, 0.5)$ and consider the difference equation (2.44), where the lag polynomials ϕ, θ are assumed to have no common zeroes. Then the solution X is causal and weakly stationary if and only if $\phi(z) \neq 0$ for all $z \in \mathbb{C}, |z| \leq 1$.*

Proof. See [Brockwell and Davis, 1991, Theorem 13.2.2]. □

Before we finish this section, we summarize the memory properties of the ARFIMA model in the following remark.

Remark 2.55 (Memory of ARFIMA) *Let X denote an ARFIMA(p, d, q) process satisfying the difference equation* (2.44)*. Recall that the process*

$$\tilde{X}_n := (1 - L)^{-d} \varepsilon_n = \sum_{j=0}^{\infty} \tilde{\psi}_j \varepsilon_{n-j}, \qquad n \in \mathbb{Z},$$

has lag coefficients $\tilde{\psi}_j$ decreasing by $\tilde{\psi}_j \sim \frac{1}{\Gamma(d)} j^{d-1}$. Following Beran et al. [2013] the crucial point is that additionally applying the ARMA filter $\frac{\theta(L)}{\phi(L)}$ does not change the asymptotics of the lag coefficients. In particular, for $d \in (0, 0.5)$ we find that the lag coeffcients ψ_j of

$$X_n = (1 - L)^{-d} \frac{\theta(L)}{\phi(L)} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{n-j}, \qquad n \in \mathbb{Z}$$

again decline by

$$\psi_j \sim \frac{1}{\Gamma(d)} j^{d-1},$$

implying that they are not summable. In the sense of Definition 2.43 this means that X has long memory. In contrast, for a fractional difference parameter $d < 0$, the coefficients ψ_j are summable, i.e. X has short memory.

2.6 Continuous ARMA models

We now turn to the continuous time analogue of ARMA processes, which are usually referred to as CARMA processes. A detailed introduction can be found in Priestley [1981]. See also Brockwell [2001a], were properties such as conditions for stationarity are discussed. As it is pointed out there, CARMA processes were originally defined with Brownian motion as driving noise. Yet, in particular for financial applications the gaussianity of a CARMA driven by a Brownian motion is too restrictive. A natural approach to incorporate more flexibility into the model, consists in substituting the driving Brownian motion by a Lévy process. This for example allows for much richer marginal distributions, such as more heavy-tailed ones.

2.6.1 Lévy-driven CARMA processes

Following Brockwell [2001b], Lévy-driven CARMA(p, q) processes with $q < p$ can be introduced in analogy to the linear ARMA(p, q) difference equation (2.37) by the following pth-order linear differential equation

$$b(D)X_t = a(D)DL_t, \qquad t \geq 0, \tag{2.45}$$

where D denotes the differential operator (with respect to t) and $b(\cdot), a(\cdot)$ denote polynomials given by

$$\begin{aligned} b(z) &= z^p + \beta_1 z^{p-1} + \ldots + \beta_p \\ a(z) &= \alpha_0 + \alpha_1 z + \ldots + \alpha_p z^p, \end{aligned} \tag{2.46}$$

with $\alpha_j = 0, q < j \leq p$. This representation is purely formal, as in general the derivative DL_t does not exist in the usual sense. Yet, transforming (2.45) into a system of *obervation* and *state* equations, Lévy-driven CARMA processes may be defined in the following way.

Definition 2.56 (Lévy-driven CARMA(p,q) process) *Let $(L_t)_{t\geq 0}$ denote a Lévy process. Then $(X_t)_{t\geq 0}$ is called* Lévy driven CARMA*(p,q)* process *with location parameter c, moving-average coefficients $\alpha_i, i = 0, \ldots, q$ and autoregressive coefficients $\beta_i, i = 1, \ldots, p$ with $p > q$ if*

$$X_t = c + \boldsymbol{a}'\boldsymbol{Y_t}, \qquad t \geq 0, \tag{2.47}$$

where $\boldsymbol{a}' = [\alpha_0, \ldots, \alpha_p], \alpha_i = 0, i > q$ and $(\boldsymbol{Y}_t)_{t\geq 0}$ is the strictly stationary solution of the SDE

$$d\boldsymbol{Y}_t = B\boldsymbol{Y}_t\, dt + \boldsymbol{e}\, dL_t, \qquad t > 0, \tag{2.48}$$

with

$$B = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \ldots & 1 \\ -\beta_p & -\beta_{p-1} & -\beta_{p-2} & \ldots & -\beta_1 \end{bmatrix} \text{ and } \boldsymbol{e} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \tag{2.49}$$

and assuming $\boldsymbol{Y}_0$ to be independent of $\{L_t, t \geq 0\}$.

In the following we assume without loss of generality that the location parameter c equals zero. The solution to the state equation (2.48) is given by

$$\boldsymbol{Y}_t = e^{Bt}\boldsymbol{Y}_0 + \int_0^t e^{B(t-u)}\boldsymbol{e} dL_u, \qquad t \geq 0. \tag{2.50}$$

The subsequent theorem gives conditions ensuring the state process and therefore the CARMA process itself to be stationary.

Theorem 2.57 (Existence of Lévy-driven CARMA processes) *Let $(L_t)_{t\geq 0}$ be a Lévy process with $E[|L_1|^r] < \infty$ for some $r > 0$. Assume that all the roots of the autoregressive polynomial $b(z)$ in (2.46) have negative real parts. Further, let $\boldsymbol{X}_0$ be independent of $(L_t)_{t\geq 0}$ with distribution*

$$\boldsymbol{X}_0 \stackrel{\mathscr{D}}{=} \int_0^\infty e^{Bu}\boldsymbol{e}\, dL_u. \tag{2.51}$$

Then the strictly stationary solution to (2.47) and (2.48) is given by

$$X_t = \boldsymbol{a}'e^{Bt}\boldsymbol{X}_0 + \int_0^t \boldsymbol{a}'e^{B(t-u)}\boldsymbol{e}\, dL_u, \qquad t \geq 0. \tag{2.52}$$

If the driving Lévy process L additionally satisfies $E[L_1^2] < \infty$, then the solution given by (2.51) and (2.52) is also weakly stationary.

Proof. See [Brockwell, 2001b, Theorem 2.2]. □

A Lévy-driven CARMA process can also be defined on the whole real line. This leads to the following definition.

Definition 2.58 *Consider a two-sided Lévy process L satisfying $E[L_1^2] < \infty$ and denote by $f : \mathbb{R} \to \mathbb{R}$ a measurable function with $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Then the process defined by*

$$Z_t = \int_{-\infty}^{\infty} f(t-u) dL_u, \qquad t \in \mathbb{R}, \tag{2.53}$$

is called (continuous-time) moving-average *process with kernel f. If additionally $f(u) = 0, u < 0$, then Z is said to be* causal*.*

Remark 2.59 *1. According to Corollary 2.11 the stochastic integral in* (2.53) *is well-defined as limit in probability of integrals of step functions approximating the kernel f.*

2. As proved in [Applebaum, 2004, Theorem 4.3.16], a continuous-time moving average process is indeed strictly stationary.

To obtain a causal moving-average representation, let $L = (L_t)_{t\in\mathbb{R}}$ be a two-sided Lévy process as defined in (2.7) and assume that $E[L_1] = 0$ and $E[L_1^2] < \infty$. As shown by Marquardt [2006], the CARMA kernel $g(t) := \boldsymbol{a}' e^{Bt} \boldsymbol{e}$ is exponentially bounded, such that $g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$.
Now assume that all the roots of the autoregressive polynomial $b(\cdot)$ have negative real parts and consider the strictly stationary solution as given by (2.51) and (2.52). Due to the construction of L we find

$$\boldsymbol{X}_0 \overset{\mathscr{D}}{=} \int_0^{\infty} e^{Bu} \boldsymbol{e}\, dL_u \overset{\mathscr{D}}{=} \int_{-\infty}^{0} e^{-Bu} \boldsymbol{e}\, dL_u. \tag{2.54}$$

Therefore the solution X can be rewritten as

$$X_t = \int_{-\infty}^{t} g(t-u)\, dL_u, \qquad t \in \mathbb{R}, \tag{2.55}$$

with deterministic kernel

$$g(t) = \boldsymbol{a}' e^{Bt} \boldsymbol{e}, \qquad t \geq 0. \tag{2.56}$$

Before we give a result conderning the memory properties of CARMA processes, we define what we understand under short and long memory in the continuous-time setting.

Definition 2.60 *A stochastic process $X = (X_t)_{t\in\mathbb{R}}$ is said to have* long memory *if its ACVF γ_X satisfies*

$$\int_0^{\infty} \gamma_X(h)\, dh = \infty.$$

Otherwise we say that X has short memory*.*

Recall that for a discrete ARMA model both the lag coefficients as well as the ACVF decay by an exponential rate. An analogous result can be shown in the continuous-time setting.

Proposition 2.61 *Let the roots $\lambda_1, \ldots, \lambda_p$ of the autoregressive polynomial of a CARMA(p,q) process X as defined in Definition 2.56 have negative real parts and be distinct, then the kernel g in* (2.56) *satisfies*

$$g(u) = \sum_{r=1}^{p} \frac{a(\lambda_r)}{b'(\lambda_r)} e^{\lambda_r u} \, 1_{(0,\infty)},$$

and the ACVF γ_X is given by

$$\gamma_X(h) = \mathrm{Cov}(X_{t+h}, X_t) = \sum_{r=1}^{p} \frac{a(\lambda_r)a(-\lambda_r)}{b'(\lambda_r)b(-\lambda_r)} e^{\lambda_r |h|}.$$

In particular, X is a short memory process.

Proof. See Brockwell [2004]. □

2.6.2 Lévy-driven FICARMA processes

In the last section we saw that (Lévy-driven) CARMA processes have exponentially decaying ACFs. In order to incorporate long memory into CARMA models, Brockwell [2004] substituted the kernel g in the causal CARMA representation (2.55) by the corresponding fractional Riemann-Liouville integral. This leads to the class of fractionally integrated CARMA (FICARMA) processes. We will see that it is important to distinguish between centered, i.e. $E[L_1] = 0$, and non-centered driving Lévy processes.

2.6.2.1 Centered driving Lévy process

Following Brockwell [2004] we calculate the fractional Riemann-Liouville integral of the CARMA kernel $g(t) = \boldsymbol{a}' e^{Bt} \boldsymbol{e} \, 1_{[0,\infty)}(t), t \in \mathbb{R}$. We get

$$\begin{aligned} g_d(t) = (I_+^d g)(t) &= \int_{-\infty}^{t} g(u) \frac{(t-u)^{d-1}}{\Gamma(d)} \, du \\ &= \int_{0}^{\infty} g(t-u) \frac{u^{d-1}}{\Gamma(d)} \, du \\ &= \int_{0}^{t} g(t-u) \frac{u^{d-1}}{\Gamma(d)} \, du. \end{aligned} \tag{2.57}$$

Proposition 2.62 *The asymptotic behavior of the fractionally integrated kernel g_d is given by*

$$g_d(t) \sim \frac{t^{d-1}}{\Gamma(d)} \frac{a(0)}{b(0)}, \qquad t \longrightarrow \infty.$$

In particular, if $d \in (0, 0.5)$, then $g_d \in L^2(\mathbb{R})$, but $g_d \notin L^1(\mathbb{R})$.

Proof. See [Brockwell, 2004, p. 381]. □

We are now in the position to give the definition of FICARMA processes.

Definition 2.63 (FICARMA(p, d, q)) *Let $d \in (0, 0.5)$ and denote by L a two-sided Lévy process with $E[L_1] = 0$ and $E[L_1^2] < \infty$. Let $\boldsymbol{a}, B, \boldsymbol{e}$ be given as in Definition 2.56. Then the stationary process defined by*

$$X_t = \int_{-\infty}^{t} g_d(t-u)\, dL_u, \qquad t \in \mathbb{R}, \tag{2.58}$$

with kernel g_d defined as in (2.57)*, is called* Lévy-driven FICARMA(p, d, q) *process.*

Before we turn to the memory properties of FICARMA processes, we summarize two observations.

Remark 2.64

1. *In view of Corollary 2.11 and the fact that the kernel g_d is not integrable, the condition $E[L_1] = 0$ is necessary to ensure the stochastic integral in* (2.58) *to be well- defined as limit in probability of integrals of approximating simple functions. Due to Corollary 2.14 it is even well-defined as limit in $L^2(\Omega)$.*

2. *Let $d \in (0, 0.5)$. The discrete ARFIMA(p, d, q) process can be represented as moving-average process by*

$$X_n = \sum_{j=0}^{\infty} \psi_j \varepsilon_{n-j} = (1-L)^{-d} \frac{\theta(L)}{\phi(L)} \varepsilon_n, \qquad n \in \mathbb{Z},$$

where $(\varepsilon_n)_{n\in\mathbb{Z}}$ is a white noise process and ϕ, θ are polynomials of degrees p, q. Observe that the lag coefficients ψ_j are given as convolution of exponentially decaying ARMA coefficients $\tilde{\psi}_j$, determined by $\sum_{j=0}^{\infty} \tilde{\psi}_j = \frac{\theta(L)}{\phi(L)}$, and hyperbolically decreasing coefficients π_j^{-d}, $\sum_{j=0}^{\infty} \pi_j^{-d} = (1-L)^{-d}$, which satisfy $\pi_j^{-d} \sim j^{d-1}/\Gamma(d)$. Consequently, fractionally integrating the CARMA kernel in the sense of (2.57) *directly corresponds to the approach taken in the discrete case.*

When it comes to the memory structure of fractionally integrated CARMA processes, the analogy to the discrete case again becomes apparent.

Proposition 2.65 *Let $d \in (0, 0.5)$ and X be a FICARMA(p, d, q) process as defined in Definition 2.63. Denote by γ_X its autocovariance function. Then*

$$\gamma_X(h) \sim h^{2d-1} \frac{E[L_1^2]\Gamma(1-2d)}{\Gamma(d)\Gamma(1-d)} \left[\frac{a(0)}{b(0)}\right]^2, \qquad h \longrightarrow \infty.$$

Consequently, X is a long memory process.

Having successfully defined integration with respect to MvN-FLP (see Theorem 2.34), Marquardt [2006] proves that long memory can be incorporated into CARMA models in to ways: Either by fractionally integrating the CARMA kernel as we did above, or by integrating memory into the driving Lévy process. We summarize this result in the following theorem.

Theorem 2.66 *Consider a FICARMA(p,d,q) process X as defined in Definition 2.63 and additionally assume the driving Lévy process L to have no Brownian motion component. Then X can be represented as*

$$X_t = \int_{-\infty}^{t} g(s)\, dL_s^d, \qquad t \in \mathbb{R},$$

where $L^d = (L_t^d)_{t\in\mathbb{R}}$ denotes the MvN-FLP driven by L and g is the CARMA kernel as defined in (2.56)*.*

Proof. See [Marquardt, 2006, Theorem 6.5]. □

2.6.2.2 Non-centered driving Lévy process

As mentioned in the previous section the non-integrability of the kernel $g_d \notin L^1(\mathbb{R})$ requires the driving Lévy process to have zero mean. However, allowing for the possibility of a non-decreasing driving Lévy process L implying $E[L_1] \geq 0$ is desirable, as this way the non-negativity of the resulting FICARMA can be ensured. Consequently, FICARMA processes could be applied when modeling volatility.
In view of the asymptotic behavior

$$g_d(t) \sim Ct^{d-1}, \qquad t \to \infty,$$

the tail of g_d can be made integrable by restricting the parameter d to be negative. In this case, however, g_d fails to be integrable at $t = 0$. Brockwell and Marquardt [2005] therefore suggest to modify the Lévy-driven FICARMA process in Definition 2.63 by substituting the kernel g_d with

$$g_{a,d}(t) = \int_0^t g(t-u)h_{a,d}(u)du, \qquad a > 0, d < 0, \tag{2.59}$$

where

$$h_{a,d}(u) = K_{a,d}\min(a^{d-1}, u^{d-1})1_{(0,\infty)}(u),$$

and $K_{a,d}$ is a constant[3]. Clearly, $g_{a,d}$ is now both integrable and square integrable, such that the driving Lévy process is not restricted to have zero mean. This leads to the following definition.

Definition 2.67 *Let $d < 0$ and L be a two-sided Lévy process with $E[L_1^2] < \infty$. Then the stationary process*

$$X_t = \int_{-\infty}^{t} g_{a,d}(t-u)dL_u, \qquad t \in \mathbb{R},$$

where the kernel $g_{a,d}$ is given by (2.59)*, is referred to as* non-centered Lévy driven FICARMA(p,d,q) *process.*

[3] Brockwell and Marquardt [2005] suggest to choose $K_{a,d} = a^{|d|}\frac{|d|}{1+|d|}$, such that $h_{a,d}$ constitutes a probability density.

Proposition 2.68 *For $d < 0$ consider a non-centered Lévy driven FICARMA(p,d,q) process X as defined above and denote by γ_X its ACVF. Then it holds that*

$$\gamma_X(h) \sim K_{a,d} \left[\frac{a(0)}{b(0)}\right]^2 t^{d-1}. \tag{2.60}$$

Proof. See Brockwell and Marquardt [2005]. □

Obviously, a non-centered Lévy driven FICARMA(p,d,q) process does not exhibit long memory in the sense of an integrable ACVF, see Definition 2.60. Yet, observe that for d tending to zero in (2.60), the ACVF approaches non-integrability.

Chapter 3

ARCH models and fractional integration

This chapter constitutes the main part of this thesis. First, we give a review of ARCH(∞) processes and their abilities to incorporate long range dependence. This leads us to FIGARCH processes. Secondly, we present the COGARCH model and address the question, whether there exists a continuous-time analogue to the FIGARCH.

3.1 Discrete Time Models

3.1.1 ARCH(∞) processes

We start with the definition of ARCH(∞) models.

Definition 3.1 (ARCH(∞)) *Let $(\varepsilon_n)_{n\in\mathbb{Z}}$ be a sequence of i.i.d random variables with zero mean. Then $(X_n)_{n\in\mathbb{Z}}$ is said to be an ARCH(∞) process, if it satisfies the system of equations*

$$\begin{aligned} X_n &= \sigma_n \varepsilon_n, \\ \sigma_n^2 &= \omega + \sum_{i=1}^{\infty} \psi_i X_{n-i}^2, \qquad n \in \mathbb{Z}, \end{aligned} \tag{3.1}$$

with $\omega \in (0,\infty)$ and $\psi_i \geq 0$ for $i \in \mathbb{N}$.

Remark 3.2 *Obviously, the famous ARCH(p) model of Engle [1982] is comprised by the class of ARCH(∞) models. It can be obtained by setting $\psi_i = 0$ for all $i > p$.*

A solution $(X_n)_{n\in\mathbb{Z}}$ of the ARCH(∞) model (3.1) is called *causal*, if X_n is measurable with respect to the σ-algebra generated by $\varepsilon_u, u \leq n$. The latter is denoted by

$$\mathcal{F}_n := \sigma(\varepsilon_u, u \leq n). \tag{3.2}$$

An important question is, what conditions need to be imposed on the lag coefficients $\psi_i, i \in \mathbb{N}$, as well as on the distribution of the noise ε_n to ensure the existence of stationary and causal solutions to (3.1). Douc et al. [2008] addressed this question by establishing sufficient conditions for the strict stationarity and showed the uniqueness of the corresponding solution. Giraitis et al. [2000] considered the squares $(X_n^2)_{n\in\mathbb{Z}}$ of an ARCH(∞) model giving sufficient conditions for weak and strict stationarity. Finally, necessary and sufficient conditions for weakly stationary solutions of (3.1) are provided by Zaffaroni [2004] for both the levels X_n and the squares X_n^2. We present their results in the following.

Theorem 3.3 (Stationarity of ARCH(∞) models) *There exists a strictly stationary and causal solution to the ARCH(∞) model* (3.1)*, if there is an $s \in (0,1]$ such that*

$$E[|\varepsilon_n|^{2s}]\sum_{i=1}^{\infty}\psi_i^s < 1. \tag{3.3}$$

The solution is then given by

$$\begin{aligned} X_n &= \sigma_n\varepsilon_n, \\ \sigma_n^2 &= \omega + \omega\sum_{k=1}^{\infty}\sum_{i_1,\dots,i_k\geq 1}\psi_{i_1}\dots\psi_{i_k}\varepsilon_{n-i_1}^2\dots\varepsilon_{n-i_1-\dots-i_k}^2, \qquad n\in\mathbb{Z}, \end{aligned} \tag{3.4}$$

and it represents the unique strictly stationary and causal solution such that $E[|X_n|^{2s}] < \infty$.

Proof. See [Douc et al., 2008, Theorem 1]. □

In the literature the equation (3.4) is known as *Volterra expansion*, cf. Giraitis et al. [2000].

Remark 3.4 *Obviously, if* (3.3) *is satisfied with $s = 1$, then the solution* (3.4) *is also weakly stationary. On the other hand, given a weakly stationary ARCH(∞) process $(X_n)_{n\in\mathbb{Z}}$, by applying expectations on both sides of* (3.1) *and rearranging terms, we obtain*

$$E[\sigma_n^2] = \frac{\omega}{1 - E[\varepsilon_n^2]\sum_{i=1}^{\infty}\psi_i}, \qquad n\in\mathbb{Z}. \tag{3.5}$$

Consequently,

$$E[\varepsilon_n^2]\sum_{i=1}^{\infty}\psi_i < 1 \tag{3.6}$$

is a necessary and sufficient condition for the weak stationarity of the ARCH(∞) model (3.1)*.*

As already mentioned, a typical observation for financial time series, such as series of price returns is the absence of autocorrelations, whereas the corresponding squared series is characterized by strong autocorrelation. This fact is captured by ARCH(∞) models. Right from the definition (3.1) it follows (assuming weak stationarity) for all $k\in\mathbb{N}$, that

$$\begin{aligned} E[X_{n+k}X_n] &= E[\sigma_{n+k}\sigma_n\varepsilon_n E[\varepsilon_{n+k}|\mathcal{F}_{n+k-1}]] \\ &= 0, \qquad n\in\mathbb{Z}, \end{aligned}$$

since ε_{n+k} is centered and independent of $\mathcal{F}_{n+k-1}$.
To analyze the second order structure of a squared ARCH(∞) process, consider the model

$$Y_n = \sigma_n^2 \varepsilon_n^2, \qquad \sigma_n^2 = \omega + \sum_{j=1}^{\infty} \psi_j Y_{n-j}, \tag{3.7}$$

with $(\varepsilon_n)_{n\in\mathbb{Z}}$ as before a zero mean, i.i.d. sequence of random variables and $\omega \in (0,\infty)$. While [Giraitis et al., 2000, Theorem 2.1] provided the sufficient condition

$$E[\varepsilon_n^4]^{1/2} \sum_{j=1}^{\infty} \psi_j < 1 \tag{3.8}$$

for weak stationarity of (3.7), the latter was completely characterized by Zaffaroni [2004].

Theorem 3.5 (Weak stationarity of squared ARCH(∞) processes) *Consider the model* (3.7) *and assume that*

$$\xi(z) := 1 - E[\varepsilon_n^2] \sum_{j=1}^{\infty} \psi_j z^j \tag{3.9}$$

is invertible, that means, it exists a sequence $\delta_n, n \in \mathbb{N}_0$ with $\delta_0 = 1$ and $\sum_{j=0}^{\infty} \delta_j^2 < \infty$ such that

$$\frac{1}{\xi(z)} = \sum_{j=0}^{\infty} \delta_j z^j. \tag{3.10}$$

Then the model (3.7) *admits a weakly stationary solution, if and only if*

$$\mathrm{Var}(\varepsilon_n^2) \sum_{u=-\infty}^{\infty} \chi_\delta(u)\chi_{\tilde{\psi}}(u) < 1, \tag{3.11}$$

where $\tilde{\psi}_0 = 0, \tilde{\psi}_k = \psi_k, k \geq 1$ and $\chi_c(u) = \sum_{k=0}^{\infty} c_k c_{k+u}, u \in \mathbb{Z}$, for a square summable sequence $(c_n)_{n\in\mathbb{N}_0}$.
In this case, it holds for $v_n := E[Y_n] - E[Y_n|\mathcal{F}_{n-1}]$ that

$$\mathrm{Cov}(Y_n, Y_0) = E[v_n^2]\chi_\delta(n), \qquad n \in \mathbb{Z}, \tag{3.12}$$

where

$$E[v_n^2] = \mathrm{Var}(\varepsilon_n^2) \left(\frac{E[Y_n]}{E[\varepsilon_n^2]}\right)^2 \frac{1}{1 - \mathrm{Var}(\varepsilon_n^2)\sum_{u=-\infty}^{\infty} \chi_\delta(u)\chi_{\tilde{\psi}}(u)}.$$

Proof. See [Zaffaroni, 2004, Theorem 1].

Remark 3.6 *Note that if the invertibility condition* (3.10) *holds, then the volatility process*

$$\sigma_n^2 = \omega + \sum_{j=1}^{\infty} \psi_j Y_{n-j}$$

can be rewritten as

$$\left(1 - E[\varepsilon_n^2]\sum_{j=1}^{\infty}\psi_j L^j\right) Y_n = E[\varepsilon_n^2](\omega - \sigma_n^2) + Y_n,$$

where L denotes the backshift operator, defined by $LY_n = Y_{n-1}$. Using the martingale difference sequence $v_n = Y_n - \sigma_n^2 E[\varepsilon_n^2], n \in \mathbb{Z}$, we finally obtain the linear representation

$$Y_n = E[\varepsilon_n^2]\omega \sum_{j=0}^{\infty}\delta_j + \sum_{j=0}^{\infty}\delta_j v_{n-j}, \qquad n \in \mathbb{Z}. \tag{3.13}$$

Due to the empirical relevance of long memory, the question arises, if ARCH(∞) models are capable of capturing this property. Recall that for MA(∞) models the memory was determined by the MA-coefficients (see Proposition 2.43). The results [Giraitis et al., 2000, Proposition 3.1] and [Kokozka and Leipus, 2000, Theorem 3.1] indicate that this relationship also holds in the case of ARCH(∞) models. Subsequently, we present a fundamental theorem in this respect which basically refines the two just mentioned findings.

Theorem 3.7 *Consider the model* (3.7)*. Assume that the invertibility condition* (3.10) *as well as* (3.6) *hold. If the coefficients $\psi_n, n \in \mathbb{N}$, tend to zero exponentially or more slowly than exponentially, i.e. if for $c \in (0,1)$*

$$\psi_n / c^n \longrightarrow \infty, \qquad n \longrightarrow \infty,$$

then

$$\chi_\delta(n) \sim C\psi_n, \qquad n \longrightarrow \infty, \tag{3.14}$$

where $C \in (0,\infty)$ and χ_δ as defined in Theorem 3.5.

Remark 3.8 (Memory of ARCH(∞) processes)

1. *ARCH(∞) models are incapable of modeling long memory (in the squares) in the sense of a non-summable ACVF. Indeed, denote by $X = (X_n)_{n\in\mathbb{Z}}$ an ARCH(∞) process. Then the condition* (3.6) *implies that the lag coefficients ψ_j are summable. Consequently, it follows from Theorem 3.7 that*

$$\sum_{j=1}^{\infty} |\mathrm{Cov}(X_{n+j}^2, X_n^2)| < \infty.$$

   This is in contrast to MA(∞) models, where it is possible to model long memory by considering non-summable MA-coefficients, see Remark 2.55.

2. *As implied by* (3.12) *and* (3.14)*, the lag coefficients ψ_j essentially determine the degree of memory of the squares of an ARCH(∞) process X. Furthermore, this degree of memory is already imparted by the condition* (3.6)*, which ensures covariance stationarity of the levels X_n but not that of the squares X_n^2.*

As covariance stationarity and long memory (in the sense of a non-summable ACVF) are not compatible, one may take into consideration to allow for an infinite second moment. This leads to the following subclass of ARCH(∞) models.

Definition 3.9 (IARCH(∞)) *A process is said to be an integrated ARCH(∞) process, if it satisfies the equations* (3.1) *with additional restriction*

$$E[\varepsilon_n^2]\sum_{i=1}^{\infty}\psi_i = 1. \tag{3.15}$$

Definition 3.10 *Following Davidson [2004], we refer to the product $E[\varepsilon_n^2]\sum_{i=1}^{\infty}\psi_i$ as* amplitude*.*

Remark 3.11 *Remark 3.4 implies that any strictly stationary IARCH(∞) process has an infinite second moment.*

Although no weakly stationary IARCH(∞) exists, conditions for strict stationarity can be established. Recall therefore the corresponding condition (3.3) for ARCH(∞) models. Douc et al. [2008] used this result to prove the existence of IARCH(∞) processes.

Corollary 3.12 (Strict stationarity of IARCH(∞)) *Consider an IARCH(∞) model as given by Definition 3.9 and assume $E[\varepsilon_n^2] = 1$. Then the condition* (3.3) *is satisfied, if and only if there is a $p \in (0, 1]$ such that $\sum_{i=1}^{\infty}\psi_i^p < \infty$ and*

$$\sum_{i=1}^{\infty}\psi_i\log(\psi_i) + E[\varepsilon_n^2\log(\varepsilon_n^2)] \in (0, \infty]. \tag{3.16}$$

In this case, the strict stationary solution of the model is given by (3.4) *and satisfies $E[|X_n^q|] < \infty$ for $q \in [0, 2)$ and $E[X_n^2] = \infty$.*

Proof. See [Douc et al., 2008, Corollary 2]. □

The second result, which is available in the literature, stems from Kazakevicius and Leipus [2003]. They weakened the conditions on the sequence $(\varepsilon_n)_{n\in\mathbb{N}}$, but asssumed the lag coefficients to decay at an exponential rate. This suggests that conditions ensuring strict stationarity represent a compromise between conditions on the driving noise and summability conditions on the lag coefficients.

Proposition 3.13 (Strict stationarity of IARCH(∞)) *Assume that ε_n^2 is nondegenerate, $E[|\log\varepsilon_n^2|] < \infty$ and that there is a $q > 1$ such that*

$$\sum_{i=1}^{\infty}\psi_i q^i < \infty.$$

Then, the IARCH(∞) model, as given by Definition 3.9, has a strictly stationary and causal solution given by (3.4).

Proof. See [Kazakevicius and Leipus, 2003, Theorem 2.1]. □

In the next section we will consider another important ARCH(∞) model, namely the well-known GARCH model.

3.1.2 GARCH(p,q) processes

We now turn to the most popular model in financial econometrics, namely the univariate GARCH model, originally introduced by Bollerslev [1986]. Before we give the usual GARCH(p,q) representation, we define it as an ARCH(∞) process.

Definition 3.14 (GARCH(p,q)) *For $p \in \mathbb{N}_0$ and $q \in \mathbb{N}$ let α and β denote two polynomials given by*

$$\alpha(z) = \sum_{j=1}^{q} \alpha_j z^j, \qquad \beta(z) = \sum_{j=1}^{p} \beta_j z^j, \qquad z \in \mathbb{C}, \tag{3.17}$$

with $1-\beta(z) \neq 0, |z| \leq 1$. Then $(X_n)_{n\in\mathbb{Z}}$ is called GARCH(p,q) process, if it is an ARCH(∞) process according to (3.1) *with lag coefficients $\psi_i, i \in \mathbb{N}$, defined by the equation*

$$\sum_{i=1}^{\infty} \psi_i z^i = \frac{\alpha(z)}{1-\beta(z)}, \qquad z \in \mathbb{C}, |z| \leq 1. \tag{3.18}$$

Remark 3.15 (Standard and ARMA-type representation) *1. To obtain the representation, which is common in the literature, let $(X_n)_{n\in\mathbb{Z}}$ be a GARCH(p,q) process as in Definition 3.14. Consequently, the volatility process satisfies*

$$\begin{aligned}\sigma_n^2 &= \omega + \sum_{i=1}^{\infty} \psi_i X_{n-i}^2 \\ &= \omega + \left(\sum_{i=1}^{\infty} \psi_i L^i\right) X_n^2 \\ &= \omega + \frac{\alpha(L)}{1-\beta(L)} X_n^2.\end{aligned}$$

Now, applying the operator $(1-\beta(L))$ on both sides, we arrive at the usual representation

$$\begin{aligned}X_n &= \sigma_n \varepsilon_n, \\ \sigma_n^2 &= \tilde{\omega} + \alpha(L) X_n^2 + \beta(L)\sigma_n^2, \qquad n \in \mathbb{Z},\end{aligned} \tag{3.19}$$

with $\tilde{\omega} = \omega(1-\beta(1))$.

2. Assume that $X = (X_n)_{n\in\mathbb{N}}$ denotes a GARCH(p,q) process such that the squared process X^2 is covariance stationary. Then it is straightforward to see that

$$\begin{aligned}\nu_n :&= X_n^2 - E[X_n^2|\mathcal{F}_{n-1}] \\ &= X_n^2 - \sigma_n^2\, E[\varepsilon_n^2], \qquad n \in \mathbb{Z},\end{aligned}$$

defines a white noise sequence, i.e. it is uncorrelated with zero mean and finite variance. Consequently, substituting $\sigma_n^2 = (X_n^2 - \nu_n)/E[\varepsilon_n^2]$ in (3.19) *we obtain the ARMA representation*

$$\left[1 - \left(E[\varepsilon_n^2]\, \alpha(L) + \beta(L)\right)\right] X_n^2 = \omega^* + (1-\beta(L))\nu_n \qquad n \in \mathbb{Z}. \tag{3.20}$$

Being a particular ARCH(∞) model, for a GARCH(p, q) process the necessary and sufficient condition for weak stationarity (3.6) turns into

$$E[\varepsilon_n^2] \sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1. \tag{3.21}$$

As we have seen in Theorem 3.3, the condition (3.21) also implies the existence of a strictly stationary solution to the GARCH(p, q), however, this condition is not necessary. But unlike for general ARCH(∞) models, a necessary and sufficient condition for strict stationarity is available in literature for GARCH models. This can be found, for example, in [Bougerol and Picard, 1992, Theorem 2.5].

Finally, notice that analogously to the IARCH(∞) model we can define an integrated GARCH(p, q) model.

Definition 3.16 *A process $(X_n)_{n\in\mathbb{Z}}$ is called integrated GARCH(p, q), shortly IGARCH(p, q), process, if it is GARCH(p, q) according to Definition 3.14 with additional amplitude restriction*

$$E[\epsilon_n^2] \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j = 1 \tag{3.22}$$

Originally, the observation that when the GARCH model is applied to financial data the estimates for the model parameters $\alpha_i, i = 1, \dots, q$ and $\beta_j, j = 1, \dots, p$ often sum up to a value close to unity - the so called *IGARCH-effect* - led Engle and Bollerslev [1986] to introduce this model.

Remark 3.17 *Note that we can establish a* purely formal *analogy between integrated ARMA and integrated GARCH models. Although being meaningless (as no second moment exists), consider therefore the ARMA-type representation* (3.19) *of an IGARCH model. Then as ARIMA models, its autoregressive polynomial* $1 - (E[\varepsilon_n^2]\, \alpha(z) + \beta(z))$ *has a unit root.*

3.1.3 FIGARCH(p, d, q) processes

As in the ARMA case, the lag coefficients ψ_j in (3.18) of the GARCH model decay exponentially, resulting in short memory. In order to allow long range dependence, Baillie et al. [1996] introduced the class of *fractionally integrated* GARCH (FIGARCH) models, intending to mimic the successful generalization of ARMA to ARFIMA models. Although they defined FIGARCH processes by some ARMA-type representation, we follow Douc et al. [2008] embedding it into the class of ARCH(∞) models.

The structure of this section is as follows. Firstly, we will present the definition and give the proof for the existence. Subsequently, we discuss the model structure and memory properties. In doing so, we will be particularly concerned with its connection to ARFIMA models.

3.1.3.1 Definition and existence

As mentioned above, we follow Douc et al. [2008] by defining the FIGARCH as particular ARCH(∞) model.

Definition 3.18 (FIGARCH(p, d, q)) *Let $d \in (0, 1)$ and $p, q \in \mathbb{N}$. Denote by $\alpha(z), \beta(z)$ two polynomials*

$$\begin{aligned} \alpha(z) &= 1 + \alpha_1 z + \alpha_2 z^2 + \cdots + \alpha_p z^p, \qquad \text{and} \\ \beta(z) &= 1 + \beta_1 z + \beta_2 z^2 + \cdots + \beta_q z^q, \qquad z \in \mathbb{Z}. \end{aligned} \tag{3.23}$$

where $\beta(z) \neq 0, |z| \leq 1$. Then $(X_n)_{n\in\mathbb{Z}}$ is said to be a fractionally integrated GARCH(p, q) process, or FIGARCH(p, d, q) process, if it is an ARCH(∞) process (3.1) *with $E[\epsilon_n^2] = 1$ and where the lag coefficients $(\psi_n)_{n\in\mathbb{N}}$ of the squared volatility process*

$$\sigma_n^2 = \omega + \sum_{j=1}^{\infty} \psi_j X_{n-j}^2, \qquad n \in \mathbb{Z}, \tag{3.24}$$

are defined by

$$\psi(z) := \sum_{j=1}^{\infty} \psi_j z^j = 1 - (1 - z)^d \frac{\alpha(z)}{\beta(z)}, \qquad |z| \leq 1. \tag{3.25}$$

Remark 3.19 (Non-negativity of variance process) *In order to ensure the FIGARCH(p, d, q) model as given by Definition 3.18 to be well-defined, conditions need to be imposed on the coefficients of the polynomials $\alpha(z), \beta(z)$ in* (3.23) *to ensure non-negativity of the variance process* (3.24). *Conrad and Haag [2006] provide necessary and sufficient conditions for FIGARCH models of orders $q \leq 2$ as well as sufficient conditions for the general case.*

In particular, for the FIGARCH$(0, d, 1)$ model, where the lag polynomial $\psi(z)$ due to (3.25) *is given by*

$$\psi(z) = 1 - (1 - z)^d \frac{1}{1 + \beta_1 z}, \qquad |z| \leq 1, \tag{3.26}$$

the variance process is non-negative (see [Conrad and Haag, 2006, Corollary 3]) if and only if

$$-d \leq \beta_1 \leq \frac{\sqrt{2(2 - d)} - d}{2}. \tag{3.27}$$

For the FIGARCH$(1, d, 0)$ process (3.25) *turns into*

$$\psi(z) = 1 - (1 - z)^d (1 + \alpha_1 z), \qquad |z| \leq 1,$$

such that due to [Conrad and Haag, 2006, Corollary 2] σ_n^2 is well-defined if and only if

$$\frac{d - 1}{2} \leq \alpha_1 \leq d. \tag{3.28}$$

We now turn to the question of the existence of a (strictly) stationary solution of the FIGARCH model. When introducing the model, Baillie et al. [1996] claimed that the "coefficients ψ_j in the infinite lag polynomial $\psi(z)$ [in (3.25)] are dominated by the

lag coefficients in the ARCH(∞) representation of an appropriately defined high-order IGARCH model". Therefore the stationarity of FIGARCH would be implied by the stationarity of the IGARCH model.
Now, observe that the lag coefficients ψ_j in (3.25) analogously to the ARFIMA case (see Remark 2.55) satisfy for $C \neq 0$

$$\psi_j \sim C j^{-1-d}, \qquad j \longrightarrow \infty \qquad d \in (0,1). \tag{3.29}$$

On the other side, the lag coefficients (3.18) in the GARCH case decay exponentially, which holds, being a special GARCH model, in particular for an IGARCH process. The question therefore arises, how to bound coefficients that decline hyperbolically by exponentially decaying ones. Yet, in the literature there seems to be general agreement that the argumentation of Baillie et al. [1996] is incorrect, see for example Douc et al. [2008] and Mikosch and Starica [2003].

Subsequently, the "FIGARCH problem" remained unsolved for a long time, until Douc et al. [2008] managed to show the existence of FIGARCH processes. More precisely, the authors were able to prove the existence of strictly stationary FIGARCH$(0, d, 0)$ processes. In doing so, they used the fact, that any FIGARCH process belongs to the IARCH(∞) class, as the amplitude restriction (3.15),

$$\begin{aligned} E[\epsilon_n^2] \sum_{j=1}^{\infty} \psi_j &= \left[1 - (1-z)^d \frac{\alpha(z)}{\beta(z)}\right]_{z=1} \\ &= 1, \end{aligned} \tag{3.30}$$

is obviously satisfied.

In the following we present their result and demonstrate that their approach can also be applied to obtain the existence of higher order FIGARCH models.

Theorem 3.20 (Existence of FIGARCH) *Consider the FIGARCH(p, d, q) model as given by Definition 3.18 and assume the squared noise $\epsilon_n^2, n \in \mathbb{N}$, to be non-degenerate.*

(a) There exists a $d^* \in (0,1)$ such that for all $d \in (d^*, 1)$ the FIGARCH$(0, d, 0)$ model has a strictly stationary solution.

(b) There exist $d^* \in (0,1)$ and $\alpha^* > 0$ such that the FIGARCH$(1, d, 0)$ model with lag polynomial

$$\psi(z) = 1 - (1-z)^d (1 + \alpha_1 z), \qquad |z| \leq 1, \tag{3.31}$$

has a strictly stationary solution for all $d \in (d^*, 1)$ and

$$\alpha_1 \in \big(0, \min(\alpha^*, d)\big) \cup \big(1 - \alpha^*, d\big], \tag{3.32}$$

where in case that $1 - \alpha^* \geq d$ we set $(1 - \alpha^*, d] := \emptyset$.

(c) There exist $d^* \in (0,1)$ and $\beta^* > 0$ such that the FIGARCH$(0, d, 1)$ model with lag polynomial

$$\psi(z) = 1 - \frac{(1-z)^d}{1 + \beta_1 z}, \qquad |z| \leq 1, \tag{3.33}$$

has a strictly stationary solution for all $d \in (d^*, 1)$ *and*

$$\beta_1 \in \big(\max(-\beta^*, -d), 0\big). \tag{3.34}$$

The solution $(X_n)_{n\in\mathbb{Z}}$ *in each of these three cases is given by* (3.4) *and satisfies* $E[|X_n^s|] < \infty$ *for* $s \in (0, 2)$.

Proof. Recall that for the FIGARCH(p, d, q) model (see Definition 3.18) the lag polynomial $\psi(z)$ is defined by

$$\begin{aligned}\psi(z) &= \sum_{j=1}^{\infty} \psi_j z^j \\ &= 1 - (1-z)^d \frac{\alpha(z)}{\beta(z)}, \qquad |z| \le 1, d \in (0,1).\end{aligned} \tag{3.35}$$

To stress the dependence of the coefficients ψ_j on d, in the following we write $\psi_j(d)$. Moreover, we set

$$\begin{aligned}\lambda(z) \; &= \sum_{j=0}^{\infty} \lambda_j z^j \\ &:= \frac{\alpha(z)}{\beta(z)}, \qquad |z| \le 1.\end{aligned} \tag{3.36}$$

Douc et al. [2008] introduced the function H defined for $d \in (0, 1)$ by

$$H(s, d) = \log \left(\sum_{j=1}^{\infty} \psi_j^s(d) \right). \tag{3.37}$$

As outlined above, the asymptotic behavior of the lag coefficients $\psi_j(d), j \in \mathbb{N}$, is given by $\psi_j(d) \sim \mathcal{O}(j^{-1-d})$, such that

$$\sum_{j=1}^{\infty} \psi_j^s(d) < \infty, \qquad d \in (0,1),$$

and $H(\cdot, d)$ is well-defined as long as $s \in (\frac{1}{d+1}, 1]$. Furthermore, observe that $H(\cdot, d)$ is convex. To see this, let $\delta \in (0, 1)$ and define the parameters $r_1 := \frac{1}{\delta}$ and $r_2 := \frac{1}{1-\delta}$ such that $r_1, r_2 \ge 1$ and

$$\frac{1}{r_1} + \frac{1}{r_2} = 1.$$

Now, for $s_1, s_2 \in (\frac{1}{1+d}, 1]$ we apply the Hölder inequality and obtain

$$\begin{aligned}
H\big(\delta s_1 + (1-\delta)s_2, d\big) &= \log\left(\sum_{j=1}^{\infty} \psi_j(d)^{\delta s_1} \psi_j(d)^{(1-\delta)s_2}\right) \\
&\le \log\left(\left(\sum_{j=1}^{\infty} (\psi_j(d)^{\delta s_1})^{r_1}\right)^{\frac{1}{r_1}} \left(\sum_{j=1}^{\infty} (\psi_j(d)^{(1-\delta)s_2})^{r_2}\right)^{\frac{1}{r_2}}\right) \\
&= \delta \log\left(\sum_{j=1}^{\infty} \psi_j(d)^{s_1}\right) + (1-\delta)\log\left(\sum_{j=1}^{\infty} \psi_j(d)^{s_2}\right) \\
&= \delta\ H(s_1, d) + (1-\delta)\ H(s_2, d).
\end{aligned}$$

We further have due to the unit amplitude (see (3.30))

$$\begin{aligned}
H(1, d) &= \log\left(\sum_{j=1}^{\infty} \psi_j(d)\right) \\
&= 0,
\end{aligned}$$

while for $s \downarrow \frac{1}{1+d}$ the lag coefficients $\psi_j^s(d)$, having rate of decay $\psi_j^s(d) \sim \mathcal{O}(j^{(-1-d)s})$, approach non-summability, such that for all $d \in (0,1)$ it holds

$$H(s, d) \to \infty, \qquad \text{for } s \downarrow \frac{1}{1+d}.$$

Consequently, the above derived convexity implies that $H(\cdot, d)$ is decreasing. Further the derivative of H with respect to s is given by

$$\frac{\partial}{\partial s} H(s, d) = \frac{\sum_{j=1}^{\infty} e^{s\log(\psi_j)} \log(\psi_j)}{\sum_{j=1}^{\infty} \psi_j^s}.$$

In particular, the left-sided derivative $L(d)$ at $s = 1$ satisfies

$$\begin{aligned}
L(d) &:= \lim_{s \uparrow 1} \frac{H(s, d) - H(1, d)}{s - 1} \\
&= \sum_{j=1}^{\infty} \psi_j(d) \log(\psi_j(d)),
\end{aligned}$$

which corresponds to the sum on the left of condition (3.16). As $H(\cdot, d)$ is decreasing and convex, its derivative $\frac{\partial}{\partial s} H(\cdot, d)$ is negative and increasing, such that we end up with the following bounds for $L(d)$,

$$0 \le -L(d) \le -\frac{H(s, d)}{s - 1}, \qquad \text{for } d \in (0,1) \text{ and } s \in \big(1/(1+d), 1\big). \tag{3.38}$$

We now show that the upper bound in (3.38) independently of the parameter s can get arbitrarily small. Using the notation in (3.35) and (3.36) the lag coefficients $\psi_j(d)$ of the FIGARCH model are given by the convolution

$$\psi_j(d) = \sum_{k=0}^{j} -\pi_k(d)\lambda_{j-k}, \qquad j \ge 1, \tag{3.39}$$

where the coefficients $\pi_k(d), k \in \mathbb{N}_0$, are defined as in (2.42) and satisfy according to [Brockwell and Davis, 1991, p. 520] $\pi_0 = 1, \pi_1 = -d$ and

$$\pi_j = \frac{-d(1-d)\cdot\ldots\cdot(j-1-d)}{j!}, \qquad j \geq 2. \tag{3.40}$$

In particular, for $d \uparrow 1$ we have that $\pi_1(d) \to -1$ and $\pi_j(d) \to 0, j \geq 2$, such that

$$\lim_{d\uparrow 1} \psi_j(d) = \lambda_{j-1} - \lambda_j, \qquad j \geq 1.$$

Further note that for $j \geq 0$ the quotient $\frac{|\pi_j(d)|}{d}$ is non-increasing in $d \in (0,1)$, which will be used in the following.

In order to apply the *dominated convergence theorem*, observe that for all $d \in [0.5, 1)$

$$\begin{aligned}
|\psi_j(d)| &\leq \sum_{k=0}^{j} |\pi_k(d)||\lambda_{j-k}| \\
&\leq \sum_{k=0}^{j} \left|\frac{\pi_k(\frac{1}{2})}{\frac{1}{2}}\right| |\lambda_{j-k}| \\
&= 2\sum_{k=0}^{j} |\pi_k(1/2)||\lambda_{j-k}| \\
&:= b_j.
\end{aligned}$$

Since $|\pi_j(1/2)| \sim C\, j^{-1.5}$, the sequence $(b_j)_{j\in\mathbb{N}}$ is summable, such that we have found an upper bound for $(\psi_j(d))_{j\in\mathbb{N}}$, that is, for all $d \in [0.5, 1)$

$$\sum_{j=1}^{\infty} \psi_j^s(d) \leq \sum_{j=1}^{\infty} b_j^s < \infty, \qquad s \in \big(1/(1+d), 1\big). \tag{3.41}$$

Thus, we can apply the dominated convergence theorem and obtain

$$\lim_{d\uparrow 1} H(s,d) = \log\left(\sum_{j=1}^{\infty} (\lambda_{j-1} - \lambda_j)^s\right), \qquad s \in (2/3, 1). \tag{3.42}$$

Summing up, we have found that

$$0 \leq \limsup_{d\uparrow 1} -L(d) \leq \frac{\log\left(\sum_{j=1}^{\infty} (\lambda_{j-1} - \lambda_j)^s\right)}{1-s}, \qquad s \in (2/3, 1). \tag{3.43}$$

In order to prove part (i), consider the lag polynomial of the FIGARCH$(0,d,0)$ model, i.e.

$$\psi(z) = 1 - (1-z)^d, \qquad |z| \leq 1.$$

Obviously, the coefficients λ_j as defined in (3.36) satisfy

$$\lambda_0 = 1, \qquad \lambda_j = 0, j \geq 1,$$

such that (3.43) turns into

$$0 \leq \limsup_{d \uparrow 1} -L(d) \leq \frac{\log(1)}{1-s} = 0. \tag{3.44}$$

Now, using the convexity of the mapping $x \mapsto x\log(x)$, Jensen's inequality implies

$$E[\epsilon_n^2 \log(\epsilon_n^2)] > E[\epsilon_n^2] \log(E[\epsilon_n^2]) = 0,$$

where the strict inequality is due to ϵ_n^2 being non-degenerate. As $L(d) \to 0$ for $d \uparrow 1$ (see (3.44)), there exists a $d^* \in (0,1)$ with

$$E[\epsilon_n^2 \log(\epsilon_n^2)] + L(d) > 0, \qquad \text{for all } d \in (d^*, 1),$$

and part (i) follows from Corollary 3.12.

For higher order FIGARCH(p,d,q) models, i.e. where not both p and q are equal to zero, additional conditions need to be imposed on the coefficients of the polynomials $\alpha(z), \beta(z)$ in (3.35), which will be illustrated when proving the parts (b) and (c). In the former case the lag polynomial $\psi(z)$ has the representation (3.31), implying that

$$\lambda_0 = 1, \lambda_1 = \alpha_1 \qquad \text{and } \lambda_j = 0, j \geq 2.$$

Let $\zeta > 0$ be arbitrary. Then from (3.43) it follows that there exists a $d^* \in (0,1)$ such that for all $d \in (d^*, 1)$

$$0 \leq -L(d) < \frac{\log((1-\alpha_1)^s + \alpha_1^s)}{1-s} + \zeta, \qquad s \in (2/3, 1), \alpha_1 > 0. \tag{3.45}$$

Obviously, the quotient on the right-hand side tends to zero for $\alpha_1 \to 0$ as well as for $\alpha_1 \to 1$. In particular, we can find an $\alpha^* > 0$ such that for all $\alpha_1 \in \mathcal{D}_{\alpha_1} := (0, \alpha^*) \cup (1-\alpha^*, 1]$

$$0 \leq -L(d) < 2\zeta.$$

Taking the condition (3.28) for the non-negativity of the variance process into account, the range $\mathcal{D}_{\alpha_1}$ of admissible values for α_1 reduces to (3.32). Now, since $\zeta > 0$ was assumed to be arbitrary, we set

$$\zeta := \frac{E[\epsilon_n^2 \log(\epsilon_n^2)]}{2} > 0$$

and therefore obtain the claim by Corollary 3.12.

On the other hand, in the case of the FIGARCH$(0,d,1)$ model with

$$\lambda(z) = \frac{1}{1+\beta_1 z}, \qquad |z| \leq 1,$$

we have $\lambda_j = (-\beta_1)^j, j \geq 0$. Again for arbitrary $\zeta > 0$ there exists a $d^* \in (0,1)$ such that for all $d \in (d^*, 1)$

$$\begin{aligned}
0 \leq -L(d) &< \frac{\log\left(\sum_{j=1}^{\infty}\left((-\beta_1)^{j-1} - (-\beta_1)^j\right)^s\right)}{1-s} + \zeta \\
&= \frac{\log\left(\left(1-(-\beta_1)\right)^s \sum_{j=1}^{\infty}\left((-\beta_1)^{j-1}\right)^s\right)}{1-s} + \zeta \\
&= \frac{\log\left((1+\beta_1)^s\right) + \log\left(\sum_{j=1}^{\infty}\left((-\beta_1)^{j-1}\right)^s\right)}{1-s} + \zeta \\
&= \frac{\log\left((1+\beta_1)^s\right) - \log(1-(-\beta_1)^s)}{1-s} + \zeta, \qquad s \in (2/3, 1), \beta_1 \in (-1, 0).
\end{aligned}$$

Analogously to the previous case the quotient on the right-hand side tends to zero for $\beta_1 \uparrow 0$, that means there exists a $\beta^* > 0$ such that for all $\beta_1 \in \mathcal{D}_\beta = (-\beta^*, 0)$

$$0 \leq -L(d) < 2\zeta.$$

Taking (3.28) into account, the range of possible values for β_1 is given by (3.34) and the result follows with Corollary 3.12. □

Finally, observe that the FIGARCH model comprises both the GARCH and the IGARCH model. We summarize this result in the following remark.

Remark 3.21 (GARCH as subclass of FIGARCH) *Consider a GARCH model as defined in Definition 3.14 with lag polynomial* $\psi(z) = \frac{\alpha(z)}{1-\beta(z)}$.

1. *Let* $d = 0$ *and define the following polynomials,*

$$\begin{aligned}
\alpha^*(z) &= 1 - \beta(z) - \alpha(z), \\
\beta^*(z) &= 1 - \beta(z).
\end{aligned}$$

   Then the above GARCH can be rewritten as FIGARCH model by defining its lag polynomial $\tilde{\psi}$ *as*

$$\begin{aligned}
\tilde{\psi}(z) &= 1 - (1-z)^d \frac{\alpha^*(z)}{\beta^*(z)} \\
&= 1 - \frac{1-\beta(z)-\alpha(z)}{1-\beta(z)} \\
&= \frac{\alpha(z)}{1-\beta(z)}.
\end{aligned}$$

2. *Now, we consider the corresponding IGARCH model, i.e. we assume that* $\alpha(1) + \beta(1) = 1$. *Then there exists a polynomial* λ *such that* $1 - \alpha(z) - \beta(z) = \lambda(z)(1 - z)$. *Consequently, this IGARCH model may be represented as FIGARCH with lag polynomial*

$$\psi(z) = 1 - (1-z)^d \frac{\lambda(z)}{1-\beta(z)},$$

   and $d = 1$.

3.1.3.2 Discussion

Now that the existence is ensured, we want to gain a deeper understanding of the structure of the FIGARCH model and analyze if and in what way it improves the memory properties of the GARCH model. Concerning the first issue, it is worthwile to revise its relationship to the class of ARFIMA models, as this was the starting point of Baillie et al. [1996], when defining the model.

Relationship between FIGARCH and ARFIMA. An ARMA(p, q) process $X = (X_n)_{n\in\mathbb{Z}}$ may be represented as

$$X_n = \frac{\theta(L)}{\phi(L)}\varepsilon_n, \qquad n \in \mathbb{Z},$$

where ϕ, θ denote finite lag polynomials and $\varepsilon = (\varepsilon_n)_{n\in\mathbb{Z}}$ denotes a white noise sequence. In particular,

$$X_n = \psi_0\varepsilon_n + \psi_1\varepsilon_{n-1} + \psi_2\varepsilon_{n-2} + \dots, \qquad n \in \mathbb{Z},$$

with exponentially decaying lag coefficients $\psi_j, j \in \mathbb{N}$. Note that these can be interpreted as being measures of sensitivity to changes in the innovation sequence ε. More precisely, ψ_j gives the impact of a unit innovation ε_{n-j} in period $n - j$ on the level of X in period n. For this reason the sequence $(\psi_j)_{j\in\mathbb{N}}$ is also referred to as *impulse response coefficients* (IRC). In this sense shocks to the process X dissipate at an exponential rate.

In contrast, assume now X to denote the corresponding non-stationary ARIMA(p, d, q) process with $d = 1$, i.e.

$$\begin{aligned}(1 - L)X_n &= \frac{\theta(L)}{\phi(L)}\varepsilon_n, \\ &= \psi_0\varepsilon_n + \psi_1\varepsilon_{n-1} + \psi_2\varepsilon_{n-2} + \dots, \qquad n \in \mathbb{Z},\end{aligned} \tag{3.46}$$

with $\theta(1) \neq 0$. We write its MA(∞)-representation as

$$X_n = \tilde{\psi}_0\varepsilon_n + \tilde{\psi}_1\varepsilon_{n-1} + \tilde{\psi}_2\varepsilon_{n-2} + \dots, \qquad n \in \mathbb{Z}.$$

Now, the impact of a unit innovation in period $n - j$ on the level X_n is given by $\tilde{\psi}_j = \sum_{k=0}^{j} \psi_k$, such that

$$\lim_{j\to\infty} \tilde{\psi}_j = \sum_{k=0}^{\infty} \psi_k = \frac{\theta(1)}{\phi(1)} \neq 0.$$

The interpretation is that shocks to the ARIMA(p, d, q) process do not dissipate, but persist indefinitely. Finally, allowing in (3.46) for fractional values $d \in (0, 0.5)$, leads to the *long-memory* ARFIMA(p, d, q) model with representation

$$(1 - L)^d X_n = \frac{\theta(L)}{\phi(L)}\varepsilon_n.$$

We have seen that the corresponding MA(∞) coefficients $\tilde{\psi}_j^d$ satisfy $\tilde{\psi}_j^d \sim C\ j^{d-1}$, i.e. shocks dissipate by a slow hyperbolic rate. Thus, from this point of view the ARFIMA model fills the gap between short memory or short persistence on the one side and complete persistence on the other side.

Now, we turn to the ARCH(∞) framework. The ARMA-type representation of the GARCH model is given by

$$(1-\alpha(L)-\beta(L))X_n^2 = \omega + (1-\beta(L))\nu_n, \qquad n \in \mathbb{Z}, \tag{3.47}$$

with notation as in Remark 3.15. This indeed defines an ARMA process if the squared GARCH X^2 is covariance stationary. The corresponding IGARCH is obtained by additionally imposing the condition $\alpha(1)+\beta(1)=1$, such that the autoregressive polynomial in (3.47) has a unit root,

$$\lambda(L)(1-L)X_n^2 = \omega + (1-\beta(L))\nu_n, \qquad n \in \mathbb{Z}. \tag{3.48}$$

with $\lambda(L)$ appropriately defined. Baillie et al. [1996] take this representation as a reason for considering the IGARCH as analogue to the ARIMA model. As a consequence, they mimic the generalization of ARMA to ARFIMA by allowing fractional differences in (3.48), thus defining the FIGARCH model by the difference equation

$$\lambda(L)(1-L)^d X_n^2 = \omega + (1-\beta(L))\nu_n, \qquad n \in \mathbb{Z}. \tag{3.49}$$

Precisely this approach is criticized in the literature. Both Davidson [2004] and Mikosch and Starica [2003] point out that the above drawn connection to the MA(∞) class is purely formal. The crucial point is the non-stationarity of the IGARCH model. Clearly, one may represent the IGARCH model using the ARMA-type representation (3.48), yet it is not meaningful as the driving process ν_n is not white noise.

A further point criticized by Mikosch and Starica [2003] concerns the defining equations (3.49), as the driving process $\nu_n = X_n^2 - \sigma_n^2$ depends on the process, which is being defined. This is why in the literature the FIGARCH model is usually defined as ARCH(∞) process and only afterwards (having ensured its existence) the equivalent representation (3.49) is derived.

Memory properties. When introducing the model, Baillie et al. [1996] justified their view of the FIGARCH as intermediate case (concerning memory) between GARCH and IGARCH by comparing their IRCs. See above, where we proceeded analogously when comparing the memory properties of ARMA and ARFIMA. Due to the ARMA-type representations (3.47)-(3.49) it is straightforward to see that the IRC of GARCH and FIGARCH decline exponentially and hyperbolically while those of the IGARCH do not even tend to zero. Thus Baillie et al. [1996] interpreted the memory properties analogously to the ARFIMA case. Again the question arises, whether it is reasonable to apply impulse response analysis, where the possibly correlated sequence ν_n is interpreted as innovation to the process.

For a general (weakly stationary) ARCH(∞) model we saw that the rate of convergence of the lag coefficients essentially determine the dependence structure of the process. Following Davidson [2004] we therefore assess in the following by means of this rate the memory properties of the FIGARCH. Recall that the FIGARCH(p, d, q) model is a particular ARCH(∞) model with lag polynomial given by

$$\psi(z) = \sum_{j=1}^{\infty} \psi_j z^j = 1 - (1-L)^d \frac{\alpha(L)}{\beta(L)}, \qquad d \in [0,1],$$

with α, β given by (3.25). The case $d = 1$ corresponds to the IGARCH model, which in the above sense has exponential memory. For $d \in (0,1)$ we saw that the lag coefficients ψ_j decline hyperbolically, namely

$$\psi_j \sim C j^{-1-d}.$$

Firstly and quite counterintuitively, for decreasing d the memory increases. Particularly for $d \to 0$, the lag coefficients approach non-summability. Yet, due to the amplitude restriction

$$\psi(1) = \sum_{j=1}^{\infty} \psi_j = 1,$$

while approaching non-summability, the individual ψ_j all tend to zero. In the limit $d = 0$, however, when obtaining the GARCH, we again have exponential memory. Davidson [2004] therefore concludes that "*the characterization of the FIGARCH model as an intermediate case between the stable GARCH and the IGARCH [...] is misleading. In fact, it has more memory than either of these models but behaves oddly owing to the rather arbitrary restriction of holding the amplitude to one.*"

3.2 Continuous Time Models

3.2.1 COGARCH processes

Referring to Brockwell et al. [2006], we now introduce the continuous time analogue of the discrete GARCH(p, q) model. Before giving its precise definition, we take a closer look at the idea the continuous approach is based on.

We start with a discrete GARCH(p, q) process X. Denote by σ_0^2 the initial value of its volatility process, such that X is given by

$$\begin{aligned} X_n &= \sigma_n \varepsilon_n, \\ \sigma_n^2 &= \omega + \alpha(L) X_n^2 + \beta(L) \sigma_n^2, \qquad n \geq r, \end{aligned}$$

where $r := \max(p, q), \alpha(z) = \sum_{j=1}^{q} \alpha_j z^j$ and $\beta(z) = \sum_{j=1}^{p} \beta_j z^j$, $(\varepsilon_n)_{n \in \mathbb{N}_0}$ denotes the driving i.i.d. noise sequence and $\sigma_0^2, \ldots, \sigma_{r-1}^2$ are i.i.d. and independent of $(\varepsilon_n)_{n \in \mathbb{N}_0}$. In particular, $(\sigma_n^2)_{n \in \mathbb{N}_0}$ fulfills

$$\sigma_n^2 (1 - \beta(L)) = \omega + (\varepsilon_n^2 \sigma_n^2) \alpha(L), \qquad n \geq r. \tag{3.50}$$

Brockwell et al. [2006] conclude that the squared volatility process can be seen as a "self-exciting" ARMA$(p, q-1)$ process, the term *self-exciting* emphasizing the particular relationship between $(\sigma_n^2)_{n\in\mathbb{N}}$ and its driving noise $(\varepsilon_n^2\sigma_n^2)_{n\in\mathbb{N}}$. Consequently, the basic idea consists in modeling σ_n^2 as CARMA process, with appropriately defined driving noise.

What remains is the question how to construct a continuous time noise process, which is equivalent to the discrete one given by $(\varepsilon_n^2\sigma_n^2)_{n\in\mathbb{N}}$. When introducing the COGARCH$(1,1)$ model, Klüppelberg et al. [2004] substituted the discrete innovations ε_n by the jumps of a Lévy process $(L_t)_{t\geq 0}$. Applying this idea to the *integrated* discrete noise

$$R_n^{(d)} = \sum_{k=0}^{n-1} \varepsilon_k^2\sigma_k^2, \qquad n \geq 1,$$

Brockwell et al. [2006] defined the continuous-time analogue by

$$\begin{aligned} R_t &= \sum_{s\leq t} \sigma_{s-}^2 (\Delta L_s)^2 \\ &= \int_{(0,t]} \sigma_{s-}^2 d[L,L]_s^{(dis)}, \qquad t \geq 0, \end{aligned} \tag{3.51}$$

where $[L,L]_t^{(dis)} = \sum_{s\leq t}(\Delta L_s)^2$ denotes the discrete part of the quadratic variation of the Lévy process $(L_t)_{t\geq 0}$.

Defining the squared volatility process as CARMA process, where in (2.48) the Lévy noise is substituted by (3.51), we arrive at the COGARCH model.

Definition 3.22 (COGARCH(p,q)) *Let $p,q \in \mathbb{N}$ with $p \geq q > 0$. For $\alpha_0 > 0, \alpha_1, \ldots, \alpha_q \in \mathbb{R}$ and $\beta_1, \ldots, \beta_p \in \mathbb{R}$ such that $\alpha_q \neq 0, \beta_p \neq 0$ we define the right-continuous (squared) volatility process $(\sigma_t^2)_{t\geq 0}$ by*

$$\sigma_t^2 = \alpha_0 + \boldsymbol{a}'\boldsymbol{Y}_t, \qquad t \geq 0, \tag{3.52}$$

where $\boldsymbol{a}' = [\alpha_1, \ldots, \alpha_p], \alpha_i = 0, i > q$ and $(\boldsymbol{Y}_t)_{t\geq 0}$ is the unique solution of the SDE

$$d\boldsymbol{Y}_t = B\boldsymbol{Y}_{t-}dt + \boldsymbol{e}(\alpha_0 + \boldsymbol{a}'\boldsymbol{Y}_{t-})d[L,L]_t^{(dis)}, \qquad t > 0, \tag{3.53}$$

with $\boldsymbol{Y}_0$ independent of $(L_t)_{t\geq 0}$ and $B, \boldsymbol{e}$ defined as in (2.49). If $(\sigma_t^2)_{t\geq 0}$ is strictly stationary and a.s. nonnegative, then the process $(G_t)_{t\geq 0}$ defined by $G_0 = 0$ a.s. and

$$dG_t = \sigma_{t-}dL_t, \qquad t > 0, \tag{3.54}$$

is called COGARCH(p,q) process.

Remark 3.23 (Definition of p and q) *Notice that compared to Brockwell et al. [2006] we interchanged the definition of the model parameters p and q. This way the resulting COGARCH model is in line with the discrete GARCH model defined in Definition 3.14.*

For simplicity, in the following we consider the COGARCH$(1,1)$ model. In this case the solution of (3.52) and (3.53) can be stated explicitly.

Proposition 3.24 (COGARCH$(1,1)$) *Consider a COGARCH process $(G_t)_{t\geq 0}$ as given by Definition 3.22 and let $p,q=1$. Then G satisfies*

$$
\begin{aligned}
dG_t &= \sigma_{t-} dL_t, \qquad t>0, \text{ with } G_0 = 0,\\
\sigma_t^2 &= e^{-X_t}\bigg(\sigma_0^2 + \alpha_0\beta_1 \int_{(0,t]} e^{X_s}\, ds\bigg), \qquad t\geq 0,
\end{aligned}
\tag{3.55}
$$

where

$$
X_t = \beta_1 t - \sum_{s\leq t} \log(1+\alpha_1(\Delta L_s)^2), \qquad t \geq 0, \tag{3.56}
$$

with $\alpha_0, \alpha_1, \beta_1 > 0$.

Remark 3.25 *When originally introducing the COGARCH(1,1), the approach taken by Klüppelberg et al. [2004] was based on an explicit representation of the discrete GARCH$(1,1)$, and in particular differed from the one presented above. Nevertheless, it lead to the same defining equations* (3.55) *and* (3.56)*, such that it is consistent with Definition 3.22.*

For a proof of Proposition 3.24 we need to introduce an important class of processes.

Definition 3.26 (Generalized Ornstein-Uhlenbeck process) *Consider a bivariate Lévy process $(\xi_t, \eta_t)_{t\geq 0}$. Then the process $(V_t)_{t\geq 0}$ given by*

$$
V_t = e^{-\xi_t}\big(V_0 + \int_{(0,t]} e^{\xi_{s-}} d\eta_s\big), \qquad t\geq 0, \tag{3.57}
$$

with initial value V_0 being independent of $(\xi_t,\eta_t)_{t\geq 0}$, is called (univariate) generalized Ornstein-Uhlenbeck *process (GOU) driven by $(\xi_t,\eta_t)_{t\geq 0}$.*

Observe that the squared volatility process $(\sigma_t^2)_{t\geq 0}$ in (3.55) represents an example of a GOU process. For a general COGARCH(p,q) model, i.e. where not both p and q are equal to one, this does not hold. Yet, due to [Behme and Lindner, 2012, Example 3.6] the state process (3.53) is always given as a *multivariate* GOU process.

Proof of Proposition 3.24. First note that for $p,q=1$ the equations (3.52) and (3.53) turn into the system

$$
\begin{aligned}
\sigma_t^2 &= \alpha_0 + \alpha_1 Y_t,\\
dY_t &= -\beta_1 Y_{t-} dt + \sigma_{t-}^2 d[L,L]_t^{(dis)}, \qquad t\geq 0.
\end{aligned}
$$

Using $Y_{t-} = \frac{\sigma_{t-}^2 - \alpha_0}{\alpha_1}$ for $\alpha_1 \neq 0$ we find that (σ_t^2) satisfies the following SDE

$$
\begin{aligned}
d\sigma_t^2 &= \alpha_1 dY_t\\
&= -\beta_1\alpha_1 \frac{\sigma_{t-}^2 - \alpha_0}{\alpha_1} dt + \alpha_1 \sigma_{t-}^2 d[L,L]_t^{(dis)}\\
&= \sigma_{t-}^2 d\Big(-\beta_1 t + \alpha_1 [L,L]_t^{(dis)}\Big) + d(\alpha_0\beta_1 t)\\
&= \sigma_{t-}^2 dU_t + dM_t, \qquad t>0,
\end{aligned}
\tag{3.58}
$$

with

$$\begin{pmatrix} U_t \\ M_t \end{pmatrix} = \begin{pmatrix} -\beta_1 t + \alpha_1 [L,L]_t^{(dis)} \\ \alpha_0\beta_1 t \end{pmatrix}$$
$$= t \begin{pmatrix} -\beta_1 \\ \alpha_0\beta_1 \end{pmatrix} + \sum_{s\le t} \begin{pmatrix} \alpha_1(\Delta L_s)^2 \\ 0 \end{pmatrix}, \qquad t \ge 0. \tag{3.59}$$

According to (3.59), $(U_t, M_t)_{t\ge 0}$ is given as the sum of a linear drift and a compound Poisson process. Hence, $(U_t, M_t)_{t\ge 0}$ represents a bivariate Lévy process and we are able to apply Theorem 3.4, part *b*), of Behme and Lindner [2012]. There it is shown, that if the Lévy measure of the process U satisfies $\nu_U(\{-1\}) = 0$ (which in our case is fulfilled, as $U_t = -\beta_1 t + \sum_{s\le t} \alpha_1(\Delta L_s)^s$ has only positive jumps), then the solution of (3.58) is a GOU process driven by the following bivariate Lévy process

$$\begin{pmatrix} \xi_t \\ \eta_t \end{pmatrix} = \begin{pmatrix} -\log(\mathcal{E}(U_t)) \\ \alpha_0\beta_1 t + \big[-\log(\mathcal{E}(U_t)), \alpha_0\beta_1 t\big] \end{pmatrix}, \qquad t \ge 0, \tag{3.60}$$

where the so-called stochastic exponential $\mathcal{E}(U_t)$ is defined by

$$\mathcal{E}(U_t) = e^{U_t - \frac{1}{2}[U,U]_t^{(con)}} \prod_{s\le t} \big(1 + \Delta U_s\big) e^{-\Delta U_s}, \qquad t \ge 0. \tag{3.61}$$

Since $U_t = -\beta_1 t + \sum_{s\le t} \alpha_1(\Delta L_s)^2$ is given as the sum of a (deterministic) drift and a pure jump process, the continuous part of its quadratic variation equals zero, i.e. $[U,U]_t^{(con)} = 0$, for all $t \ge 0$. Thus we obtain

$$\xi_t = -U_t - \sum_{s\le t} \big[\log(1+\Delta U_s) - \Delta U_s\big]. \tag{3.62}$$

Note that the jumps of $(U_t)_{t\ge 0}$ are given by $\Delta U_t = \alpha_1(\Delta L_s)^2$ such that (3.62) turns into

$$\xi_t = \beta_1 t - \alpha_1 [L,L]_t^{(dis)} - \sum_{s\le t} \log(1+\alpha_1(\Delta L_s)^2) + \sum_{s\le t} \alpha_1(\Delta L_s)^2$$
$$= \beta_1 t - \sum_{s\le t} \log(1+\alpha_1(\Delta L_s)^2), \qquad t \ge 0.$$

We now consider the second component, namely $(\eta_t)_{t\ge 0}$, which is given by

$$\eta_t = \alpha_0\beta_1 t + \alpha_0\beta_1 \Big[\beta_1 t - \sum_{s\le t} \log(1+\alpha_1(\Delta L_s)^2), t\Big]$$
$$= \alpha_0\beta_1 t - \alpha_0\beta_1 \Big[\sum_{s\le t} \log(1+\alpha_1(\Delta L_s)^2), t\Big]$$
$$= \alpha_0\beta_1 t, \qquad t \ge 0.$$

Hence, we see that $\eta_t = \alpha_0\beta_1 t$. Summing up, $(\sigma_t^2)_{t\ge 0}$ is a GOU driven by

$$\begin{pmatrix} \xi_t \\ \eta_t \end{pmatrix} = \begin{pmatrix} \beta_1 t - \sum_{s\le t} \log(1+\alpha_1(\Delta L_s)^2) \\ \alpha_0\beta_1 t \end{pmatrix}, \qquad t \ge 0, \tag{3.63}$$

which yields the assertion. □

The following result ensures the existence of the COGARCH$(1,1)$ model.

Theorem 3.27 *Let $(G_t)_{t\geq 0}$ denote a COGARCH$(1,1)$ process given by (3.55) and assume that*

$$\int_{\mathbb{R}} \log(1+\alpha_1 x^2)\ \nu_L(dx) < -\log(\beta_1), \tag{3.64}$$

where ν_L denotes the Lévy measure of the driving Lévy process $(L_t)_{t\geq 0}$. Then the squared volatility process $(\sigma_t^2)_{t\geq 0}$ with initial value satisfying

$$\sigma_0^2 \stackrel{\mathscr{D}}{=} \alpha_0\beta_1 \int_0^\infty e^{-X_s}\ ds, \tag{3.65}$$

independent of $(L_t)_{t\geq 0}$, is strictly stationary. Furthermore, $(G_t)_{t\geq 0}$ has strictly stationary increments.

Proof. See [Klüppelberg et al., 2004, Theorem 3.2, Corollary 3.1]. □

As its discrete counterpart the COGARCH model is characterized by short memory. More precisely, the autocovariance function of the squared volatility (as well as the one of the squared increments of the COGARCH process itself) decays by an exponential rate. For the subsequent result we define for $r > 0$

$$G_t^{(r)} := G_{t+r} - G_t = \int_{t+}^{t+r} \sigma_s\ dL_s, \qquad t \geq 0.$$

Proposition 3.28 *Consider the COGARCH$(1,1)$ model according to (3.55), (3.56) and assume that the condition (3.64) is satisfied, such that $(\sigma_t^2)_{t\geq 0}$ with initial value (3.65) is stationary. Denote by ψ the log-Laplace transform of X_1, i.e.*

$$\psi(u) = \log\left(E[e^{-uX_1}]\right), \qquad u > 0.$$

Then we have that

$$\operatorname{Cov}(\sigma_t^2, \sigma_{t+h}^2) = \beta_1^2 \left(\frac{2}{\psi(1)\psi(2)} - \frac{1}{\psi^2(1)}\right) e^{h\psi(1)}, \qquad t, h \geq 0.$$

Assume furthermore that the driving Lévy process has zero mean and no Brownian motion component. If $E[L_1^4] < \infty$ and $\psi(2) < 0$ then for any $t \geq 0$ and $h \geq r > 0$

$$\operatorname{Cov}(G_t^{(r)}, G_{t+h}^{(r)}) = 0, \tag{3.66}$$

but

$$\operatorname{Cov}((G_t^{(r)})^2, (G_{t+h}^{(r)})^2) = \left(\frac{e^{-r\psi(1)} - 1}{-\psi(1)}\right) E[L_1^2] \operatorname{Cov}(G_r^2, \sigma_r^2) e^{h\psi(1)}. \tag{3.67}$$

Proof. See [Klüppelberg et al., 2004, Corollary 4.1, Proposition 5.1].

3.2.2 Fractionally integrated COGARCH

As seen in the last section, the COGARCH preserves the short memory property of the GARCH model. In the following we modify the volatility process σ_t^2 of the COGARCH$(1,1)$ to allow for long range dependence. In view of the fact that the COGARCH$(1,1)$ was obtained by modeling σ_t^2 as "self-exciting" CARMA$(1,0)$ process, a natural approach to incorporate long range dependence consists in assuming σ_t^2 to follow some kind of "self-exciting" FICARMA process.
We start by recalling that the squared volatility σ_t^2 of the COGARCH$(1,1)$ satisfies

$$d\sigma_t^2 = -\beta_1(\sigma_{t-}^2 - a_0)dt + \alpha_1\sigma_{t-}^2 d[L,L]_t^{(dis)}, \qquad t > 0, \tag{3.68}$$

with parameters $\alpha_0, \alpha_1, \beta_1 > 0$ and where L is a Lévy process. Whereas the solution of the CARMA$(1,0)$ model is given by an ordinary OU process, the solution of the above SDE constitutes a generalized OU process, namely

$$\sigma_t^2 = e^{-X_t}\left(\sigma_0^2 + \alpha_0\beta_1\int_0^t e^{X_s}ds\right), \qquad t \geq 0, \tag{3.69}$$

where X is defined as in (3.56). In particular, no moving-average representation of σ_t^2 is available. As a consequence the FICARMA approach of Brockwell [2004] (see Definition 2.63), which is essentially based on the moving-average representation of an OU process, cannot be applied here.
However, Marquardt [2006] showed that the (centered) FICARMA model alternatively can be obtained by substituting the driving Lévy process with the corresponding MvN-FLP (see Theorem 2.66). At first glance, this approach also seems to be inappropriate. More precisely, the driving Lévy process of σ_t^2 is given by $\left([L,L]_t^{(dis)}\right)_{t\geq 0}$, which is a subordinator, such that the corresponding MvN-FLP is not well-defined, as mentioned in Remark 2.31. In the following we show how to cope with this problem.

3.2.2.1 MG-FLP driven volatility

While the MvN-FLP driven by the subordinator $\left([L,L]_t^{(dis)}\right)_{t\geq 0}$ is not well-defined, the corresponding MG-FLP is, as long as $E[([L,L]_1^{(dis)})^2] < \infty$. This is due to the MG-kernel, which unlike the MvN-kernel is both integrable and square integrable, such that the resulting FLP is well-defined also for driving Lévy processes with non-zero expectation, see Corollary 2.11. In the following we therefore denote by L^d the FLP given by

$$L_t^d = \int_0^t f_d^{MG}(t,u)\, d[L,L]_u^{(dis)}, \qquad t \geq 0, d \in (0, 0.5). \tag{3.70}$$

Proposition 3.29 *Let $L = (L_t)_{t\geq 0}$ be a Lévy process satisfying $E[L_1^4] < \infty$. Then L_t^d as defined in (3.70) exists for all $t \geq 0$ as limit in $L^2(\Omega)$-sense.*

Proof. Clearly, the discrete part $[L,L]^{(dis)}$ of the quadratic variation of L is itself a Lévy process. Its Lévy measure $\nu_{[L,L]}$ satisfies for $A \in \mathcal{B}(\mathbb{R}\setminus\{0\})$

$$\begin{aligned}\nu_{[L,L]}(A) &= E\Big[\sum_{s\leq 1} 1_{\{\Delta[L,L]_s^{(dis)}\in A\}}\Big]\\ &= E\Big[\sum_{s\leq 1} 1_{\{(\Delta L_s)^2\in A\}}\Big]\\ &= E\Big[\sum_{s\leq 1} 1_{\{\Delta L_s\in\{x\in\mathbb{R}:x^2\in A\}\}}\Big]\\ &= \nu_L(\{x\in\mathbb{R}:x^2\in A\}),\end{aligned}$$

where ν_L denotes the Lévy measure of L. That means, $\nu_{[L,L]}$ is the image measure of ν_L under the mapping $x\mapsto x^2$. Consequently, $[L,L]^{(dis)}$ has a finite second moment, since

$$\int_{|x|>1} x^2\,\nu_{[L,L]}(dx) = \int_{|x|>1} x^4\,\nu_L(dx) < \infty,$$

which is equivalent to $E[([L,L]_t^{(dis)})^2] < \infty, t\geq 0$. The assertion now follows from Proposition 2.27 and Corollary 2.15. □

According to Proposition 2.37 L^d has a continuous modification. It is this version of L^d we will always consider. The following result shows that L^d additionally constitutes a semimartingale (see Definition A.1).

Proposition 3.30 *Let L be a Levy process with $E[L_1^4]<\infty$. Then the FLP L^d defined by* (3.70) *is a continuous finite variation process. In particular, it constitutes a semimartingale with characteristics given by $(L^d,0,0)$.*

Proof. Observe that for $t\geq 0$ the Molchan-Golosov kernel satisfies for $s\in[0,t]$

$$f_d^{MG}(t,s) = c_d s^{-d}\int_s^t (u-s)^{d-1}u^d\,du \geq 0.$$

Further, $f_d^{MG}(\cdot,s)$ is increasing for each $s>0$ and $[L,L]^{(dis)}$ is a subordinator such that L^d is obviously a.s. increasing.
Now, being a.s. increasing, L^d has paths of finite variation. Further L^d is adapted to the filtration generated by L (see Proposition 2.37) such that [Protter, 2004, section II, Theorem 7] implies the semimartingale property. □

We are now in a position to modify the squared volatility process of the COGARCH$(1,1)$ process. That is, in analogy to Marquardt [2006] modifying the CARMA process (see also Theorem 2.66) we substitute the driving Lévy process $[L,L]^{(dis)}$ in (3.68) by the corresponding MG-FLP (3.70), leading to the following definition.

Definition 3.31 (MG-FICOGARCH$(1,d,1)$) *Let $\alpha_0,\alpha_1,\beta_1>0$ and $d\in(0,0.5)$. Assume L to be a Lévy process with $E[L_1^4]<\infty$. Then the process G satisfying $G_0=0$ a.s. and*

$$dG_t = \sigma_t dL_t \qquad t>0,$$

where the squared volatility $(\sigma_t^2)_{t\geq 0}$ *is given as the solution of the SDE*

$$d\sigma_t^2 = -\beta_1(\sigma_t^2 - a_0)\,dt + \alpha_1\sigma_t^2\,dL_t^d, \qquad t > 0, \tag{3.71}$$

with the MG-FLP L^d *given by* (3.70), *is called* Molchan-Golosov fractionally integrated COGARCH$(1,1)$ *with fractional integration parameter* d, *shortly MG-FICOGARCH*$(1,d,1)$.

Remark 3.32 *In Proposition 3.30 we saw that the paths* $t \mapsto L_t^d(\omega)$ *of the MG-FLP* (3.70) *are almost surely of bounded variation and continuous, i.e. in particular predictable. Consequently, integration with respect to* L^d *is understood in the Stieltjes-sense.*

The subsequent result shows that the structure of the COGARCH$(1,1)$ model is preserved.

Proposition 3.33 *Consider the MG-FICOGARCH*$(1,d,1)$ *model as defined above. Then the solution of the SDE* (3.71) *with initial value* σ_0^2 *is given by*

$$\sigma_t^2 = e^{-X_t}\left(\sigma_0^2 + \alpha_0\beta_1\int_0^t e^{X_s}\,ds\right), \qquad t \geq 0,$$

with

$$X_t = \beta_1 t - \alpha_1 L_t^d, \qquad t \geq 0.$$

Proof. Recall that L^d and therefore $X = (X_t)_{t\geq 0}$ has paths which are a.s. continuous and of bounded variation. Consequently, we can apply *integration by parts*, which leads to

$$\begin{aligned} d\sigma_t^2 &= \left(\sigma_0^2 + \alpha_0\beta_1\int_0^t e^{X_s}\,ds\right)\,d(e^{-X_t}) + e^{-X_t}\,d\left(\sigma_0^2 + \alpha_0\beta_1\int_0^t e^{X_s}\,ds\right) \\ &= \sigma_t^2\,d(-\beta_1 t + \alpha_1 L_t^d) + e^{-X_t}e^{X_t}\alpha_0\beta_1\,dt \\ &= -\beta_1(\sigma_t^2 - \alpha_0)\,dt + \alpha_1\sigma_t^2\,dL_t^d. \end{aligned}$$

□

However, the basic problem of this approach lies in the complexity of the Molchan-Golosov kernel f_d^{MG}: Neither we were able to compute moments of the MG-FICOGARCH or its volatility process, nor could we derive conditions ensuring stationarity. Basically, such conditions cannot be expected to exist as the MG-kernel in general does not allow the corresponding FLP to have stationary increments.

3.2.2.2 Modified MvN-SIMA driven volatility

Recall that whereas the centered FICARMA process was defined in analogy to the long memory (i.e. $d \in (0, 0.5)$ ARFIMA(p,d,q) process X (cf. Remark 2.64),

$$X_n = (1-L)^{-d}\frac{\theta(L)}{\phi(L)}\varepsilon_n, \qquad n \in \mathbb{Z}, 0 < d < 0.5, \tag{3.72}$$

the *non-centered* Lévy driven FICARMA is basically obtained by restricting d to be negative, which corresponds to inverting the fractional difference operator in (3.72). This is in analogy to the FIGARCH, where according to Definition 3.18 the volatility is modeled as

$$\sigma_n^2 = \omega + \left(1 - (1-L)^d \frac{\alpha(L)}{\beta(L)}\right) \varepsilon_n^2 \sigma_n^2, \qquad n \in \mathbb{Z},$$

i.e. in comparison with the ARFIMA process (3.72) the fractional difference operator is again applied being inverted.

To mimic this approach, we consider the MvN-kernel $f_d^{MvN}(t,\cdot)$ for d restricted to be negative, that is

$$f_d^{MvN}(t,s) = \frac{1}{\Gamma(1+d)}\left((t-s)_+^d - (-s)_+^d\right), \qquad d < 0. \tag{3.73}$$

Proposition 3.34 *Let $\delta > 0$ and consider the kernel $f_d^{MvN}(t,\cdot)$ for $d < 0$ and $t \in \mathbb{R}$. Then $|f_d^{MvN}(t,\cdot)|^\delta$ is integrable*

(i) at $-\infty$ if and only if $\delta > \frac{1}{1-d}$,

(ii) at $s = 0$ and $s = t$ if and only if $\delta < \frac{1}{-d}$.

Proof. See [Engelke and Woerner, 2013, Proposition 2]. □

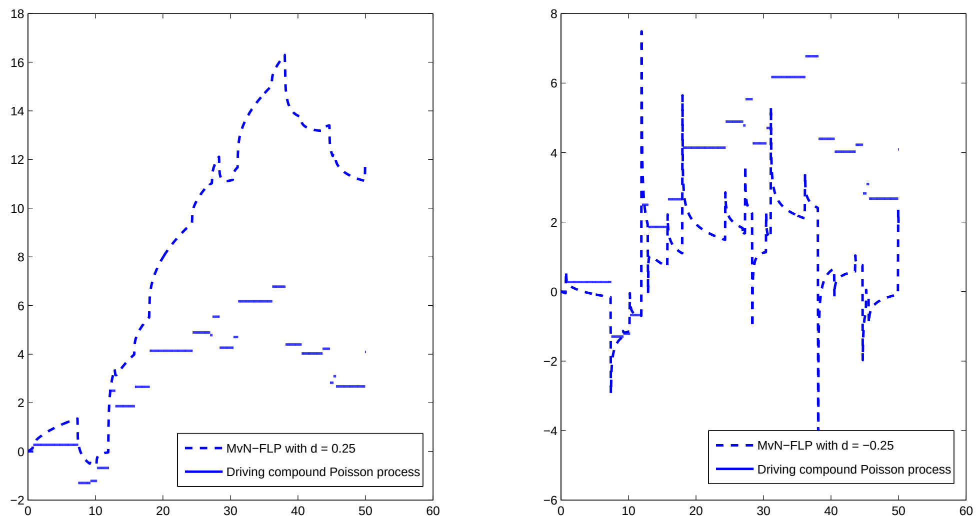

Figure 3.1: Simulated Mandelbrot-van-Ness FLP (dashed line) with fractional integration parameter $d = 0.25$ (left) and $d = -0.25$ (right) driven by a compound Poisson process (solid line) with rate 0.4 and standard normally distributed jump size.

Clearly, if follows that $f_d^{MvN}(t,\cdot) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ as long as $d \in (-0.5, 0)$, such that the MvN-FLP L^d driven by $[L, L]^{(dis)}$,

$$L_t^d = \int_{\mathbb{R}} f_d^{MvN}(t,u)\, d[L,L]_u^{(dis)}, \qquad t \in \mathbb{R}, d \in (-0.5, 0), \tag{3.74}$$

is well-defined. However, having the same covariance structure as fractional Brownian motion, L^d has negatively correlated increments for $d \in (-0.5, 0)$, see Proposition 2.24. Further, due to the singularity at $t = s$ the function $f_d^{MvN}(\cdot, s)$ is discontinuous for all $s \in \mathbb{R}$, such that L^d has discontinuous sample path with positive probability (cf. [Rosinski, 1989, Theorem 4]).

To overcome these drawbacks, it is worth remembering that when modifying the CARMA kernel to allow for a non-centered driving Lévy process, Brockwell and Marquardt [2005] proceeded in two steps: Firstly, they restricted the *fractional integration parameter* d to negative values, as we have done in (3.73). Secondly, they bounded the resulting kernel at its singularity.

Observe that the MvN-kernel $f_d^{MvN}(t,\cdot)$ up to a constant is given by $s \mapsto g_d(t-s) - g_d(-s)$ where the function g_d is defined by $g_d(x) := x_+^d$. This suggests to bound g_d at the singularities $s = 0$ and $s = t$ by incorporating a shift, leading to

$$g_{a,d}(x) := (a + x_+)^d, \qquad d < 0, a > 0. \tag{3.75}$$

In the following we give the definition of the resulting modification of the Mandelbrot-van-Ness kernel.

Definition 3.35 *Let $d < 0$ and $a > 0$. For each $t \in \mathbb{R}$ the* modified MvN-kernel *is given by*

$$\begin{aligned} f_{a,d}(t,s) &= c_{a,d}\left(g_{a,d}(-s) - g_{a,d}(t-s)\right) \\ &= c_{a,d}\left((a + (-s)_+)^d - (a + (t-s)_+)^d\right), \qquad s \in \mathbb{R}, \end{aligned} \tag{3.76}$$

where $c_{a,d}$ is a normalizing constant possibly depending on a and d (see subsequent remark).

Remark 3.36 *By including the constant $c_{a,d}$ in the previous definition, we would like to draw attention to the fact, that it might be appropriate or even necessary to normalize the kernel function for several applications. However, at this point, we do not yet see this necessity, such that in the following we set $c_{a,d} = 1$ and consequently*

$$f_{a,d}(t,s) = (a + (-s)_+)^d - (a + (t-s)_+)^d\,.$$

Remark 3.37 *Observe that besides substituting g_d by $g_{a,d}$, we changed the signs. This way we ensure $f_{a,d}(t,\cdot)$ to be non-negative for $t \geq 0$.*

Proposition 3.38 *For $a > 0$ and $d < 0$ consider the modified MvN-kernel $f_{a,d}$ and let $t \in \mathbb{R}$. Then the following holds.*

(i) $f_{a,d}(t,\cdot)$ is continuous,

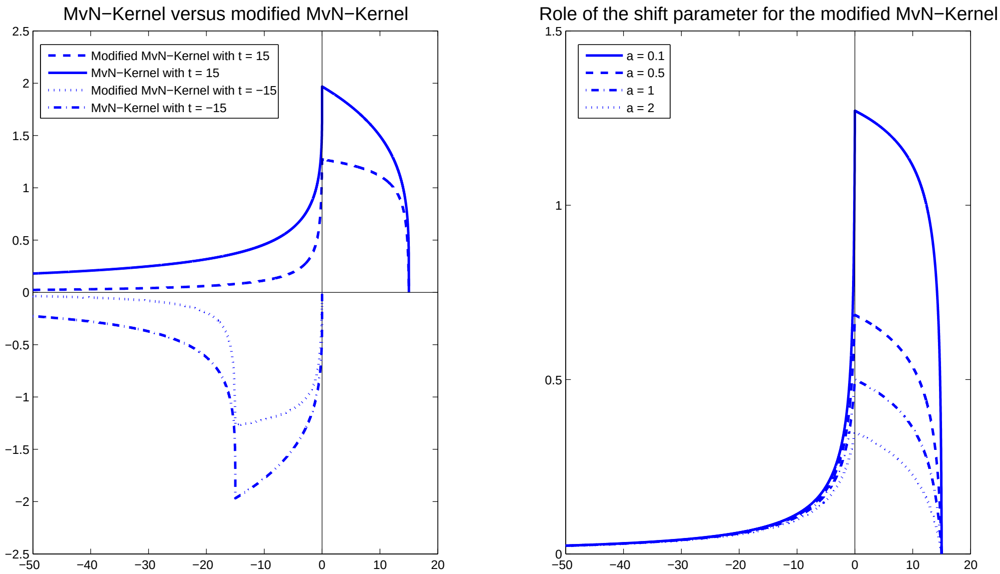

Figure 3.2: Comparison of Mandelbrot-van-Ness kernel $f_d^{MvN}(t,\cdot)$ and modified Mandelbrot-van-Ness kernel $f_{a,-d}(t,\cdot)$ for $d = 0.25, a = 0.1$ (left) and Mandelbrot-van-Ness kernel $f_{a,d}(t,\cdot)$ for $d = -0.25$ and different values of the shift parameter a (right).

(ii) bounded by a^d, i.e.

$$|f_{a,d}(t,s)| \le a^d, \qquad s \in \mathbb{R},$$

(iii) and for $s \longrightarrow -\infty$ its asymptotics are given by

$$|f_{a,d}(t,s)| \sim |s|^{d-1}.$$

In particular, the kernel satisfies

$$f_{a,d}(t,\cdot) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}).$$

Proof. The first two claims are obvious. Further note that for $d < 0$

$$\begin{aligned}\lim_{s\to-\infty} \frac{|f_{a,d}(t,s)|}{|s|^{d-1}} &= \lim_{s\to-\infty} \frac{(a-s)^d - (t+a-s)^d}{|s|^{d-1}} \\ &= \lim_{v\to-\infty} \frac{(-v)^d - (t-v)^d}{|v+a|^{d-1}} \\ &= \lim_{v\to-\infty} \frac{|f_d^{MvN}(t,v)|}{|v+a|^{d-1}}.\end{aligned}$$

In [Engelke and Woerner, 2013, Proposition 2] it is proved that

$$|f_d^{MvN}(t,v)| \sim |v|^{d-1}, \qquad v \longrightarrow -\infty,$$

such that the assertion follows. □

By substituting the MvN-kernel $f_d^{MvN}(t,\cdot)$ in (3.74) with $f_{a,d}(t,\cdot)$, we define the process $M^{a,d}$ as

$$M_t^{a,d} = \int_{\mathbb{R}} f_{a,d}(t,u)\; d[L,L]_u^{(dis)}, \qquad t \in \mathbb{R}, \tag{3.77}$$

where $d < 0$ and $a > 0$.

Proposition 3.39 *Let $a > 0$ and $d < 0$. Further denote by $L = (L_t)_{t\in\mathbb{R}}$ a two-sided Lévy process satisfying $E[L_1^4] < \infty$. Then $M_t^{a,d}$ as defined in* (3.77) *exists for all $t \in \mathbb{R}$ as limit in $L^2(\Omega)$-sense.*

Proof. The proof is similar to the one of Proposition 3.29. □

The advantages of the new defined kernel $f_{a,d}$ towards the MG-kernel f^{MG} do not only lie in its much simpler structure. In contrast to L^d in (3.70) the convolution $M^{a,d}$ does also have stationary increments. This will be especially important when it comes to modifying the squared volatility of the COGARCH(1, 1) process.
Apart from this the question arises if $f_{a,d}$ serves its purpose in the sense that the increments of $M^{a,d}$ exhibit an *appropriate* dependence structure. We address these issues in the subsequent result.

Proposition 3.40 *Let $a > 0$ and $d \in (-0.5, 0)$. Denote by $M^{a,d}$ the process as defined in* (3.77)*, where the driving Lévy process L satisfies $E[L_1^4] < \infty$. Then the following holds.*

(i) $M^{a,d}$ has stationary increments, i.e. for all $s,t \in \mathbb{R}, s < t$, we have

$$M_t^{a,d} - M_s^{a,d} \overset{\mathcal{D}}{=} M_{t-s}^{a,d}.$$

(ii) The covariance γ_h of two increments $M_{t+h}^{a,d} - M_t^{a,d}$ and $M_{s+h}^{a,d} - M_s^{a,d}$ of length $h > 0$, where $s + h \le t$ such that $t - s = rh$,

$$\gamma_h(r) := \operatorname{Cov}\left(M_{s+(r+1)h}^{a,d} - M_{s+rh}^{a,d}, M_{s+h}^{a,d} - M_s^{a,d}\right),$$

satisfies for $r \longrightarrow \infty$

$$\gamma_h(r) \sim \operatorname{Var}\left([L,L]_1^{(dis)}\right)(-d)a^d h^2\,(rh+a)^{d-1}. \tag{3.78}$$

Before we give the proof, we need the following lemma.

Lemma 3.41 *Let $a > 0$ and $d \in (-0.5, 0)$. Then the modified MvN-kernel $f_{a,d}$ satisfies for $t > 0$*

$$\int_{\mathbb{R}} f_{a,d}^2(t,u)\; du = a^{2d}\, t - \frac{2a^d}{d+1}(t+a)^{d+1} + \frac{1}{2d+1}(t+a)^{2d+1} + c(t)\; t^{2d+1} + C,$$

with $C = a^{2d+1}\left(\frac{2}{d+1} - \frac{1}{2d+1}\right)$ and

$$c(t) \longrightarrow \frac{\Gamma(d+1)}{\Gamma(2d+2)\sin(\pi(d+0.5))} + \frac{1}{2d+1}, \qquad t \longrightarrow \infty. \tag{3.79}$$

Proof. By substituting $y := \frac{u-a}{t}$ we obtain

$$\begin{aligned}
\int_{\mathbb{R}} f_{a,d}^2(t,u)\, du &= \int_{-\infty}^{0} \left[(a-u)^d - (a+t-u)^d\right]^2\, du + \int_0^t \left[a^d - (a+t-u)^d\right]^2\, du \\
&= \int_{-\infty}^{0} t^{2d} \left[\left(1 - \frac{u-a}{t}\right)^d - \left(-\frac{u-a}{t}\right)^d\right]^2\, du + \int_0^t \left[a^d - (a+t-u)^d\right]^2\, du \\
&= t^{2d+1} \int_{-\infty}^{-a/t} \left[(1-y)^d - (-y)^d\right]^2\, dy + \int_0^t a^{2d} - 2a^d\,[t+a-u]^d + [t+a-u]^{2d}\, du \\
&= t^{2d+1} c(t) + a^{2d}\, t + \frac{2a^d}{d+1}\,[t+a-u]^{d+1}\Big|_0^t - \frac{1}{2d+1}\,[t+a-u]^{2d+1}\Big|_0^t,
\end{aligned}$$

where $c(t) = \int_{-\infty}^{-a/t} \left[(1-y)^d - (-y)^d\right]^2\, dy$. Further note that using equation (2.34) it holds

$$\begin{aligned}
\int_{-\infty}^{1} \left[(1-y)_+^d - (-y)_+^d\right]^2\, dy &= \Gamma(d+1) \int_{-\infty}^{1} \left(f_d^{MvN}(1,y)\right)^2\, dy \\
&= \frac{\Gamma(d+1)}{\Gamma(2d+2)\sin(\pi(d+0.5))}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
c(t) &= \frac{\Gamma(d+1)}{\Gamma(2d+2)\sin(\pi(d+0.5))} - \int_{-a/t}^{0} \left[(1-y)^d - (-y)^d\right]^2\, dy - \int_0^1 (1-y)^{2d}\, dy \\
&= \frac{\Gamma(d+1)}{\Gamma(2d+2)\sin(\pi(d+0.5))} - \int_{-a/t}^{0} \left[(1-y)^d - (-y)^d\right]^2\, dy + \frac{1}{2d+1}.
\end{aligned}$$

Since $\int_{-a/t}^{0} \left[(1-y)^d - (-y)^d\right]^2\, dy \longrightarrow 0$ for $t \longrightarrow \infty$, the assertion holds. □

We are now ready to give the proof of Proposition 3.40.

Proof. For any $s, t \in \mathbb{R}, s < t$ we find using the stationarity of the increments of $[L,L]^{(dis)}$ that

$$\begin{aligned}
M^{a,d}(t) - M^{a,d}(s) &= \int_{\mathbb{R}} \left(a + (s-u)_+\right)^d - \left(a + (t-u)_+\right)^d\, d[L,L]_u^{(dis)} \\
&\overset{\mathscr{D}}{=} \int_{\mathbb{R}} \left(a + (-v)_+\right)^d - \left(a + (t-s-v)_+\right)^d\, d[L,L]_v^{(dis)} \\
&= M^{a,d}(t-s).
\end{aligned}$$

In order to prove the second part, let us introduce the following notation, namely $\tilde{L}_t := [L,L]_t^{(dis)} - E\big[[L,L]_t^{(dis)}\big]$ and $\tilde{M}_t^{a,d} := \int_{\mathbb{R}} f_{a,d}(t,u)\, d\tilde{L}_u$. For $t, s \geq 0$ it then holds that

$$\begin{aligned}
\operatorname{Cov}(M_t^{a,d}, M_s^{a,d}) &= \operatorname{Cov}(\tilde{M}_t^{a,d}, \tilde{M}_s^{a,d}) \\
&= E\big[\tilde{M}_t^{a,d}\tilde{M}_s^{a,d}\big] \\
&= \frac{1}{2}\left(E\Big[\left(\tilde{M}_t^{a,d}\right)^2\Big] + E\Big[\left(\tilde{M}_s^{a,d}\right)^2\Big] - E\Big[\left(\tilde{M}_t^{a,d} - \tilde{M}_s^{a,d}\right)^2\Big]\right) \\
&= \frac{1}{2}\left(E\Big[\left(\tilde{M}_t^{a,d}\right)^2\Big] + E\Big[\left(\tilde{M}_s^{a,d}\right)^2\Big] - E\Big[\left(\tilde{M}_{t-s}^{a,d}\right)^2\Big]\right).
\end{aligned}$$

In the last step we used that the increments are stationary. Furthermore, applying (2.20), we find that

$$\begin{aligned}
E\Big[\left(\tilde{M}_t^{a,d}\right)^2\Big] &= \operatorname{Var}\left([L,L]_1^{(dis)} - E\big[[L,L]_1^{(dis)}\big]\right)\int_{\mathbb{R}} f_{a,d}^2(t,u)\,du \\
&= \operatorname{Var}\left([L,L]_1^{(dis)}\right)\int_{\mathbb{R}} f_{a,d}^2(t,u)\,du,
\end{aligned}$$

such that the covariance between increments of length $h>0$ satisfies

$$\begin{aligned}
\gamma_h(r) &= \operatorname{Cov}\left(M_{s+(r+1)h}^{a,d}, M_{s+h}^{a,d}\right) - \operatorname{Cov}\left(M_{s+(r+1)h}^{a,d}, M_s^{a,d}\right) - \operatorname{Cov}\left(M_{s+rh}^{a,d}, M_{s+h}^{a,d}\right) \\
&\quad + \operatorname{Cov}\left(M_{s+rh}^{a,d}, M_s^{a,d}\right) \\
&= \frac{1}{2}\left(E\Big[\left(\tilde{M}_{(r+1)h}^{a,d}\right)^2\Big] + E\Big[\left(\tilde{M}_{(r-1)h}^{a,d}\right)^2\Big] - 2E\Big[\left(\tilde{M}_{rh}^{a,d}\right)^2\Big]\right) \\
&= \frac{1}{2}\operatorname{Var}\left([L,L]_1^{(dis)}\right)\left[\int_{\mathbb{R}} f_{a,d}^2\big((r+1)h,u\big)\,du + \int_{\mathbb{R}} f_{a,d}^2\big((r-1)h,u\big)\,du - 2\int_{\mathbb{R}} f_{a,d}^2(rh,u)\,du\right].
\end{aligned}$$

Now, using Lemma 3.41 we obtain

$$\begin{aligned}
\gamma_h(r) = \frac{1}{2}\operatorname{Var}\left([L,L]_1^{(dis)}\right)\Big[&-\frac{2a^d}{d+1}\left(((rh+a)+h)^{d+1} + ((rh+a)-h)^{d+1} - 2(rh+a)^{d+1}\right) \\
&+ \frac{1}{2d+1}\left(((rh+a)+h)^{2d+1} + ((rh+a)-h)^{2d+1} - 2(rh+a)^{2d+1}\right) \\
&+ c(rh+h)(rh+h)^{2d+1} + c(rh-h)(rh-h)^{2d+1} - 2c(rh)(rh)^{2d+1}\Big],
\end{aligned}$$

where $c(t)$ is defined as in Lemma 3.41 and according to (3.79) converges for $t \longrightarrow \infty$ to a constant, which will be denoed in the following by C. Consequently, Taylor expansion and the fact that $\frac{c(rh\pm h)}{c(rh)} \longrightarrow 1$ gives for $r \longrightarrow \infty$

$$\begin{aligned}
\gamma_h(r) = \frac{1}{2}\operatorname{Var}\left([L,L]_1^{(dis)}\right)\Big[&-\frac{2a^d}{d+1}(rh+a)^{d+1}\left((d+1)d\frac{h^2}{(rh+a)^2} + \mathcal{O}\left(\frac{1}{(rh+a)^4}\right)\right) \\
&+ \frac{(rh+a)^{2d+1}}{2d+1}\left((2d+1)2d\frac{h^2}{(rh+a)^2} + \mathcal{O}\left(\frac{1}{(rh+a)^4}\right)\right) \\
&+ C\left((2d+1)2d\frac{h^2}{(rh)^2} + \mathcal{O}\left(\frac{1}{(rh)^4}\right)\right)\Big] \\
\sim \operatorname{Var}\left([L,L]_1^{(dis)}\right)(-d)a^d h^2\,(rh+a)^{d-1}&.
\end{aligned}$$

□

Remark 3.42 *The increments of $M^{a,d}$ do not have long memory in the sense of Definition 2.60. However, for d close to zero we approximately obtain long memory. This is in analogy to the asymptotic rate of decay of the modified CARMA kernel $g_{a,d}$ in the case of non-centered Lévy-driven CARMA processes, see section 2.6.2.2.*

Remark 3.43 *The result* (3.78) *implies that asymptotically the increments of $M^{a,d}$ are positively correlated. Further, defining the mapping*

$$I_f(t) := \int_{\mathbb{R}} f_{a,d}^2(t,u)\, du,$$

we get from the proof of Proposition 3.40 that

$$\begin{aligned}\gamma_h(t,s) :&= \operatorname{Cov}\left(M_{t+h}^{a,d} - M_t^{a,d}, M_{s+h}^{a,d} - M_s^{a,d}\right)\\ &= \frac{1}{2}\operatorname{Var}\left([L,L]_1^{(dis)}\right)\Big[\int_{\mathbb{R}} f_{a,d}^2\big(t-s+h,u\big)\, du + \int_{\mathbb{R}} f_{a,d}^2\big(t-s-h,u\big)\, du\\ &\qquad\qquad - 2\int_{\mathbb{R}} f_{a,d}^2(t-s,u)\, du\Big]\\ &= \frac{1}{2}\operatorname{Var}\left([L,L]_1^{(dis)}\right)\big[\left(I_f(t-s+h) - I_f(t-s)\right) - \left(I_f(t-s) - I_f(t-s-h)\right)\big].\end{aligned}$$

Consequently, if $t \mapsto I_f(t)$ is convex, then the increments are always positively correlated. We conjecture that the convexity holds, but could not proof it. However, several plots of I_f support this hypothesis.

Both the MvN-FLP as well as the new process $M^{a,d}$ are members of the large class of *stationary increment moving-average*, shortly SIMA, processes. These are stochastic processes, which can be represented in the form

$$X_t = \int_{\mathbb{R}} f(t-s) - f_0(-s)\, dL_s, \qquad t \geq 0, \tag{3.80}$$

where $f, f_0 : \mathbb{R} \longrightarrow \mathbb{R}$ are measurable functions, satisfying $f(x) = f_0(x) = 0$ as soon as $x < 0$, and L denotes a Lévy process, such that (3.80) is well-defined in probability (see Theorem 2.10). By setting

$$f(x) = f_0(x) := x_+^d, \qquad d \in (0, 0.5)$$

we obtain the MvN-FLP, while

$$f(x) = f_0(x) := a^d - (a + x_+)^d, \qquad a > 0,\ d \in (-0.5, 0),$$

leads to the SIMA $M^{a,d}$. SIMA processes are studied concerning their finite variation and semimartingale property in O'Connor and Rosinski [2013] and O'Connor and Rosinski [2014] respectively as particular subclass of stationary increment infinitely divisible processes. The subsequent result characterizes SIMA processes, which have the semimartingale property.

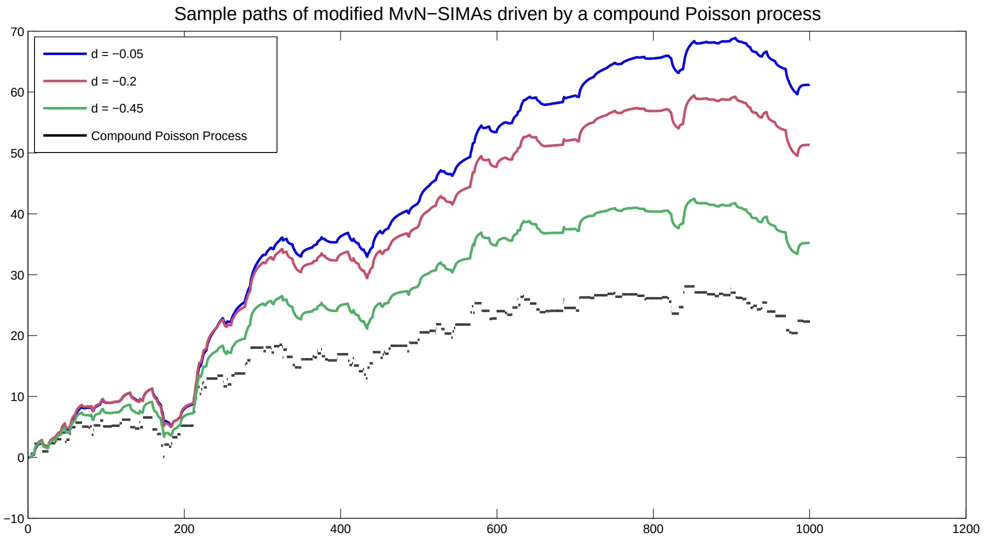

Figure 3.3: Simulated sample paths of SIMAs with modified MvN-kernel $f_{a,d}$ with $a = 1$ and different fractional integration parameters d driven by a compound Poisson process with rate 0.2 and standard normally distributed jumps sizes.

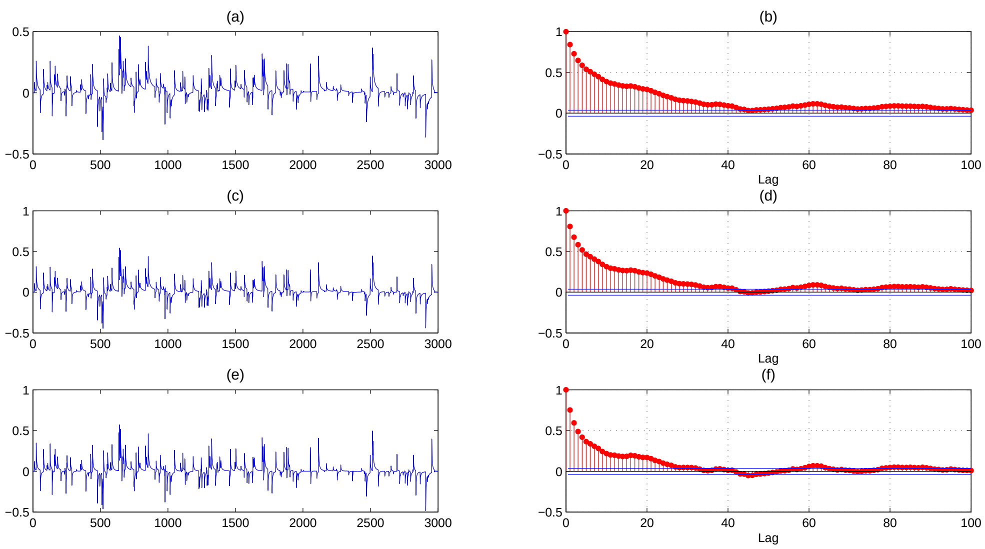

Figure 3.4: Increments and corresponding sample ACF for the SIMAs shown in Figure 3.3 for fractional integration parameter $d = -0.05$ ((a), (b)), $d = -0.2$ ((c), (d)), $d = -0.45$ ((e), (f))

Theorem 3.44 *Let $X = (X_t)_{t\geq 0}$ be a SIMA process of the form* (3.80) *driven by the Lévy process $L = (L_t)_{t\in\mathbb{R}}$. Then X is a semimartingale with respect to the filtration generated by L if and only if*

$$X_t = X_0 + M_t + A_t, \qquad t \geq 0,$$

where $M = (M_t)_{t\geq 0}$ is the Lévy process given by $M_t = f(0)L_t$, $t \geq 0$, and $A = (A_t)_{t\geq 0}$ is a predictable process of finite variation given by

$$A_t = \int_{\mathbb{R}} f(t-s) - f(-s) dL_s - f(0)L_t, \qquad t \geq 0.$$

Proof. See [O'Connor and Rosinski, 2014, Theorem 4.1]. □

Furthermore, O'Connor and Rosinski [2014] establish sufficient conditions on f and the driving Lévy process L for a SIMA to be a semimartingale. Recall therefore that a function $h : I \to \mathbb{R}$, where $I \subset \mathbb{R}$ is an interval, is said to be *absolutely continuous*, if there exists a locally integrable function $\dot{h}$ such that

$$h(t) - h(s) = \int_s^t \dot{h}(u)\, du, \qquad s < t.$$

Theorem 3.45 *Let $X = (X_t)_{t\geq 0}$ be a SIMA process of the form* (3.80) *driven by the Lévy process $L = (L_t)_{t\in\mathbb{R}}$ with characteristic triplet $(\gamma_L, \sigma_L^2, \nu_L)$. Suppose that f is absolutely continuous on $[0, \infty)$ with derivative $\dot{f}$ satisfying*

$$\int_0^\infty |\dot{f}(s)|^2 ds\, \sigma_L^2 < \infty,$$

and

$$\int_0^\infty \int_{\mathbb{R}} \min\left(|x\dot{f}(s)|, |x\dot{f}(s)|^2\right)\, \nu_L(dx)ds < \infty.$$

Then X is a semimartingale with respect to the filtration generated by L.

Proof. See [O'Connor and Rosinski, 2014, Theorem 4.2]. □

Using these results it is straightforward to show that the SIMA process $M^{a,d}$ is a semimartingale. We summarize this result in the subsequent Corollary.

Corollary 3.46 *Let $M^{a,d}$ be the process as defined in* (3.77) *with driving Lévy process $L = (L_t)_{t\in\mathbb{R}}$, satisfying $E[L_1^4] < \infty$. Then $(M^{a,d})_{t\geq 0}$ is predictable and has path of finite variation.*
In particular, it constitutes a semimartingale (with respect to the filtration generated by L) with characteristics given by $(M^{a,d}, 0, 0)$.

Proof. Recall that the discrete part of the quadratic variation $[L, L]^{(dis)}$ is itself a Lévy process with characteristics $(\gamma_{[L,L]}, 0, \nu_{[L,L]})$, where $\nu_{[L,L]}$ is the image measure of ν_L under the mapping $x \mapsto x^2$. Consequently,

$$\int_0^\infty \int_{\mathbb{R}} \min\left(|x\dot{f}(s)|, |x\dot{f}(s)|^2\right)\ \nu_{[L,L]}(dx)ds \le \int_0^\infty \int_{\mathbb{R}} |x||\dot{f}(s)|\ \nu_{[L,L]}(dx)ds$$
$$= \int_{\mathbb{R}} x^2 \nu_L(dx) \int_0^\infty -d(a+s)^{d-1}\ ds$$
$$< \infty.$$

Consequently, $M^{a,d}$ is a semimartingale. Now, it follows from Theorem 3.44 that its characteristics are given by $(M^{a,d}, 0, 0)$. □

Due to Corollary 3.46 integration with respect to $M^{a,d}$ can be defined in the Stieltjes-sense. Thus, by substituting the driving noise $[L,L]^{(dis)}$ in the squared volatility process (3.68) with $M^{a,d}$, we are now in the position to introduce a further *fractionally integrated* COGARCH$(1,1)$ process.

Definition 3.47 (Modified-MvN FICOGARCH$(1,d,1)$) *Let $\alpha_0, \alpha_1, \beta_1 > 0$ and $d \in (-0.5, 0)$. Assume L to be a Levy process with $E[L_1^4] < \infty$. Then the process G satisfying $G_0 = 0$ a.s. and*

$$dG_t = \sigma_{t-} dL_t \qquad t > 0,$$

where the squared volatility $(\sigma_t^2)_{t\ge 0}$ is given as the solution of the SDE

$$d\sigma_t^2 = -\beta_1(\sigma_{t-}^2 - a_0)\ dt + \alpha_1 \sigma_{t-}^2\ dM_t^{a,d}, \qquad t > 0, \tag{3.81}$$

with the modified MvN-SIMA $M^{a,d}$ given by (3.80)*, is called* Modified-MvN fractionally integrated COGARCH$(1,1)$ *with fractional integration parameter d, shortly modified-MvN FICOGARCH$(1,d,1)$.*

Again, due to the finite variation property of $M^{a,d}$, we can state the solution of the SDE (3.81) explicitly. This is shown in the following proposition.

Proposition 3.48 *Consider the modified MvN-FICOGARCH$(1,d,1)$ model as defined above. Then the solution of the SDE* (3.81) *with initial value σ_0^2 is given by*

$$\sigma_t^2 = e^{-X_t}\left(\sigma_0^2 + \alpha_0\beta_1 \int_0^t e^{X_s}\ ds\right), \qquad t \ge 0, \tag{3.82}$$

with

$$X_t = \beta_1 t - \alpha_1 M_t^{a,d}, \qquad t \ge 0.$$

Proof. As $M^{a,d}$ is a semimartingale, we can apply *integration by parts*. Thus, taking the finite variation property of $M^{a,d}$ into consideration, (3.82) fulfills the SDE (3.81) as shown in the proof of Proposition 3.33. □

3.2.2.3 Simulation

Before we turn to the simulation of the FICOGARCH, we are concerned with the simulation of convoluted Lévy processes, such as MG-FLPs or Lévy driven SIMAs.
For $T > 0$ we consider therefore the process $M = (M_t)_{t\in[0,T]}$ given by

$$M_t = \int_{\mathbb{R}} f(t,s)\, dL_s, \qquad t \in [0,T],$$

where f denotes a deterministic kernel with $f(t,s) = 0, s \geq t$, and $L = (L_t)_{t\in\mathbb{R}}$ is a Lévy process. For $n \in \mathbb{N}$ we set $\Delta_n := T/n,\ f_{i,k}^{(n)} := f(i\Delta_n, k\Delta_n)$ and approximate the paths of M at the grid points $i\Delta_n, i = 1,\dots,n$ by the Riemann sums

$$M_{i\Delta_n}^{(n)} = \sum_{k=-n}^{i-1} f_{i,k}^{(n)} \left(L_{(k+1)\Delta_n} - L_{k\Delta_n}\right).$$

This leads to the following simulation routine for a path of $(M_t)_{t\in[0,T]}$.

(i) Generate the increments of the driving Lévy process $\Delta L_k^{(n)} := L_{(k+1)\Delta_n} - L_{k\Delta_n}$, $k = -n,\dots,n-1$.

(ii) Calculate the kernel matrix $K \in M(n \times 2n)$, which is given by $K_{i,j}^{(n)} := f_{i,j-1-n}^{(n)}$, $i = 1,\dots,n,\ j = 1,\dots,2n$. Note that $f_{i,k}^{(n)} = 0$ for $k \geq i$.

(iii) Calculate the discretized path $M^{(n)} = (M_{i\Delta_n}^{(n)})_{i=1,\dots,n}$ by

$$M^{(n)} = K^{(n)}\,(\Delta L^{(n)})'.$$

Remark 3.49 *When simulating paths of the MG-FLP with fractional integration parameter $d \in (0, 0.5)$, then the kernel f, given by*

$$f(t,s) = c_d\, d \int_s^t (u-s)^{d-1} \left(\frac{u}{s}\right)^d du\ 1_{[0,t]}(s), \qquad t > 0,$$

has to be evaluated numerically. Hence, the calculation of the kernel matrix $K^{(n)}$, which in this case reduces to $K_{i,j}^{(n)} = f_{i,j-1}^{(n)}, i,j = 1,\dots,n$ is very time-consuming. However, some calculation time can be saved by using that

$$f_{i,k}^{(n)} = f_{i-1,k}^{(n)} + c_d\, d \int_{(i-1)\Delta_n}^{i\Delta_n} (u-s)^{d-1} \left(\frac{u}{s}\right)^d du, \qquad i > k.$$

We now turn to the simulation of processes $(\sigma_t^2)_{t\geq 0}$ satisfying

$$d\sigma_t^2 = -\beta_1(\sigma_{t-}^2 - a_0)\, dt + \alpha_1 \sigma_{t-}^2\, dM_t, \qquad t > 0. \tag{3.83}$$

Observe that by setting $M_t := [L,L]_t^{(dis)}$, where $[L,L]^{(dis)}$ denotes the discrete part of the quadratic variation of a Lévy process L, we obtain the squared volatility process of

the COGARCH(1, 1) model, see (3.58). Moreover, $M_t = L^d_t$ with L^d given by (3.70) and $M_t = M^{a,d}_t$ with $M^{a,d}$ defined as in (3.77) leads to the FICOGARCH models given in Definition 3.31 and Definition 3.47, referently.
Applying Euler-discretization, the following recursion gives a disretized path of σ^2,

$$\sigma^2_{(i+2)\Delta_n} = \sigma^2_{(i+1)\Delta_n} - \beta_1(\sigma^2_{(i+1)\Delta_n} - \alpha_0)\Delta_n + \alpha_1 \sigma^2_{i\Delta_n}\left(M_{(i+2)\Delta_n} - M_{(i+1)\Delta_n}\right), \qquad i = 0, \ldots, n-2,$$

where we set $\sigma^2_0 = \sigma^2_{\Delta_n} = \alpha_0$.

Remark 3.50 *If the driving process M in* (3.83) *is given by the MG-FLP* (3.70)*, then its a.s. continuity implies*

$$d\sigma^2_t = -\beta_1(\sigma^2_t - a_0)\, dt + \alpha_1 \sigma^2_t \, dM_t, \qquad t > 0. \tag{3.84}$$

In this case, we shall use the following recursion for the simulation of a path of σ^2, namely

$$\sigma^2_{(i+1)\Delta_n} = \sigma^2_{i\Delta_n} - \beta_1(\sigma^2_{i\Delta_n} - \alpha_0)\Delta_n + \alpha_1 \sigma^2_{i\Delta_n}\left(M_{(i+1)\Delta_n} - M_{i\Delta_n}\right), \qquad i = 0, \ldots, n-1,$$

with $\sigma^2_0 = \alpha_0$.

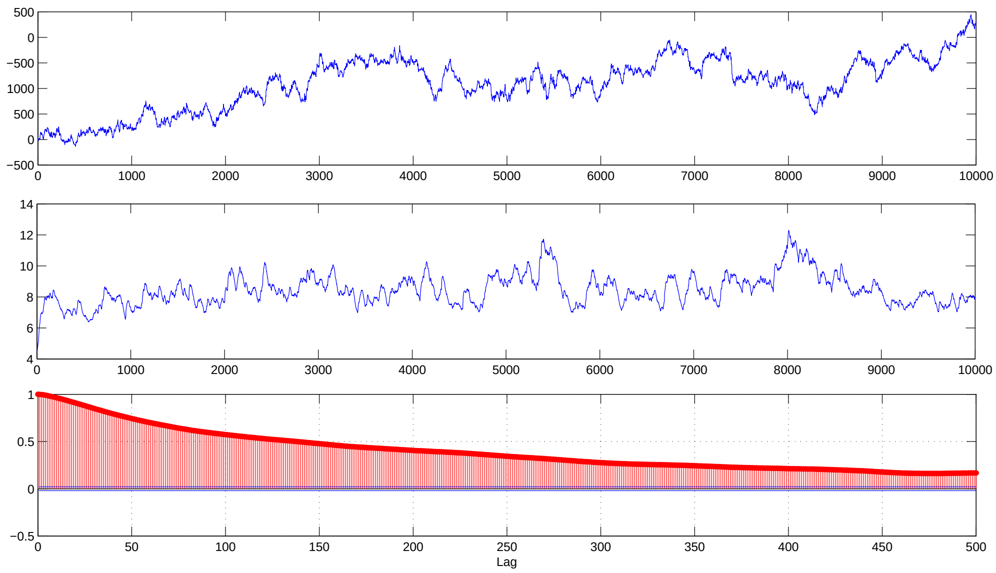

Figure 3.5: Simulation of the modified MvN-FICOGARCH process (top) with corresponding volatility process σ (middle) and ACF of σ (bottom) driven by a compound Poisson process with rate 1 and standard normally distributed jump sizes. The model parameters were $\alpha_0 = 19.4957, \alpha_1 = 0.0105, \beta_1 = 0.0513, d = -0.1$ and $a = 1$.

Chapter 4

Conclusion

In this thesis we successfully generalized the ideas, which were used by Douc et al. [2008] to show the existence of the FIGARCH$(0, d, 0)$, and proved the existence of the FIGARCH$(1, d, 0)$ and FIGARCH$(0, d, 1)$.

In the continuous-time setting we proposed two approaches to incorporate long range dependence into the COGARCH model. For the first one we substituted the driving subordinator of the COGARCH volatility with the corresponding Molchan-Golosov fractional Lévy process (MG-FLP). However, due to the complexity of the MG-kernel, we neither were able to analyze the memory structure of the resulting Molchan-Golosov fractionally integrated COGARCH model (MG-FICOGARCH), nor to derive results concerning stationarity. In the light of the fact that MG-FLPs in general do not have stationary increments, the question arises if the latter is to be expected at all.

For the second approach we introduced a modified Mandelbrot-van-Ness (MvN) kernel. In its structure this new kernel is as simple and flexible as the ordinary MvN-kernel. In contrast, however, it allows for integration with respect to non-centered Lévy processes. We showed that the convoluted Lévy process resulting from our modified MvN-kernel belongs to the class of stationary increment moving averages (SIMA). Further, we proved that its increments decay by a slow hyperbolic rate.

Conclusively, we simulated the resulting Modified MvN-FICOGARCH. The simulation results suggest that the memory properties of the modified MvN-SIMA propagate to the volatility process of the FICOGARCH model.

Concerning the proposed Modified-MvN-FICOGARCH many questions remain open. Firstly, conditions ensuring the stationarity of the model need to be derived. Furthermore, the second order structure of both the FICOGARCH volatility process as well as that of the squared FICOGARCH increments need to be analyzed in detail.

Finally, the modified MvN-kernel SIMA raises the question, whether it can be used to define a non-centered FICARMA process in analogy to Marquardt [2006] deriving the centered FICARMA by substituting the zero-mean driving Lévy process of the CARMA model by the corresponding MvN-FLP.

Appendix A

Semimartingales

Among all stochastic processes the class of *semimartingales* plays a very special role. This is due to the fact that it is the most general class of processes with respect to which stochastic integration can be defined in a meaningful way. See, for example, Protter [2004], where semimartingales are in fact defined this way. For the subsequently presented results we refer to Jacod and Shiryaev [2002].
Before we give its definition, recall that a process $(X_t)_{t\geq 0}$ is called *local martingale*, if there exists an increasing sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ with $\tau_n \to \infty$ a.s. for $n \to \infty$, such that for all $n \in \mathbb{N}$ the stopped process $(X_{\min(t,\tau_n)})_{t\geq 0}$ is a martingale.

Definition A.1 *Consider a stochastic process $X = (X_t)_{t\geq 0}$ on a filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\geq 0})$. Then X is a* semimartingale *if it has a decomposition*

$$X_t = X_0 + A_t + M_t, \qquad t \geq 0, \tag{A.1}$$

where $A = (A_t)_{t\geq 0}$ is a càdlàg adapted process with paths of finite variation, $M = (M_t)_{t\geq 0}$ is a local martingale, X_0 is $\mathcal{F}_0$-measurable and $A_0 = M_0 = 0$.
If A is additionally predictable, then X is called special semimartingale. *In this case the decomposition* (A.1) *is unique up to P-null sets and is called the* canonical decomposition.

Notice that *Lévy processes*, introduced in section 2.1, belong to the class of semimartingales. According to [Jacod and Shiryaev, 2002, Theorem 4.18], the local martingale M in (A.1) can be uniquely (up to P-null sets) decomposed into $M = M^c + M^d$, where M^c is a continuous local martingale and M^d denotes a *purely discontinuous local martingale*, i.e. it is a local martingale such that for all continuous local martingales N the product $M^d N$ is again a local martingale. Consequently, we can rewrite the decomposition (A.1) into

$$X_t = X_0 + A_t + M_t^c + M_t^d, \qquad t \geq 0. \tag{A.2}$$

When describing the jump behavior of semimartingales the concept of *random measures* is very useful.

Definition A.2 *Let $(\Sigma, \mathcal{S})$ be a measurable space. Then a mapping*

$$\mu : \Omega \times \Sigma \longrightarrow [0, \infty) \tag{A.3}$$

is called random measure *if*

(i) $\omega \mapsto \mu(\omega, S)$ *is a random variable for each* $S \in \mathcal{S}$,

(ii) $S \mapsto \mu(\omega, S)$ *is a measure for all* $\omega \in \Omega$.

Assume further that there exists a σ*-finite measure* ν *on* $(\Sigma, \mathcal{S})$ *and let* $S, S_1, \ldots, S_n \in \mathcal{S}$ *be disjoint such that* $\nu(S) < \infty$ *and* $\nu(S_i) < \infty, i = 1, \ldots, n$. *Then we say that* μ *is a* Poisson random measure *on* Σ *with* intensity measure ν *if it additionally holds that*

(ii) the random variable $\mu(S)$ *is Poisson distributed with intensity* $\nu(S)$*, and*

(ii) $\mu(S_1), \ldots, \mu(S_n)$ *are independent.*

With a semimartingale X we can associate a random measure by setting for any measurable subset $S \subset \mathbb{R}_0$

$$\mu_X(\omega, [0,t] \times S) := \#\left\{0 \leq s \leq t : \Delta X_s(\omega) \in S)\right\}. \tag{A.4}$$

If the closure $\overline{S}$ of S does not contain 0, then the process $\mu_X(t, S)$ defined by $t \mapsto \mu_X(\omega, [0,t] \times S)$ is càdlàg, adapted and non-decreasing. [Jacod and Shiryaev, 2002, Theorem 1.18] implies that there exists a unique random measure ν_X on $(0, \infty) \times \mathbb{R}_0$, called *compensator* of μ_X, such that $\mu_X(t, S) - \nu_X(t, S)$ (with $t \mapsto \nu_X(t, S)$ being defined as $\mu_X(t, S)$) is a local martingale. This leads to the following definition.

Definition A.3 *Denote by* $X = (X_t)_{t \geq 0}$ *a special semimartingale with canonical decomposition*

$$X_t = X_0 + A_t + M_t^c + M_t^d, \qquad t \geq 0.$$

Then the triplet $((A_t)_{t \geq 0}, (C_t)_{t \geq 0}, \nu_X)$ *is called* characteristics *of* X*, where*

(i) $C = [M^c, M^c]$ *with* $[\cdot, \cdot]$ *denoting the* quadratic variation*, and*

(ii) ν_X *denotes the compensator of the random measure* μ_X *associated to jumps of* X*, see* (A.4)*.*

Appendix B

MATLAB-Code

B.1 Main classes

```
classdef LP
   properties
       % Structure containing the name and the parameters
       % specifying the distribution of the increments of the Levy process
       % E.g.: distribution.Name = 'poisson', distribution.Param = 1
       Distribution
   end

   methods
       function this = LP(varargin)
           % Default distribution = Poisson(1)
           if nargin < 1 || ~ischar(varargin{1})
               this.Distribution.Name = 'poisson';
               this.Distribution.Param = 1;
           else
               this.Distribution.Name = varargin{1};
               varargin(1) = [];
               this.Distribution.Param = cell2mat(varargin);
           end
       end

       function [paths, discQuadVariationPaths] = simulate(this, T, ...
                                                   numPaths, numSteps, seed)
         %> function simulates path of the Levy process

           % Default seed
           if nargin < 5
               seed = 1;
           end

           rand('seed', seed);
           randn('seed', seed);
           deltaT = T / numSteps;
```

```
            % Simulation of increments of Levy Process
            switch this.Distribution.Name
                case 'compPoisson' % Compound Poisson process with
                                   %normally distributed jumps
                    % jumps ~ Normal(mean,variance)
                    lambda = this.Distribution.Param(1);
                    mean = this.Distribution.Param(2);
                    variance = this.Distribution.Param(3);

                    % Number of jumps is Poisson distributed
                    numJumps = poissrnd(lambda * deltaT, numPaths, numSteps);

                    % Construct increments
                    increments = zeros(numPaths, numSteps);
                    incrementsQuadVar = zeros(numPaths, numSteps);
                    for i=1:numPaths
                        for j=1:numSteps
                            jumps = normrnd(mean, variance, numJumps(i,j), 1);
                            increments(i,j) = sum(jumps);
                            incrementsQuadVar(i,j) = sum(jumps.^2);
                        end
                    end
                case 'varGamma' % variance gamma process
                    % Vol of Brownian motion
                    volatility = this.Distribution.Param(1);
                    % Variance rate of the gamma time change (Mean rate = 1)
                    varRate = this.Distribution.Param(2);
                    % Drift of the Brownian motion
                    drift = this.Distribution.Param(3);

                    % Generate Gamma and Normal random variables
                    gammaRV = gamrnd(deltaT/varRate, varRate, numPaths, numSteps);
                    normalRV = normrnd(0, 1, size(gammaRV));

                    % Construct increments of the variance gamma process
                    increments = drift * gammaRV + volatility ...
                                    * sqrt(gammaRV) .* normalRV;
                    incrementsQuadVar(i,j) = increments .^ 2;
            end

            % Construct paths using the simulated increments
            paths = [zeros(numPaths, 1), cumsum(increments, 2)];
            discQuadVariationPaths = [zeros(numPaths, 1), ...
                                        cumsum(incrementsQuadVar, 2)];
        end
    end

    methods (Static)
        function path = twoSidedPath(pathsLeft, pathsRight)
            pathsLeft = -pathsLeft;
            path = [pathsLeft(:, end:-1:2) pathsRight];
        end
    end
end
```

```
classdef FLP_new
    properties
        Type

        Kernel

        % Instance of a Lévy Process (LP)
        DrivingLevyProcess

        % Fractional integration parameter
        FracDiffParam

        % a -> see modified MvN-Kernel
        ShiftParameter
    end

    methods
        function this = FLP_new(drivingLevyProcess, d, type, a)
            this.DrivingLevyProcess = drivingLevyProcess;
            this.FracDiffParam = d;
            if nargin < 4
                a=0.5;
            end
            if nargin < 3
                type = 'MG';
            end
            this.Type = type;
            this.ShiftParameter = a;
            this.Kernel = this.setKernel;
        end

        function kernel = setKernel(this)
            switch this.Type
                case 'MG'
                    kernel = @(uBound,lBound,param)MolchanGolosovKernel...
                            (uBound,lBound,param,this.FracDiffParam);
                case 'mMvN'
                    kernel = @(t,s,d,a)ModifiedMandelbrotVanNess(t,s,d,a);
                case 'MvN'
                    kernel = @(t,s,d)MandelbrotVanNess(t,s,d);
            end
        end

        function [pathsFlp, pathsLp, pathsDiscQuadVar, incDrivingProcess, kernelMat]...
                 = simulate(this, T, numPaths, numStepsTime, numStepsPath, ...
                   seed, isDrivenByQuadVar, negativeStartTime)
            %
            % INPUT Parameters
            % @isDrivenByQuadVar:   Is a boolean. If isDrivenByQuadVar==true,
            %                       then the FLP is driven by the discrete
```

```
%                       part of the quadratic variation of
%                       this.DrivingLevyProcess
% @numStepsPath:    Determines the times where the FLP is
%                   simulated, i.e. FLP is simulated on the
%                   grid [0:T/numStepsPath:T]
% @numStepsTime:    Determines the grid on which the Lévy and
%                   the kernel is calculated;
%                   Must be a multiple of numStepsPath.
%
% Example: T=10, numStepsTime = 100, numStepsPath = 10
%          Path is simulated on the grid 1:1:10.
%          Therefore, for t=1:1:10
%          the kernel f(t,.) and the Lévy process is calculated
%          on the grid
%          0:1/100:t

if nargin < 6
    isDrivenByQuadVar = false;
end

if nargin < 8
    negativeStartTime = -T;
end

deltaT = T/numStepsTime;
quot = numStepsTime/numStepsPath;
negativeStartTime = -deltaT * floor(-negativeStartTime/deltaT);

switch this.Type
    case 'MG'
        calcTime = zeros(2,numStepsTime-1);

        % kernelMat(i,j) = MolchanGolosovKernel(delta*i, delta*j)
        if T <= 8500 && ~mod(T,1) && T == numStepsTime && ...
           T == numStepsPath && sum(this.FracDiffParam == ...
           [0.1 0.2 0.3 0.4]) == 1
            % If available, load precalculated values of the kernel
            switch this.FracDiffParam
                case 0.4
                    data = load('kernelData_d0.4_8500.mat');
                case 0.3
                    data = load('kernelData_d0.3_5000.mat');
                case 0.2
                    data = load('kernelData_d0.2_5000.mat');
                case 0.1
                    data = load('kernelData_d0.1_5000.mat');
            end
            kernelMat = data.kernelMat(1:numStepsTime, 1:numStepsTime);
        else
            kernelMat = zeros(numStepsPath, numStepsTime);
            tol = 10e-16;
            counter = 0;
            for j = 1:(numStepsTime-1)
                %tic
```

```
                if j >= counter * quot
                    counter = counter + 1;
                end
                for i = (quot*counter):quot:numStepsTime
                    int = this.Kernel(deltaT * i, ...
                            deltaT * max(i-quot, j), deltaT * j);
                    if i/quot==1
                        kernelMat(i/quot,j) = int;
                    else
                        kernelMat(i/quot,j) = kernelMat((i/quot-1),j)+int;
                    end
                end
                disp('------------------------------------------------');
            end
        end

        % Simulation of paths of driving process
        [pathsLp, pathsDiscQuadVar] = ...
            this.DrivingLevyProcess.simulate(T, numPaths, ...
            numStepsTime, seed);
        if  isDrivenByQuadVar == false
            incDrivingProcess = pathsLp(:, 2:end) -...
            pathsLp(:, 1:(end-1));
        else
            incDrivingProcess = pathsDiscQuadVar(:, 2:end) ...
                - pathsDiscQuadVar(:, 1:(end-1));
        end

        pathsFlp = [zeros(numPaths, 1), (kernelMat * incDrivingProcess')'];

    case {'MvN','mMvN'}

        % kernel matrix
        time = negativeStartTime:deltaT:T;
        kernelMat = zeros(numStepsPath, length(time));
        if strcmp(this.Type,'MvN')
            for i=1:numStepsPath
                kernelMat(i,:) = this.Kernel(i*quot*deltaT, time,...
                    this.FracDiffParam);
            end
        else
            for i=1:numStepsPath
                i
                kernelMat(i,:) = this.Kernel(i*quot*deltaT,...
                    time, this.FracDiffParam,this.ShiftParameter);
            end
        end

        % Simulation of two-sided driving process
         [pathsLpRightSide, pathsDiscQuadVarRightSide] =...
         this.DrivingLevyProcess.simulate(T, numPaths, numStepsTime, seed);

        [pathsLpLeftSide, pathsDiscQuadVarLeftSide] =...
         this.DrivingLevyProcess.simulate(-negativeStartTime, numPaths,...
```

```
                                -negativeStartTime/deltaT, seed+1000);

                             pathsLp = LP.twoSidedPath(pathsLpLeftSide, pathsLpRightSide);

                             pathsDiscQuadVar = LP.twoSidedPath(pathsDiscQuadVarRightSide,...
                                 pathsDiscQuadVarLeftSide);

                            if  isDrivenByQuadVar == false
                                incDrivingProcess = pathsLp(:, 2:end) - pathsLp(:, 1:(end-1));
                            else
                                incDrivingProcess = ...
                                    pathsDiscQuadVar(:, 2:end) - pathsDiscQuadVar(:, 1:(end-1));
                            end

                            % paths of FLP
                            pathsFlp = [zeros(numPaths, 1),...
                                (kernelMat(:,1:(end-1)) * incDrivingProcess')'];
                    end
                end
            end
        end



classdef FICOGARCH
   properties
    % Instance of a Lévy process
    DrivingLevyProcess

    % d
    FracDiffParam

    % Model Parameters: [alpha_0 alpha_1 beta_1]
    Parameter
   end

   methods
       function this = FICOGARCH(levyProcess, d, params)
           this.DrivingLevyProcess = levyProcess;
           this.FracDiffParam = d;
           this.Parameter = params;
       end

       function [pathsFicogarch, pathsSquaredVol, pathsReturns]= ...
               simulate(this, V0, T, numPaths, numStepsTime, numStepsPath,...
               seed, type,a,negativeStartTime)
           if nargin < 10
                negativeStartTime = -T;
           end

           if nargin < 8
                type = 'MG';
           end
           if nargin < 9
```

```
                 a=1;
            end

            alpha0 = this.Parameter(1);
            alpha1 = this.Parameter(2);
            beta1 = this.Parameter(3);

            if strcmp(type, 'mMvN') || strcmp(type, 'MG')
                isDrivenByQuadVar = true;
                flp = FLP(this.DrivingLevyProcess, this.FracDiffParam, type, a);
                [pathsFlp, pathsLp, pathsDiscQuadVar, incDrivingProcess, kernelMat]...
                    = flp.simulate(T, numPaths, numStepsTime, numStepsPath, seed,...
                      isDrivenByQuadVar,negativeStartTime);
                incrementsFLP = pathsFlp(:, 2:end) - pathsFlp(:, 1:(end-1));
            else
                % For the COGARCH the increments (subsequently called
                % incrementsFLP) are given by the discrete
                % part of the quadratic variation
                [pathsLp, discQuadVariationPaths] = this.DrivingLevyProcess.simulate...
                    (T, numPaths, numStepsPath, seed);
                incrementsFLP = discQuadVariationPaths(:,2:end) ...
                    - discQuadVariationPaths(:, 1:end-1);
            end
            deltaT = T / numStepsPath;

            pathsSquaredVol = zeros(numPaths, numStepsPath + 1);
            pathsSquaredVol(:, 1:2) = ones(numPaths, 2) * V0;

            if strcmp(type, 'mMvN') || strcmp(type, 'COGARCH')
                for i = 1:(numStepsPath-1)
                   % Euler Step
                   pathsSquaredVol(:,i+2) = pathsSquaredVol(:,i+1) ...
                       + (beta1 * alpha0 * ones(numPaths, 1) ...
                       - beta1 * pathsSquaredVol(:,i+1)) * deltaT ...
                       + alpha1 * pathsSquaredVol(:,i) .* incrementsFLP(:,i+1);
                end
            else
                for i = 1:numStepsPath
                   % Euler Step
                   pathsSquaredVol(:,i+1) = pathsSquaredVol(:,i) ...
                             + (beta1 * alpha0 * ones(numPaths, 1) - beta1 ...
                             * pathsSquaredVol(:,i)) * deltaT ...
                             + alpha1 * pathsSquaredVol(:,i) .* incrementsFLP(:,i);
                end
            end

            if strcmp(type, 'COGARCH')
                incLp = pathsLp(:,2:end)-pathsLp(:,1:end-1);
            elseif strcmp(type, 'MG')
                quot = numStepsTime / numStepsPath;
                pathsLp = [pathsLp(1) pathsLp((1+quot):quot:end)];
                incLp = pathsLp(:,2:end)-pathsLp(:,1:end-1);
            elseif strcmp(type, 'mMvN')
                quot = numStepsTime / numStepsPath;
```

```
                pathsLp = pathsLp(end-numStepsTime:end);
                pathsLp = [pathsLp(1) pathsLp((1+quot):quot:end)];
                incLp = pathsLp(:,2:end)-pathsLp(:,1:end-1);
            end
            pathsReturns = sqrt(pathsSquaredVol(:,1:(end-1))) .* incLp;

            pathsFicogarch = cumsum(pathsReturns);
        end
    end
end
```

B.2 Auxiliary functions

```
function value = PositivePart(x,d)
value = x.^d;
value(isinf(value))=0;
value(~imag(value)==0)=0;
value = value .* (x>0);
end


function output = MolchanGolosovKernel(uBound, lBound, param, d)
    % Calculation of:
    % constant * int_lBound^uBound((u-param).^(d-1).*(u/param).^d) du

    tol = 10e-8;
    if uBound<0 || ~(d>0 && d<0.5)
        warning('Kernel is not well defined');
        output=-1;
    elseif param>uBound || param<0 || param == uBound || uBound < tol
        output = 0;
    else
     constant =  d^2 * (2*d+1)*gamma(1-d) / gamma(1+d) / gamma(1-2*d);
     constant = sqrt(constant);

     kernel = @(u)((u-param).^(d-1).*(u/param).^d);
     int =  quadgk(@(u)kernel(u),lBound,uBound,'RelTol',0,...
         'AbsTol',1e-12, 'MaxIntervalCount', 1886);

     output = constant * int;
    end
end


function f = MandelbrotVanNess(t,s,d)
const = 1/(gamma(1+d));
f = (PositivePart(t-s,d) - PositivePart(-s,d))*const;
end


function f = ModifiedMandelbrotVanNess(t,s,d,a)
if d>=0 || a<0
```

```
    warning('Kernel is not well defined');
    exit function;
end
const = gamma((-d)^0.8);
f = ((a+max(-s,0)).^d - (a+max(t-s,0)).^d).* const;
end
```

Bibliography

D. Applebaum. *Lévy Processes and Stochastic Calculus*. Cambridge University Press, 2004.

Baillie et al. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, 74(1):3–30, 1996.

R.T. Baillie. Long memory processes and fractional integration in econometrics. *J. Econometrics*, 73(1):5–59, 1996.

A. Behme and A. Lindner. Multivariate generalised Ornstein-Uhlenbeck processes. *Stochastic Processes and their Applications*, 122(4):1487–1518, 2012.

A. Benassi, S. Cohen, and J. Istas. On roughness indices for fractional fields. *Bernoulli*, 10(3):357–373, 2004.

C. Bender and T. Marquardt. Integrating volatility clustering into exponential Lévy models. *Journal of Applied Probability*, 46:609–628, 2009.

C. Bender, A. Lindner, and M. Schicks. Finite variation of fractional Lévy processes. *Journal of Theoretical Probability*, 25(2):594–612, 2012.

Beran et al. *Long-Memory Processes*. Springer, 1st edition, 2013.

T. Bollerslev. Generalized autoregressive conditional heteroscedasticity. *J. Econometrics*, 31(3):307–327, 1986.

P. Bougerol and N. Picard. Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics*, 52:115–127, 1992.

G. E. P. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control*. Holden Day, 1970.

P. Brockwell. Continuous-time ARMA processes. *Handbook of Statistics*, 19:249–276, 2001a.

P. Brockwell. Lévy driven CARMA processes. *Annals of the Institute of Statistical Mathematics*, 53(1):113–124, 2001b.

P. Brockwell. Representations of continuous-time ARMA processes. *Journal of Applied Probabilities*, pages 375–382, 2004.

P. Brockwell and T. Marquardt. Lévy-driven and fractionally integrated ARMA processes with continuous time parameter. *Statistica Sinica*, 15:477–494, 2005.

P. Brockwell, E. Chadraa, and A. Lindner. Continuous-time GARCH processes. *The Annals of Applied Probability*, 16(2):790–826, 2006.

P.J. Brockwell and R.A. Davis. *Time Series: Theory and Models*. Springer, 2nd edition, 1991.

C. Conrad and B. R. Haag. Inequality constraints in the fractionally integrated GARCH model. *Journal of Financial Econometrics*, 4(3):413–449, 2006.

J. Davidson. Moment and memory properties of linear conditional heteroscedasticity models, and a new model. *Journal of Business and Economic Statistics*, 22(1):16–29, 2004.

R. Douc, F. Roueff, and P. Soulier. On the existence of some ARCH(∞) processes. *Stochastic Processes and their Applications*, 118:755–761, 2008.

S. Engelke and J. H. C. Woerner. A unifying approach to fractional Lévy processes. *Stochastics and Dynamics*, 13, 2013.

R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1006, 1982.

R. F. Engle and T. Bollerslev. Modelling the persistence of conditional variances. *Econometric Reviews*, 5(1):1–50, 1986.

H. Fink. Conditional characteristic functions of Molchan-Golosov fractional Lévy processes with applications to credit risk. *Journal of Applied Probability*, 50:983–1005, 2013.

Giraitis et al. Stationary ARCH models: Dependence structure and central limit theorem. *Econometric Theory*, 16:3–22, 2000.

C.W.J. Granger and R. Joyeux. An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1):15–29, 1980.

J.R.M. Hosking. Fractional differencing. *Biometrika*, 68(1):165–176, 1981.

J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*. Springer, 2002.

C. Jost. Integral Transformations of Volterra Gaussian Processes. *Dissertation, University of Helsinki*, 2007.

V. Kazakevicius and R. Leipus. A new theorem on the existence of invariant distributions with applications to ARCH processes. *J. Appl. Probab.*, 40:147–162, 2003.

C. Klüppelberg, A. Lindner, and R. Maller. A continuous-time GARCH process driven by a lévy process: stationarity and second-order behaviour. *Journal of Applied Probability*, 41(3):601–638, 2004.

P. Kokozka and R. Leipus. Change-point estimation in ARCH models. *Bernoulli*, 6(3): 513–539, 2000.

A. M. Lindner. Continuous time approximations to GARCH and stochastic volatility models. *In: Handbook of Financial Time Series*, pages 481–496, 2009.

T. Marquardt. Fractional lévy processes with an application to long memory moving average processes. *Bernoulli*, 12(6):943–1126, 2006.

T. Mikosch and C. Starica. Long range dependence effects and ARCH modeling. *Theory and Applications of Long Range Dependence*, pages 439–460, 2003.

D. B. Nelson. ARCH models as diffusion approximations. *Journal of Economatrics*, 45 (1-2):7–38, 1990.

A. B. O'Connor and J. Rosinski. Characterization of the finite variation property for a class of stationary increment infinitely divisible processes. *Stochastic Processes and their Applications*, 123(6):1871–1890, 2013.

A. B. O'Connor and J. Rosinski. Structure of infinitely divisible semimartingales. *arXiv:1209.1644*, 2014.

M.B. Priestley. *Spectral Analysis and Time Series*. Academic Press, 1981.

P.E. Protter. *Stochastic Integration and Differential Equations*. Springer, 2004.

B.S. Rajput and J. Rosinski. Spectral representation of infinitely divisible processes. *Probability Theory and Related Fields*, 82:451–487, 1989.

J. Rosinski. On path properties of certain infinitely divisible processes. *Stochastic Processes and their Applications*, 33:73–87, 1989.

S. G. Samko, A. Kilbas, and O. I. Marichev. *Fractional Integrals and Derivatives. Theory and Applications*. Gordon and Breach Science Publishers, 1993.

K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, 1999.

N. Shephard. Stochastic volatility. *In: Andersen, T.G., Davis, R.A., Kreiß , J.-P.and Mikosch, T. (eds.): Handbook of Financial Time Series.*, 2008.

H.J. Tikanmäki and Y. Mishura. Fractional Lévy processes as a result of compact interval integral transformation. *Stochastic Analysis and Applications*, 29:1081–1101, 2011.

P. Zaffaroni. Stationarity and memory of ARCH(∞) models. *J. Econometrics*, 20:147–160, 2004.

M. Zähle. Integration with respect to fractal functions and stochastic calculus. *Probability Theory and Related Fields*, 111:333–374, 1998.