University of Milano-Bicocca

Department of Statistics and Quantitative Methods

Ph.D. School in Statistics and Mathematical Finance
Ph.D. in Statistics and Applications

COGARCH Processes: Theory and Asymptotics for the Pseudo-Maximum Likelihood Estimator

Supervisors:
Prof. Ilia Negri
Prof. Claudia Klüppelberg

Author:
Mauro Iannace

Academic Year 2013-2014

COGARCH Processes: Theory and Asymptotics for the Pseudo-Maximum Likelihood Estimator

Mauro Iannace

December 12, 2014

Abstract

In order to capture the so-called stylized facts and model high-frequency and irregularly time spaced financial data continuous time GARCH processes are becoming popular. In 2004 Klüppelberg, Lindner and Maller introduced the COGARCH model as a continuous time analogue to the successful GARCH model. Like the GARCH process, the COGARCH is based on a single source of randomness, which is a driving Lévy process. Once introduced Lévy processes and stochastic calculus for semimartingales we go into detail to discuss some properties of the COGARCH process.
Motivated by the fact that many data are asymmetric we also study some extensions; in particular a continuous time GJR-GARCH is analysed. We go on to focus on markovianity, stability, stationarity and moments. These are prerequisites for proposing a new version of the pseudo-maximum likelihood estimator, which is, under some regularity conditions, consistent. Finally the empirical quality of our estimator is investigated in a simulation study based on a comparison with the method of moments.

Acknowledgments

First of all I would like to thank my advisors Prof. Ilia Negri and Prof. Claudia Klüppelberg for having confidence in me despite my non-mathematical background. I express my gratitude to them for giving me support, encouragement and advice. I am grateful to them, that introduced me to the theory of Lévy processes and their application to continuous time financial time series. Without their precious help and inspiration this thesis would have never been written. Many thanks also to the Department of Mathematics-Chair of Mathematical Statistics, for the possibility to work on my research project at Technical University of Munich with Prof. Claudia Klüppelberg.

I would like to thank my colleagues at Milano-Bicocca University, in particular Fede and Failo for the very pleasant working atmosphere and the amazing time we spend together in our free time. I also am thankful to Jan, Marina and Zara for their friendship, which gives me support in bad times and laughs in good times.
Finally, I would like to thank my family for providing the opportunity to do all this.

Contents

List of Figures

Introduction

Stochastic volatility is the main concept used in Financial Econometrics and Mathematical Finance to deal with the random and time-varying dispersion of returns of a given asset or market index. We need it to take decisions concerning risk analysis, portfolio selection and derivative pricing. It was also clear to the funding fathers of the modern continuous time finance that homogeneity was not realistic; Black and Scholes wrote in [10]: "*there is evidence of non-stationarity in the variance. More work must be done to predict variances using the available information.*" In financial time series the volatility clustering phenomenon is observed too. As Mandelbrot [37] noted, large changes tend to be followed by large change, of either sign, and small changes are usually followed by small changes. Therefore extreme returns tend to cluster together. Empirical evidence also shows that, while returns are uncorrelated, absolute returns or their squares display a positive, significant and slowly decaying autocorrelation function. Moreover returns have fat tails such that their distributions are leptokurtic. Such phenomena intrigued many researchers and oriented the development of stochastic models in finance to model the so-called *stylized facts*. Engle [15] first developed the famous ARCH (autoregressive conditionally heteroscedastic) model and then Bollerslev [11] generalized this process obtaining the GARCH (generalized ARCH) model. Processes with generalized autoregressive conditional heteroscedasticity provide the volatility by means of the previous values of the process. These kinds of models capture the main characteristics of financial data.
For a long period volatility and prices were modeled with discrete time models, but over the last years the frequency of data has increased. This growing amount of available data is called *high-frequency data*. Taking the data, which can be unequally spaced in time, only at fixed time intervals neglects some of the information. Then, for modelling this huge amount of data it seems natural to model price and volatility processes in continuous time, which should also reflect as many as possible the well-known stylized facts.
A first approach goes back to Nelson [43], who extended the discrete time GARCH process by making diffusion approximations. This process leads to a stochastic volatility model driven by two independent Brownian motions, in contrast with GARCH models driven by only one source of randomness. Real data show that volatility and price processes also have jumps, which cannot be captured by Nelson model. In order to capture that Barndorff-Nielsen and Shephard [4] introduced a new model where the volatility is an Ornstein-Uhlenbeck process driven by a Lévy process. Modelling jumps with this model is possible, but it still contains two sources of randomness. Therefore Klüppelberg, Lindner and Maller introduced in [27] a new continuous time GARCH (COGARCH) process driven by a single Lévy process. Such a model preserves the

structure and the main characteristics of the discrete time GARCH model.
Another typical characteristic of financial data is the so-called *leverage effect*, i.e. a negative correlation between current returns and future volatility. Like in the discrete time case, the COGARCH process cannot model this phenomenon. Therefore, extensions, based on the discrete time GJR-GARCH and APGARCH models developed in [20] and [45], were recently introduced in [31] and [5].
Concerning the estimation of COGARCH parameters, Haug, Klüppelberg, Lindner and Zapp [23] proposed a method of moment estimator, which can be applied for equally spaced time series. Its consistency and asymptotic normality were verified under regularity conditions. In order to work with irregularly spaced time series Müller [41] considered an MCMC estimator for COGARCH models driven by a compound Poisson. No restrictions to the driven Lévy process are contemplated by the pseudo-maximum likelihood estimator, proposed by Maller, Müller and Szimayer [36], which is suitable for irregularly spaced data too. A modified version of this estimator (see [26]) has been studied together with consistency and asymptotic normality. The most recent procedure is due to Bibbona and Negri [9]; by means of higher moments they use the optimal prediction-based estimating functions method proposed in [50]. Inferential techniques for asymmetric processes are just related to the method of moments and maximum likelihood (cf. [5]).

The aim of this work is to give an overview of the continuous time GARCH processes, both symmetric and asymmetric, and study inferential procedures for the asymmetric COGARCH process via pseudo-maximum likelihood method. This thesis is divided in three chapters.

In Chapter 1 we give an introduction to the theory of Lévy processes and stochastic calculus for semimartingales. Especially we will focus on infinitely divisible laws, Lévy-Itô decomposition, properties about bounded variation and moments, subordinators and stochastic integration.

In Chapter 2 we introduce COGARCH and asymmetric COGARCH models and prove a few properties like stationarity, markovianity and moments. We will also consider examples showing simulated trajectories, for both processes, in order to see how log-prices, log-returns and volatilities behave under such models. As driving Lévy process we will choose the variance gamma process.

Chapter 3 is dedicated to estimate the asymmetric COGARCH model. Following Behme, Klüppelberg and Mayr [5] we analyse the available methods: the method of moment and the pseudo-maximum likelihood estimator. We will propose a new pseudo-maximum likelihood method, which guarantees asymptotic properties. Such method can be used with irregularly spaced time series and is based on an approximation of the continuous time asymmetric GARCH process by an embedded sequence of discrete time asymmetric GARCH series, which converges in probability to the continuous time model in the Skorokhod distance. After a detailed description of this approach we will prove the consistency of the estimator (as Kim and Lee have done for the COGARCH model in [26]). We conclude with a simulation study in order to apply to simulated data the developed algorithms.

Chapter 1

Introduction to Lévy processes

This chapter is dedicated to the theory of Lévy processes. Refering to [2], [30], [48] and [46] we introduce infinitely divisible laws to go into detail about Lévy processes and stochastic calculus for semimartingales. Examples will simplify theoretical concepts and simulations give an idea about trajectories. Applications to Finance and Actuarial Sciences can be found in [49], [33] and [40].

1.1 Infinitely divisible distributions

Definition 1.1. A probability measure μ on $\mathbb{R}^d$ endowed with the Borel σ-field $\mathcal{B}(\mathbb{R}^d)$ is infinitely divisible[1] if $\forall\, n \in \mathbb{N}$ there exists another probability measure μ_n such that

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

i.e. if μ has a convolution n-root.

Remark 1. The law μ_X of a random variable X which takes values in $\mathbb{R}^d$ is infinitely divisible if for all $n \in \mathbb{N}$ μ_X has a convolution n-root. Equivalently X is infinitely divisible if for all $n \in \mathbb{N}$ there exist i.i.d. random variables $X_1^{(n)}, \dots, X_n^{(n)}$ such that

$$X \stackrel{\mathrm{d}}{=} X_1^{(n)} + \cdots + X_n^{(n)}.$$

Moreover by Kac's theorem[2] the law of X is infinitely divisible if, for all $n \in \mathbb{N}$, there exists a random variable $X^{(n)}$ such that

$$\phi_X(u) = \left[\phi_{X^{(n)}}(u)\right]^n,$$

where $\phi_X(u) := \mathbb{E}(\mathrm{e}^{i\langle u, X\rangle})$ is the characteristic function of X.

Example 1.1. (Multivariate Gaussian r.v.) Let X be a gaussian $\mathbb{R}^d$-valued random vector, with density with respect to the Lebesgue measure on $\mathbb{R}^d$ given by

$$\mu_X(\mathrm{d}x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{-1/2}} \exp\left(-\frac{1}{2}\langle x - m, \Sigma^{-1}(x-m)\rangle\right) \mathbf{1}_{\mathbb{R}^d}(x)\mathrm{d}x$$

[1] De Finetti was the first to introduce the notion of infinitely divisible distribution.

[2] The $\mathbb{R}^d$-valued random variables $X_1, \dots, X_n$ are independent if and only if $\mathbb{E}\mathrm{e}^{i\sum_{j=1}^n \langle u_j, X_j\rangle} = \phi_{X_1}(u_1)\cdots\phi_{X_n}(u_n)$ for all $u_1, \dots, u_n \in \mathbb{R}^d$.

for all $m \in \mathbb{R}^d$ and strictly[3] positive definite symmetric $d \times d$ matrix Σ. It is well known that

$$\phi_X(u) = \exp\left(i\langle m, u\rangle - \frac{1}{2}\langle u, \Sigma u\rangle\right) = \left[\exp\left(i\langle \frac{m}{n}, u\rangle - \frac{1}{2}\langle u, \frac{\Sigma}{n} u\rangle\right)\right]^n,$$

so we see that X is infinitely divisible, where $\phi_{X^{(n)}}$ is the characteristic function of a normal vector with expectation m/n and covariance matrix Σ/n.

Example 1.2. (Poisson r.v.) We consider the Poisson distribution with $d = 1$ taking values in the set $\mathbb{N}_0$. X is Poisson distributed if its law, absolutely continuous respect to the counting measure, is such that

$$\mu_X(\{x\}) = \mathrm{e}^{-\lambda} \frac{\lambda^x}{x!} \mathbf{1}_{\mathbb{N}_0}(x)$$

for all $\lambda > 0$. It is easy to verify that this law is infinitely divisible too, indeed

$$\phi_X(u) = \exp\left(\lambda(\mathrm{e}^{iu} - 1)\right) = \exp\left(\lambda/n(\mathrm{e}^{iu} - 1)\right)^n.$$

Example 1.3. (Compound Poisson r.v.) Let $(J_k)_{k\in\mathbb{N}}$ a sequence of i.i.d. and $\mathbb{R}^d$-valued random variables with common law μ_J independent of N which is Poisson distributed. The random variable $X := \sum_{k=1}^N J_k$ is called compound Poisson random variable and we can think of it as a random walk with a random number of steps controlled by a Poisson random variable. Conditioning on N, using independence and fixing $J_0 = 0$ we proceed to calculate the characteristic function

$$\begin{aligned}
\phi_X(u) &= \mathbb{E}(\mathbb{E}(\mathrm{e}^{iuX}|N)) \\
&= \sum_{n=0}^{\infty} \mathbb{E}(\mathrm{e}^{iuX}|N = n)\mu_N(\{n\}) \\
&= \sum_{n=0}^{\infty} (\phi_J(u))^n \,\mathrm{e}^{-\lambda} \frac{\lambda^n}{n!} \\
&= \mathrm{e}^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda\phi_J(u))^n}{n!} = \exp[\lambda(\phi_J(u) - 1)].
\end{aligned}$$

Similarly to the Poisson case we can prove also for the compound Poisson the infinite divisibility.

Other examples of infinitely divisible laws are the gamma, negative binomial, Cauchy and strictly stable distributions. Counter-examples are the binomial and uniform distributions. In particular every random variable with bounded range is not infinitely divisible, unless is constant.

Proposition 1.1.1. *Let X and Y be independent and infinitely divisible random numbers. Then the same holds for $X + Y$, $-X$ and $X - Y$.*

[3] We avoid the definite positive case because the density could not exist if the matrix were singular.

Proof. By hypothesis μ_X and μ_Y are infinitely divisible probability measures, so by characterization via characteristic function we have

$$\phi_{X+Y}(u) = \phi_X(u)\phi_Y(u) = [\phi_{X^{(n)}}(u)\phi_{Y^{(n)}}(u)]^n,$$

hence the convolution $\mu_X * \mu_Y$ is infinitely divisible. Similarly for the second point; it is enough to observe that $\phi_{-X}(u) = \overline{\phi_X(u)}$. About the third part we notice that $\mu_{X-Y} = \mu_X * \mu_{-Y}$ and $\phi_{X-Y} = \phi_X \overline{\phi_Y} = |\phi_X|^2$, and the result follows. □

About infinitely divisible laws it is useful to analyze the characteristic function behavior. For this reason we state and prove the following theorem.

Theorem 1.1.2. *The characteristic function of an infinitely divisible law never vanishes.*

Proof. We know that if ϕ_X is the characteristic function of an infinitely divisible law μ_X, then the same holds for $|\phi_X|^2$ that is the characteristic function of the convolution $\mu_X * \mu_{-Y}$, where Y is an independent copy of X. Hence if $\mu_{X,n}$ is the n-root of μ_X then $\mu_{X,n} * \mu_{-Y,n}$ is the n-root of $\mu_X * \mu_{-Y}$ and $|\phi_{X,n}|^2 = |\phi_X|^{2/n}$. We notice that $\forall u \in \mathbb{R}^d$

$$\psi(u) := \lim_{n\to\infty} |\phi_{X,n}(u)|^2 = \lim_{n\to\infty} |\phi_X(u)|^{2/n} = \mathbf{1}_{\{\phi_X(u)\neq 0\}}.$$

As $\phi_X(0) = 1$ and ϕ is continuous, there exists a positive ϵ such that for $u \in (-\epsilon, \epsilon)$ $\phi_X(u) \neq 0$. Then for $u \in (-\epsilon, \epsilon)$ $\psi(u) = 1$, i.e. it is continuous at 0. In particular, by Lévy theorem, it is a characteristic function and by continuity of the characteristic function, as it takes values from the set $\{0, 1\}$, it must be $\psi(u) = 1$ $\forall u$. Consequently for all $u \in \mathbb{R}^d$ $\phi_X(u) \neq 0$. □

Proposition 1.1.3. *Let $(\mu_k)_{k\in\mathbb{N}}$ a sequence of infinitely divisible probability measures. If μ_k converges weakly to μ as $k \to \infty$, then μ is infinitely divisible.*

Proof. μ_k converges weakly to μ, then $\phi_k \to \phi$ as $k \to \infty$, which implies $\phi_k^{1/n} \to \phi^{1/n}$ for every $n \in \mathbb{N}$ when $k \to \infty$. We know that $\phi_k^{1/n}$ is a characteristic function as ϕ_k is such that for $u \in \mathbb{R}^d$ $\phi_k(u) = (\phi_{k,n}(u))^n$, where $\phi_{k,n}$ is a characteristic function for every n. Moreover $\phi^{1/n}$ is continuous at 0, then it is a characteristic function by Lévy continuity theorem. Finally $\phi = (\phi^{1/n})^n$, hence ϕ is the characteristic function of a infinitely divisible probability measure. □

Proposition 1.1.4. *Any infinitely divisible probability law can be obtained as weak limit of a sequence of compound Poisson distributions.*

Proof. Let ϕ_X be the characteristic function of an arbitrary infinitely divisble probability measure μ_X, so that $\phi_X^{1/n}$ is the characteristic function of $\mu_{X^{(n)}}$. We define

$$\phi_n(u) = \exp\left[n\left(\phi_X^{1/n}(u) - 1\right)\right]$$

so that ϕ_n is the characteristic function of a compound Poisson random variable for $n \in \mathbb{N}$. Then, since $\mathrm{e}^x - 1 \sim x$ as $x \to 0$

$$\phi_n(u) = \exp[n(\mathrm{e}^{(1/n)\log(\phi_X(u))} - 1)] \sim \exp[\log(\phi_X(u))] = \phi_X(u)$$

as $n \to \infty$ and the result follows by Glivenko theorem[4]. □

Corollary 1.1.5. *The set of all infinitely divisible probability measures on* $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ *coincides with the weak closure of the set of all compound Poisson laws on* $\mathbb{R}^d$.

Proof. It follows from Propositions 1.1.3 and 1.1.4. □

1.1.1 Characterisation via Lévy-Khintchine formula

Theorem 1.1.6. *(Lévy-Khintchine) A probability distribution* μ_X *is infinitely divisible if and only if there exist a vector* $b \in \mathbb{R}^d$, *a positive definite symmetric* $d \times d$ *matrix* C *and a measure* ν *on* $\mathbb{R}^d$ *satisfying* $\nu(\{0\}) = 0$ *and* $\int_{\mathbb{R}^d}(1 \wedge |x|^2)\nu(\mathrm{d}x) < \infty$, *such that, for all* $u \in \mathbb{R}^d$

$$\phi_X(u) = \exp\left\{ i\langle b, u\rangle - \frac{1}{2}\langle u, Cu\rangle + \int_{\mathbb{R}^d} [\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle \mathbf{1}_B(x)]\nu(\mathrm{d}x) \right\}$$

where $B = B_1(0)$ *is the unit ball of* $\mathbb{R}^d$.

Proof. We only prove the sufficient condition (a complete proof can be found in Sato); so we need to show that the right-hand side is a characteristic function. Consider a monotonic decreasing to zero sequence a_n in $\mathbb{R}^d$ and define the following function

$$\begin{aligned}\phi_n(u) = &\exp\left[i\langle b - \int_{[-a_n,a_n]^c \cap B} x\nu(\mathrm{d}x), u\rangle - \frac{1}{2}\langle u, Cu\rangle \right] \times \\ &\times \exp\left[\int_{[-a_n,a_n]^c} (\mathrm{e}^{i\langle u,x\rangle} - 1)\nu(\mathrm{d}x) \right].\end{aligned}$$

This represents the characteristic function of the convolution of a gaussian law with an independent compound Poisson distribution (with intensity $\lambda_n = \nu([-a_n, a_n]^c)$ and jump magnitude $\tilde{\nu}(\mathrm{d}x) = \frac{\mathbf{1}_{[-a_n,a_n]^c}(x)\nu(\mathrm{d}x)}{\nu([-a_n,a_n]^c)}$). Moreover it is clear that

$$\lim_{n\to\infty} \phi_n(u) = \phi_X(u).$$

In order to apply Lévy theorem we show now that this limit function is continuous at zero; this boils down to proving for each $u \in \mathbb{R}^d$ the continuity at zero of

$$\begin{aligned}\psi(u) &= \int_{\mathbb{R}^d} [\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle \mathbf{1}_B(x)]\nu(\mathrm{d}x) \\ &= \int_B (\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle)\nu(\mathrm{d}x) + \int_{B^c} (\mathrm{e}^{i\langle u,x\rangle} - 1)\nu(\mathrm{d}x).\end{aligned}$$

We use Taylor expansion for $\mathrm{e}^{i\langle u,x\rangle} = 1 + i\langle u, x\rangle - \frac{1}{2}\langle u, x\rangle^2 + \cdots$ in order to obtain

$$\begin{aligned}|\psi(u)| &\leq \int_B |\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle|\nu(\mathrm{d}x) + \int_{B^c} |\mathrm{e}^{i\langle u,x\rangle} - 1|\nu(\mathrm{d}x) \\ &\leq \int_B \frac{1}{2}|\langle u, x\rangle^2|\nu(\mathrm{d}x) + \int_{B^c} |\mathrm{e}^{i\langle u,x\rangle} - 1|\nu(\mathrm{d}x).\end{aligned}$$

[4]If ϕ_n and ϕ respectively are for $n \in \mathbb{N}$ the characteristic functions of the probability distributions μ_n and μ, then if $\phi_n(u) \to \phi(u)$ for all $u \in \mathbb{R}^d$ when $n \to \infty$, then μ_n converges weakly to μ as $n \to \infty$.

Using now hyphotesis about ν, Cauchy-Schwarz inequality to find a bound for $\langle u,x\rangle$ and dominated convergence theorem in the second integral

$$|\psi(u)| \leq \frac{|u|^2}{2}\int_B |x|^2\nu(\mathrm{d}x) + \int_{B^c} |\mathrm{e}^{i\langle u,x\rangle} - 1|\nu(\mathrm{d}x) \to 0$$

as $u \to 0$. We proved that ϕ_X is a characteristic function via Lévy theorem; then the infinite divisibility follows applying Glivenko thereom and Proposition 1.1.3. □

Remark 2. The triplet (b, C, ν) is called *Lévy triplet* or *characteristic triplet* and the exponent $\eta : \mathbb{R}^d \to \mathbb{C}$

$$\eta(u) = i\langle b,u\rangle - \frac{1}{2}\langle u, Cu\rangle + \int_{\mathbb{R}^d} [\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u,x\rangle \mathbf{1}_B(x)]\nu(\mathrm{d}x)$$

is called *Lévy exponent* or *characteristic exponent* (or *symbol*). Since for all $u \in \mathbb{R}^d$ $|\phi(u)| \leq 1$ for any probability measure, if μ is infinitely divisible then $\mathcal{R}(\eta(u)) \leq 0$, where $\mathcal{R}(\cdot)$ is the real part. Moreover $b \in \mathbb{R}^d$ is called *drift term*, C *gaussian* or *diffusion coefficient* and ν *Lévy measure.*

Remark 3. The Lévy measure[5] is such that $\nu(\{0\})$ and

$$\int_{\mathbb{R}^d} (1 \wedge |x|^2)\nu(\mathrm{d}x) = \nu\big(\{|x|^2 \geq 1\}\big) + \int_{|x|^2<1} |x|^2\nu(\mathrm{d}x) < \infty. \tag{1.1}$$

These two conditions are sufficient to ensure that integral in the Lévy-Khintchine formula converges, since the integrand function is $O(1)$ if $|x|^2 \geq 1$ ($|\mathrm{e}^{i\langle u,x\rangle}| = 1$) and $O(|x|^2)$ if $|x|^2 < 1$ ($|\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u,x\rangle\mathbf{1}_{\{|x|<1\}}| \leq \frac{1}{2}\langle u,x\rangle^2 \leq \frac{1}{2}|u|^2|x|^2$).
We know that $\nu(\{|x|^2 \geq 1\}) < \infty$, but $\nu(\{|x|^2 < 1\})$ can be finite or infinite. If the Lévy measure is finite then $\nu(\{|x|^2 < 1\}) < \infty$, but if $\nu(\{|x|^2 < 1\}) = \infty$ then (1.1) implies that $\nu(\{|x| \geq \epsilon\}) < \infty$ and $\nu(\{|x| < \epsilon\}) = \infty$ $\forall\epsilon \in (0,1)$. This fact follows by the following inequalities

$$\nu(\{|x| \geq \epsilon\}) = \int_{|x|\geq\epsilon} \frac{|x|^2}{1+|x|^2}\frac{1+|x|^2}{|x|^2}\nu(\mathrm{d}x) \leq \frac{1+\epsilon^2}{\epsilon^2}\int_{\mathbb{R}^d} \frac{|x|^2}{1+|x|^2}\nu(\mathrm{d}x) < \infty$$

since the map $y \mapsto \frac{1+y^2}{y^2}$ is decreasing on $\mathbb{R}^+$. Therefore this measure is always bounded on sets which are disjoint with respect to a neighbourhood of the origin, but the measure can be infinite in a neighbourhood of zero. It is clear that any finite measure on $\mathbb{R}^d$ is a Lévy measure and furthermore any Lévy measure is a σ-finite measure.

---

[5] An alternative definition for the Lévy measure on $\mathbb{R}^d\setminus\{0\}$ is given by the equivalent characterisation

$$\int_{\mathbb{R}^d\setminus\{0\}} \frac{|x|^2}{1+|x|^2}\nu(\mathrm{d}x) < \infty$$

as $\forall x \in \mathbb{R}^d$

$$\frac{|x|^2}{1+|x|^2} \leq 1 \wedge |x|^2 \leq \frac{2|x|^2}{1+|x|^2}.$$

Example 1.4. It is straightforward to obtain characteristic triplets for Gaussian, Poisson and compound Poisson distributions. If X has gaussian law then $b = \mathbb{E}X$, C is the covariance matrix and $\nu = 0$. In particular an infinitely divisible probability is gaussian if and only if $\nu = 0$. For the Poisson case $b = 0$, $c = 0$ and $\nu = \lambda\delta_1$. In effect

$$\exp\left[\int_{\mathbb{R}} (e^{iux} - 1 - iux\mathbf{1}_{\{|x|<1\}})\lambda\delta_1(\mathrm{d}x)\right] = \exp[\lambda(\mathrm{e}^{iu} - 1)].$$

Under the hyphotesis that X is compound Poisson distributed since

$$\int_{\mathbb{R}^d} i\langle u, x\rangle \mathbf{1}_B(x)\nu(\mathrm{d}x) = i\langle u, \int_{\mathbb{R}^d} x\mathbf{1}_B(x)\nu(\mathrm{d}x)\rangle$$

then $b = \int_{\mathbb{R}^d} x\mathbf{1}_B(x)\nu(\mathrm{d}x)$, $C = 0$ and $\nu = \lambda\mu_J$.

Remark 4. If the Lévy measure is such that $\int_B |x|\nu(\mathrm{d}x) < \infty$ then the symbol η can be written in the following way

$$\eta(u) = i\langle\tilde{b}, u\rangle - \frac{1}{2}\langle u, Cu\rangle + \int_{\mathbb{R}^d} (\mathrm{e}^{i\langle u,x\rangle} - 1)\nu(\mathrm{d}x)$$

where $\tilde{b} = b - \int_B x\nu(\mathrm{d}x)$. Moreover if μ_X has characteristic triplet $(b, 0, \nu)$ where ν is a finite measure and $b = \int_B x\nu(\mathrm{d}x)$, then μ_X is a compound Poisson law with intensity $\lambda = \nu(\mathbb{R}^d)$ and jump size $\tilde{\nu} = \nu/\lambda$. Actually

$$\begin{aligned}
\phi_X(u) &= \exp\left\{i\langle b, u\rangle - \frac{1}{2}\langle u, Cu\rangle + \int_{\mathbb{R}^d} [\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle\mathbf{1}_B(x)]\nu(\mathrm{d}x)\right\} \\
&= \exp\left[\lambda\int_{\mathbb{R}^d} (\mathrm{e}^{i\langle u,x\rangle} - 1)\tilde{\nu}(\mathrm{d}x)\right] = \exp\left[\lambda(\phi_J(u) - 1)\right].
\end{aligned} \tag{1.2}$$

Proposition 1.1.7. *The Lévy exponent η is continuous at every $u \in \mathbb{R}^d$ and such that $|\eta(u)| \le K(1 + |u|^2)$, for each $u \in \mathbb{R}^d$, where $K > 0$.*

Proof. Continuity follows by a well known result according to which if $\phi : \mathbb{R}^d \to \mathbb{C}$ is continuous and such that $\phi(0) = 1$ and $\phi(u) \neq 0$ for every u, then there exists a function $\eta : \mathbb{R}^d \to \mathbb{C}$ continuous and such that $\eta(0) = 0$ and $\mathrm{e}^{\eta(u)} = \phi(u)$. $\square$

Theorem 1.1.8. *The map $\eta : \mathbb{R}^d \to \mathbb{C}$ is a Lévy exponent if and only if it is continuous, hermitian*[6]*, conditionally positive definite*[7] *and such that $\eta(0) = 0$.*

---

[6] A map $f : \mathbb{R}^d \to \mathbb{C}$ is hermitian if $f(-u) = \overline{f(u)}$ $\forall u \in \mathbb{R}^d$.

[7] We say that $f : \mathbb{R}^d \to \mathbb{C}$ is conditionally positive definite if for all $n \in \mathbb{N}$ and $c_1, \dots, c_n \in \mathbb{C}$ such that $\sum_{j=1}^n c_j = 0$ we have

$$\sum_{i=1}^n \sum_{j=1}^n c_j\bar{c}_i f(u_j - u_i) \ge 0$$

for all $u_1, \dots, u_n \in \mathbb{R}^d$. If the complex numbers $c_1, \dots, c_n$ have no constraints then f is called positive definite.

Proof. If η is a Lévy symbol then the same holds for $t\eta$ for $t > 0$, consequently there exists a probability measure μ_t such that $\phi_{\mu_t}(u) = \mathrm{e}^{t\eta(u)}$ for $u \in \mathbb{R}^d$. Continuity follows by Proposition 1.1.7, and by Schoenberg correspondence[8] η is also hermitian and conditionally positive definite since $\mathrm{e}^{t\eta(u)}$ is a characteristic function. It is clear that $\eta(0) = 0$.
We now suppose that η is continuous, hermitian and conditionally positive definite with $\eta(0) = 0$. By Schoenberg correspondence e^{η} is positive definite, but by Bochner's theorem[9], since e^{η} is positive definite, continuous and such that $\mathrm{e}^{\eta(0)} = 1$, this function is the characteristic function of a measure μ for each $u \in \mathbb{R}^d$. Finally also η/n for $n \in \mathbb{N}$ is hermitian, conditionally positive definite, continuous and it vanishes at the origin, then $\mathrm{e}^{\eta/n}$ is a characteristic function too, hence μ is infinitely divisible. $\square$

1.2 Lévy processes

1.2.1 Definition and examples

Definition 1.2. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in T}, P)$ be a filtered probability space satisfying the usual conditions[10]. An adapted and $\mathbb{R}^d$-valued stochastic process $L = (L_t)_{t\in T}$ is called Lévy Process if the following statements are satisfied:

1. $L_0 = 0$ P-a.s.

2. $L_t - L_s$ is independent of $\mathcal{F}_s$ $\forall\ 0 \le s < t$.

3. $L_{t+s} - L_t$ has law independent of t $\forall\ s, t \in T$.

4. L is stochastically continuous, i.e. for every $t \in T$ and each $\epsilon > 0$ $\lim_{s\to t} P(|L_t - L_s| > \epsilon) = 0$.

Remark 5. Conditions 2. and 3. imply respectively that L has independent and stationary increments. Hence for $n \in \mathbb{N}$ and each $0 \le t_1 < t_2 < \cdots < t_{n+1} < \infty$ the random variables $(L_{t_{j+1}} - L_{t_j})_{1\le j\le n}$ are independent and the vector $L_t - L_s$ has the same law of L_{t-s} $\forall\ s < t$, as $L_0 = 0$. One can notice that under the first condition and the stationarity hypothesis the last condition is equivalent to $\lim_{t\downarrow 0} P(|L_t| > \epsilon) = 0$ for $\epsilon > 0$ since $\lim_{s\to t} P(|L_t - L_s| > \epsilon) = \lim_{s\to t} P(|L_{t-s}| > \epsilon) = \lim_{k\to 0} P(|L_k| > \epsilon) = 0$. In addition if X and Y are two stochastically continuous processes then $X + Y$ is still stochastically continuous. This fact follows from the well known inequality

$$P(|X_t + Y_t| > \epsilon) \le P(|X_t| > \epsilon/2) + P(|Y_t| > \epsilon/2) \quad \forall \epsilon > 0.$$

[8]The mapping $f : \mathbb{R}^d \to \mathbb{C}$ is hermitian and conditionally positive definite if and only if e^{tf} is positive definite $\forall\ t > 0$.

[9]If $\phi : \mathbb{R}^d \to \mathbb{C}$ is positive definite, continuous at the origin and such that $\phi(0) = 1$, then ϕ is a characteristic function.

[10]We assume completeness and right continuity for the filtration $(\mathcal{F}_t)_{t\in T}$. Given a complete probability space $(\Omega, \mathcal{F}, P)$, a filtration $(\mathcal{F}_t)_{t\in T}$ on $(\Omega, \mathcal{F}, P)$ is called *complete* is $\mathcal{F}_t$ contains every negligible set for all $t \in T$. Moreover $(\mathcal{F}_t)_{t\in T}$ is *right continuous* if $\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{u>t} \mathcal{F}_u$ $\forall\ t < \sup(T)$. Under these hypothesis we call the filtered probability space also *standard filtered probability space*.

The term Lévy process honours the French mathematician Paul Lévy who played an instrumental role in bringing together an understanding and characterisation of processes with stationary independent increments. In earlier literature these processes were called *additive processes* (processes with independent increments). Lévy himself referred to them as a sub-class of additive processes.

We highlight now the relationship between infinite divisibility and Lévy processes.

Proposition 1.2.1. *Given a Lévy process L, then L_t is infinitely divisible for each $t \in T$.*

Proof. For any $n \in \mathbb{N}$ and any $t \in T$

$$L_t = L_{\frac{t}{n}} + \left(L_{\frac{2t}{n}} - L_{\frac{t}{n}}\right) + \cdots + \left(L_t - L_{\frac{(n-1)t}{n}}\right).$$

By stationarity and independence of the increments $\left(L_{\frac{kt}{n}} - L_{\frac{(k-1)t}{n}}\right)_{k\geq 1}$ is a sequence of i.i.d. random variables. □

Remark 6. We proved that L_t is infinitely divisible, hence from Theorem 1.1.6 $\phi_{L_t}(u) = \mathrm{e}^{\eta(t,u)}$ for each $t \in T$ and $u \in \mathbb{R}^d$, where η is the characteristic exponent.

To better understand the future results we introduce the following useful lemma.

Lemma 1.2.2. *If $X = (X_t)_{t\in T}$ is a stochastically continuous process, then the map $t \mapsto \phi_{X_t}(u)$ is continuous for each $u \in \mathbb{R}^d$.*

Proof. We start by considering the function $x \mapsto \mathrm{e}^{i\langle u,x\rangle}$; it is continuous at the origin for each $u \in \mathbb{R}^d$. Then, fixed u, $\forall \epsilon > 0\ \exists \delta_1 > 0$ such that $\sup_{|x|<\delta_1} |\mathrm{e}^{i\langle u,x\rangle} - 1| < \epsilon/2$. By stochastic continuity and definition of limit we can find $\delta_2 > 0$ such that if $|t-s| < \delta_2$ $P(|X_t - X_s| > \delta_1) < \epsilon/4$ for each $s, t \in T$.
Hence, if we call μ_{t-s} the law of $X_t - X_s$, for $|t-s| < \delta_2$

$$\begin{aligned}
|\phi_{X_t}(u) - \phi_{X_s}(u)| &= \left|\int_\Omega \mathrm{e}^{i\langle u, X_{t-s}(\omega)\rangle} P(\mathrm{d}\omega)\right| \\
&= \left|\int_\Omega \mathrm{e}^{i\langle u, X_s(\omega)\rangle}\left[\mathrm{e}^{i\langle u, X_t(\omega) - X_s(\omega)\rangle} - 1\right] P(\mathrm{d}\omega)\right| \\
&\leq \int_{\mathbb{R}^d} |\mathrm{e}^{i\langle u,x\rangle} - 1| \mu_{t-s}(\mathrm{d}x) \\
&= \int_{B_{\delta_1}(0)} |\mathrm{e}^{i\langle u,x\rangle} - 1| \mu_{t-s}(\mathrm{d}x) + \int_{B_{\delta_1}(0)^c} |\mathrm{e}^{i\langle u,x\rangle} - 1| \mu_{t-s}(\mathrm{d}x) \\
&\leq \sup_{B_{\delta_1}(0)} |\mathrm{e}^{i\langle u,x\rangle} - 1| + 2P(|X_t - X_s| > \delta_1) < \epsilon.
\end{aligned}$$

□

We are now ready to show an important theorem about characteristic function of L_t, with $t \in T$.

Theorem 1.2.3. *If L is a Lévy process, then $\phi_{L_t}(u) = \mathrm{e}^{t\eta(u)}$ $\forall u \in \mathbb{R}^d$ and $t \in T$, where η is the characteristic exponent of L_1.*

Proof. Let introduce $\phi_u(t) := \phi_{L_t}(u)$ the characteristic function as function of t, which is continuous for every $u \in \mathbb{R}^d$ by Lemma (1.2.2). By independence and stationarity of the increments

$$\phi_u(s+t) = \mathbb{E}(\mathrm{e}^{i\langle u, L_{s+t}\rangle}) = \mathbb{E}(\mathrm{e}^{i\langle u, L_{s+t}-L_s\rangle}\mathrm{e}^{i\langle u, L_s\rangle}) = \mathbb{E}(\mathrm{e}^{i\langle u, L_{s+t}-L_s\rangle})\mathbb{E}(\mathrm{e}^{i\langle u, L_s\rangle}),$$

then $\phi_u(s+t) = \phi_u(t)\phi_u(s)$. We are looking for solutions of this functional equation, with the constraint $\phi_u(0) = 1$. It is clear that the only function satisfying these conditions is $\phi_u(t) = \mathrm{e}^{t\alpha(u)}$, where $\alpha : \mathbb{R}^d \to \mathbb{C}$. We know that L_1 is infinitely divisible, so α must be its characteristic exponent. □

We can now formulate the Lévy-Khintchine expression for a Lévy process $L = (L_t)_{t\in T}$,

$$\phi_{L_t}(u) = \exp\left\{t\left[i\langle b, u\rangle - \frac{1}{2}\langle u, Cu\rangle + \int_{\mathbb{R}^d}[\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle\mathbf{1}_B(x)]\nu(\mathrm{d}x)\right]\right\}$$

for $u \in \mathbb{R}^d$, where (b, C, ν) is the Lévy triplet of L_1.

Remark 7. If L is a Lévy process with triplet (b, C, ν), then $-L$ is a Lévy process with characteristic triplet $(-b, C, \tilde{\nu})$, where $\tilde{\nu}(A) = \nu(-A)$ $\forall A \in \mathcal{B}(\mathbb{R}^d)$. In effect

$$\begin{aligned}
&\phi_{-L_t}(u) = \mathbb{E}(\mathrm{e}^{i\langle u,-L_t\rangle}) = \phi_{L_t}(-u) = \\
&= \exp\left\{t\left[i\langle -b, u\rangle - \frac{1}{2}\langle u, Cu\rangle + \int_{\mathbb{R}^d}[\mathrm{e}^{i\langle u,-x\rangle} - 1 - i\langle u, -x\rangle\mathbf{1}_B(x)]\nu(\mathrm{d}x)\right]\right\} \\
&= \exp\left\{t\left[i\langle -b, u\rangle - \frac{1}{2}\langle u, Cu\rangle + \int_{\mathbb{R}^d}[\mathrm{e}^{i\langle u,y\rangle} - 1 - i\langle u, y\rangle\mathbf{1}_B(y)]\nu(-\mathrm{d}y)\right]\right\}
\end{aligned}$$

where $-x = y$. If we define the new measure $\tilde{\nu}$ we obtain the triplet $(-b, C, \tilde{\nu})$ for the new process $-L$, that is certainly a Lévy process, because the conditions of the Definition 1.2 are immediately verified.

For $\beta \in \mathbb{R}^d$ the process $(L_t + \beta t)_{t\in T}$ is a Lévy process too: the first three conditions are clearly satisfied, about the fourth condition we have $\forall\epsilon > 0$

$$\begin{aligned}
P(|L_t + \beta t| > \epsilon) &\leq P(|L_t| > \epsilon/2) + P(t|\beta| > \epsilon/2) \\
&= P(|L_t| > \epsilon/2) + \mathbf{1}_{\left(\frac{\epsilon}{2|\beta|},\infty\right)}(t)
\end{aligned}$$

and taking the limit as $t \to 0$ we have stochastic continuity. About its characteristic function $\forall u \in \mathbb{R}^d$

$$\phi_{L_t+\beta t}(u) = \mathbb{E}(\mathrm{e}^{i\langle u,t\beta\rangle})\mathrm{e}^{i\langle u,L_t\rangle} = \mathrm{e}^{i\langle u,t\beta\rangle}\phi_{L_t}(u),$$

hence the triplet is $(b + \beta, C, \nu)$.

Example 1.5. (Brownian motion and Poisson process) The simplest example of Lévy process is the deterministic linear drift. Other examples are the Brownian motion and the Poisson process. About Brownian motion we prove only continuity condition since the other ones are obvious. Since $B_t \sim \mathcal{N}(0, tI)$ for every $t \in T\setminus\{0\}$

$$\begin{aligned}
&P(|B_t| > \epsilon) = 1 - P(-\epsilon/\sqrt{t} \leq Z \leq \epsilon/\sqrt{t}) = 1 - \Phi(\epsilon/\sqrt{t}) + \Phi(-\epsilon/\sqrt{t}) = \\
&1 - \int_{-\infty}^{\frac{\epsilon_1}{\sqrt{t}}} \cdots \int_{-\infty}^{\frac{\epsilon_d}{\sqrt{t}}} \mu_Z(\mathrm{d}x_1, \ldots, \mathrm{d}x_d) + \int_{-\infty}^{-\frac{\epsilon_1}{\sqrt{t}}} \cdots \int_{-\infty}^{-\frac{\epsilon_d}{\sqrt{t}}} \mu_Z(\mathrm{d}x_1, \ldots, \mathrm{d}x_d) \to 0
\end{aligned}$$

as $t \to 0$, and where Z is an $\mathbb{R}^d$-valued gaussian vector with zero mean, covariance matrix I and law μ_Z. Similarly for the Poisson process on $\mathbb{R}$ $\forall \lambda > 0$ and as $t \to 0$

$$P(N_t > \epsilon) = 1 - \sum_{x=0}^{\lfloor \epsilon \rfloor} e^{-\lambda t} \frac{(\lambda t)^x}{x!} = 1 - e^{-\lambda t} - \sum_{x=1}^{\lfloor \epsilon \rfloor} e^{-\lambda t} \frac{(\lambda t)^x}{x!} \to 0.$$

By Remark 7 the linear Brownian motion $(X_t)_{t \in T} = (\mu t + \sigma B_t)_{t \in T}$ with $\sigma > 0$ and $\mu \in \mathbb{R}$ and the compensated compound Poisson process $(N_t - \lambda t)_{t \in T}$ for $\lambda > 0$ are Lévy processes too.

In the following figures one can see simulations about paths of some of these processes.

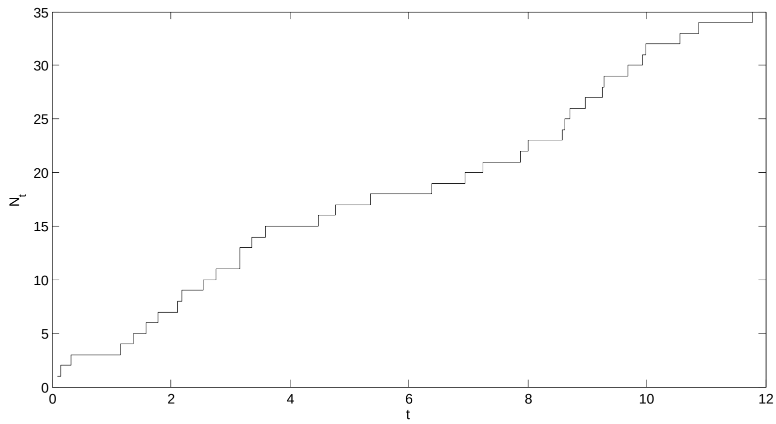

Figure 1.1: Poisson process sample path with $\lambda = 3$

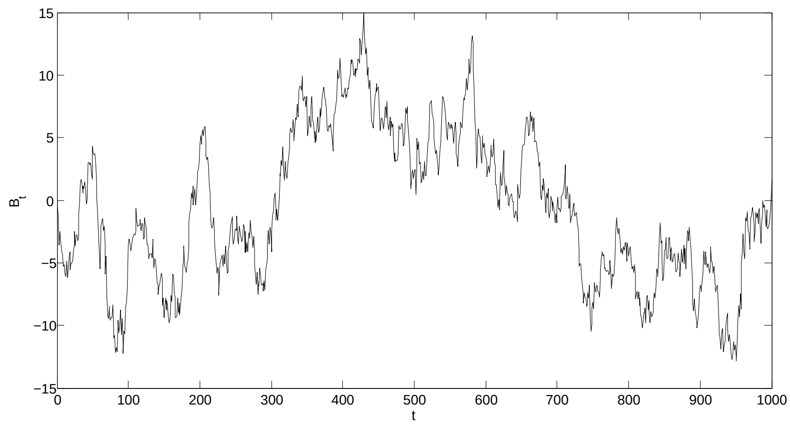

Figure 1.2: Brownian motion sample path

Example 1.6. (Compound Poisson process) Let $(J_n)_{n \in \mathbb{N}}$ and $N = (N_t)_{t \in T}$ be a sequence of i.i.d. random variables taking values in $\mathbb{R}^d$ having no atom at 0 with common

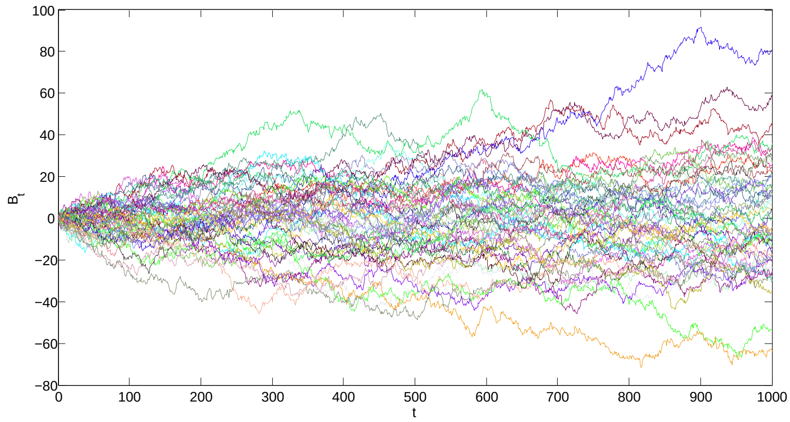

Figure 1.3: 50 Brownian motion sample paths

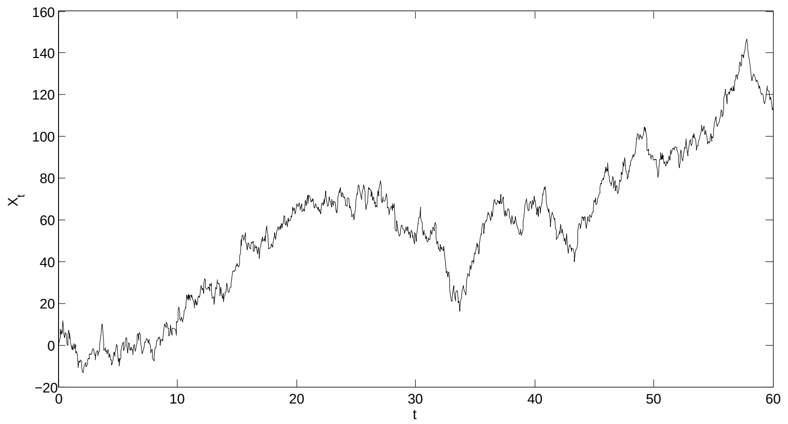

Figure 1.4: Linear Brownian motion sample path with $\sigma = 10$ and $\mu = 3$

law μ_J and a Poisson process with intensity $\lambda > 0$ independent of this sequence. The compound Poisson process[11] is defined as follows

$$Y_t := \sum_{n=1}^{N_t} J_n, \quad t \in T.$$

We have already calculated in the Example 1.3 the expression of its characteristic function

$$\phi_{Y_t}(u) = \exp\left[\lambda t(\phi_J(u) - 1)\right].$$

We now prove via definition that $Y = (Y_t)_{t \in T}$ is a Lévy process. $Y_0 = 0$ a.s. since we use the convention that for any $n \in \mathbb{N}_0$ $\sum_{n+1}^{n} = 0$. About independence and stationarity of the increments we introduce the filtration $\tilde{\mathcal{F}}_t := \sigma((N_u : u \leq t) \cup (J_n \mathbf{1}_{\{n \leq N_t\}} : n \in \mathbb{N}))$

[11] This kind of process is widely used in the classical actuarial risk process to model the total claim amount up to time t. In particular $(J_n)_{n \in \mathbb{N}}$ are the (positive) claim sizes and N_t for $t \geq 0$ is the number of claims in $[0, t]$. More details and applications can be found in Mikosh [40].

and we can see that

$$Y_t - Y_s = \sum_{n=N_s+1}^{N_t} J_n = \sum_{n=1}^{N_t-N_s} J_{N_s+n}\mathbf{1}_{\{N_s+n\leq N_t\}} = \sum_{n=1}^{N_t-N_s} J_{N_s+n}$$

and

$$Y_s = \sum_{n=1}^{\infty} J_n \mathbf{1}_{\{n\leq N_s\}}.$$

To prove independence of the increments we introduce the following lemma.

Lemma 1.2.4. *The random variables J_{N_s+n} are independent of $\tilde{\mathcal{F}}_s$ for $n \in \mathbb{N}$ and distributed as J_1.*

Proof. Let $A \in \mathcal{B}(\mathbb{R}^m)$, $B \in \mathcal{B}(\mathbb{R}^d)$ and $E \in \sigma(N_u : u \leq t)$. Then

$$\begin{aligned}
&P\left(\{(J_{N_s+1},\ldots,J_{N_s+m}) \in A\} \cap E \cap \{(J_1,\ldots J_d) \in B\} \cap \{N_s \geq i\}\right) = \\
&= \sum_{h=i}^{\infty} P\left(\{(J_{h+1},\ldots,J_{h+m}) \in A\} \cap E \cap \{(J_1,\ldots,J_d) \in B\} \cap \{N_s = h\}\right) \\
&= \sum_{h=i}^{\infty} P\left(\{(J_{h+1},\ldots,J_{h+m}) \in A\}\right) P(\{(J_1,\ldots,J_d) \in B\}) P(E \cap \{N_s = i\}) \\
&= \sum_{h=i}^{\infty} P\left(\{(J_1,\ldots,J_m) \in A\}\right) P(\{(J_1,\ldots,J_d) \in B\}) P(E \cap \{N_s = h\}) \\
&= P(\{(J_1,\ldots,J_m) \in A\}) \sum_{h=i}^{\infty} P(\{(J_1,\ldots,J_d) \in B\}) P(E \cap \{N_s = h\}) \\
&= P(\{(J_1,\ldots,J_m) \in A\}) P(\{(J_1,\ldots,J_d) \in B\}) P(E \cap \{N_s \geq i\}).
\end{aligned}$$

By choosing $E = \Omega$, $i = 0$ and $B = \mathbb{R}^d$ we obtain that the two vectors $(J_1,\ldots,J_m)$ $(J_{N_s+1},\ldots,J_{N_s+m})$ have the same law. Since the family of the events $E\cap\{(J_1,\ldots,J_d) \in B\}\cap$ $\{N_s \geq i\}$ is stable under finite intersections and generates $\tilde{\mathcal{F}}_s$ for $E \in \sigma(N_u : u \leq t)$, $A \in \mathcal{B}(\mathbb{R}^m)$, $B \in \mathcal{B}(\mathbb{R}^d)$ and $i \in \mathbb{N}_0$ the equality

$$\begin{aligned}
&P\left(\{(J_{N_s+1},\ldots,J_{N_s+m}) \in A\} \cap E \cap \{(J_1,\ldots,J_d) \in B\} \cap \{N_s \geq i\}\right) = \\
&= P(\{(J_1,\ldots,J_m) \in A\}) P(\{(J_1,\ldots,J_d) \in B\}) P(E \cap \{N_s \geq i\})
\end{aligned}$$

ends the proof. □

It is now obvious that the increments are independent. Thanks to this fact we have $\forall$ $0 \leq s < t$

$$\phi_{Y_t}(u) = \phi_{Y_t-Y_s}(u)\phi_{Y_s}(u)$$

hence

$$\phi_{Y_t-Y_s}(u) = \frac{\phi_{Y_t}(u)}{\phi_{Y_s}(u)} = \exp[\lambda(t-s)(\phi_J(u)-1)] = \phi_{Y_{t-s}}(u),$$

therefore Y has also stationary increments. About stochastic continuity we have that $\phi_{Y_t-Y_s}(u) \to 1$ as $s \to t$, hence, since as $s \to t$ μ_{t-s} converges weakly to δ_0 $Y_t - Y_s$ converges in probability to 0.

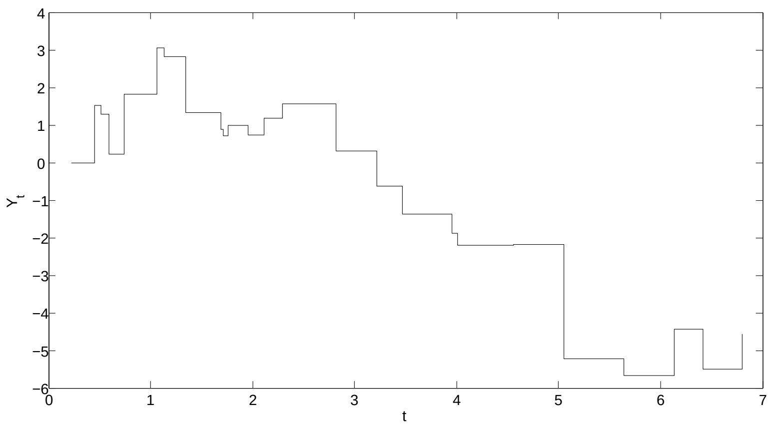

Figure 1.5: Compound Poisson process sample path with $\lambda = 3$ and jump-size standard Gaussian distribution

Remark 8. If X and Y are independent Lévy processes, then also $X + Y$ is a Lévy process. We will prove this fact later via Lévy-Itô theorem.

Example 1.7. (Lévy-jump diffusion) Assume that the process L in $\mathbb{R}$ is a sum between a linear Brownian motion and a compensated compound Poisson process, i.e.

$$L_t = bt + \sqrt{c}B_t + \sum_{n=1}^{N_t} J_n - t\lambda\mathbb{E}(J_1)$$

where $b \in \mathbb{R}$, $c \geq 0$ and $\lambda > 0$. Assume that all source of randomness are mutually independent. Then the characteristic function of L_t is

$$\begin{aligned}
\mathbb{E}(\mathrm{e}^{iuL_t}) &= \mathbb{E}\left[\exp\left(iu\left(bt + \sqrt{c}B_t + \sum_{n=1}^{N_t} J_n - t\lambda\mathbb{E}(J_1)\right)\right)\right] \\
&= \mathrm{e}^{iubt}\mathbb{E}[\exp(iu\sqrt{c}B_t)]\mathbb{E}\left[\exp\left(iu\sum_{n=1}^{N_t} J_n - iut\lambda\mathbb{E}(J_1)\right)\right] \\
&= \mathrm{e}^{iubt}\mathrm{e}^{-\frac{1}{2}u^2ct}\exp\left[\lambda t(\mathbb{E}[\mathrm{e}^{iuJ_1} - 1 - iuJ_1])\right] \\
&= \exp\left[t\left(iub - \frac{1}{2}u^2c + \int_{\mathbb{R}}(\mathrm{e}^{iux} - 1 - iux)\lambda\mu_J(\mathrm{d}x)\right)\right].
\end{aligned}$$

Example 1.8. Let $B = (B_t)_{t\in T}$ a Brownian motion on $\mathbb{R}$; we introduce the stopping time[12] $\tau_t = \inf\{u : B_u = t\}$. By the iterated logarithm law it is a finite stopping time

[12]Reflection principle helps us to describe the law μ_{τ_t}. $P(\tau_t \leq x) = 2P(B_x > t) = 2(1 - \Phi(t/\sqrt{x}))$, hence $\mu_{\tau_t}(\mathrm{d}x) = f_{\tau_t}(x)\mathrm{d}x$ and the density is absolutely continuous

$$f_{\tau_t}(x) = \frac{d}{\mathrm{d}t}F_{\tau_t}(x) = \frac{t}{x^{3/2}\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2x}}.$$

$\forall t \in T$ τ_t does not have finite mean, in effect $\mathbb{E}(\tau_t) = \int_0^\infty \frac{t}{\sqrt{2\pi x}}\mathrm{e}^{-\frac{t^2}{2x}}\mathrm{d}x$ and the integrand function behaves as $1/\sqrt{x}$ as $x \to \infty$. This process is also stable since $\mu_{\tau_{nt}}$ is equal to the image of μ_{τ_t} by the homothety $x \mapsto n^2x$. We saw that $\tau_{2t} = \tau_t + \tilde{\tau}_t$, so by recurrence the sum of n i.i.d. distributions process. Our aim is to show that it is a Lévy process. Clearly $\tau_0 = \inf\{u : B_u = 0\} = 0$, and as τ_t is finite $\tilde{B}_u = B_{u+\tau_t} - B_{\tau_t}$ is a Brownian motion independent of $\mathcal{F}_{\tau_t}$. Given $s \in T$ $\tilde{\tau}_s = \inf\{u : \tilde{B}_u = s\}$ and τ_s has the same distribution, and moreover $\tilde{\tau}_s$ is independent of τ_t. We can write

$$B_{\tau_t+\tilde{\tau}_s} = B_{\tau_t+\tilde{\tau}_s} - B_{\tau_t} + B_{\tau_t} = s + t.$$

Therefore one can deduce $\tau_t + \tilde{\tau}_s = \tau_{t+s}$, that is $\tau_{t+s} - \tau_t = \tilde{\tau}_s \overset{\mathrm{d}}{=} \tau_s$. It follows that the stopping time process has independent and homogeneous increments. Furthermore as $t \to 0$ by the reflection principle

$$P(\tau_t > \epsilon) = 1 - P(\tau_t \leq \epsilon) = 1 - 2P(B_\epsilon \geq t) \to 0$$

that guarantees that τ_t is a Lévy process.

1.2.2 Lévy-Itô decomposition

Theorem 1.2.5. *(Lévy-Itô decomposition) Given a vector $b \in \mathbb{R}^d$, a positive definite symmetric $d \times d$ matrix C and a measure ν on $\mathbb{R}^d$ satisfying $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}^d}(1 \wedge |x|^2)\nu(\mathrm{d}x) < \infty$, then there exists a probability space $(\Omega, \mathcal{F}, P)$ on which three independent Lévy processes exist, where L^1 is a linear Brownian motion, L^2 a compound Poisson process and L^3 a square integrable martingale with an a.s. countable number of jumps of magnitude less than 1 on each finite time interval. Taking $L = L^1 + L^2 + L^3$ one can define a Lévy process on a probability space with the following characteristic exponent*

$$\eta(u) = i\langle b, u\rangle - \frac{1}{2}\langle u, Cu\rangle + \int_{\mathbb{R}^d} [\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle \mathbf{1}_B(x)]\nu(\mathrm{d}x)$$

$\forall u \in \mathbb{R}^d$.

Remark 9. Any characteristic exponent η belonging to an infinitely divisible distribution can be written

$$\eta(u) = \left\{ i\langle b, u\rangle - \frac{1}{2}\langle u, Cu\rangle \right\} + \left\{ \nu(\mathbb{R}^d \backslash B) \int_{B^c} (\mathrm{e}^{i\langle u,x\rangle} - 1)\frac{\nu(\mathrm{d}x)}{\nu(\mathbb{R}^d\backslash B)} \right\} + \\ + \left\{ \int_B (\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle)\nu(\mathrm{d}x) \right\}$$

for all $u \in \mathbb{R}^d$. Call the three bracktes η^1, η^2 and η^3; η^1 and η^2 correspond, respectively, to a linear Brownian motion with drift b and diffusion C and a compound Poisson process with rate $\nu(\mathbb{R}^d\backslash B)$ and jump law $\mathbf{1}_{B^c}(x)\frac{\nu(\mathrm{d}x)}{\nu(\mathbb{R}^d\backslash B)}$. The proof of existence of the theorem boils down to showing the existence of the process L^3, whose characteristic exponent is given by η^3.

$\tau_1, \ldots, \tau_n$ with law of μ_{τ_t} have the same law of τ_{nt}. It follows that about the density of $X := \frac{\tau_1+\cdots+\tau_n}{n^2}$

$$f_X(x) = n^2 f_{\tau_{nt}}(n^2 x) = n^2 \frac{nt}{\sqrt{2\pi}(n^2x)^{3/2}}\mathrm{e}^{-\frac{t^2n^2}{2xn^2}} = \frac{t}{x^{3/2}\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2x}} = f_{\tau_t}(x)$$

that proves that τ_t has a strictly stable distribution with $\alpha = 1/2$.

However the proof of the theorem needs some preparatory results about square integrable martingales.
Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in T}, P)$ be a standard filtered probability space, let introduce a new space: the space of real valued, right continuous, zero mean and square integrable martingales with respect to the given filtration over the time period $[0, \tilde{t}]$. We call this space $\mathcal{M}^2_{\tilde{t}}$ and let see immediately that it is a complete space.

Proposition 1.2.6. *$(\mathcal{M}^2_{\tilde{t}}, \langle\cdot,\cdot\rangle)$ is a Hilbert space, where $\langle M, N\rangle = \mathbb{E}(M_{\tilde{t}}N_{\tilde{t}})$, $M, N \in \mathcal{M}^2_{\tilde{t}}$.*

Proof. We verify first of all that $\langle M, N\rangle = \mathbb{E}(M_{\tilde{t}}N_{\tilde{t}})$, $M, N \in \mathcal{M}^2_{\tilde{t}}$ is an inner product. It is symmetric ($\langle M, N\rangle = \langle N, M\rangle$), linear ($\langle \alpha M + \beta N, O\rangle = \alpha\langle M, O\rangle + \beta\langle N, O\rangle$) and such that $\langle M, M\rangle \geq 0$ by definition of expectation. It is also obvious that $M = 0$ implies $\langle M, M\rangle = 0$, but it is more difficult to prove that $\langle M, M\rangle = 0$ implies $M = 0$. However by Doob's maximal inequality, for $M \in \mathcal{M}^2_{\tilde{t}}$

$$\mathbb{E}(\sup_{s\leq\tilde{t}} M_s^2) \leq 4\mathbb{E}(M_{\tilde{t}}^2)$$

then $\sup_{t\leq\tilde{t}} M_t = 0$ a.s., and since M is right continuous it follows that $M_t = 0\ \forall t \in [0, \tilde{t}]$ a.s.
We now can show the completeness property; let $M^{(n)}$ a Cauchy sequence in $\mathcal{M}^2_{\tilde{t}}$, then $\forall t\ M_t^{(n)}$ is a Cauchy sequence in the Hilbert space of zero mean and square integrable random variables defined on $(\Omega, \mathcal{F}_t, P)$, $L^2(\Omega, \mathcal{F}_t, P)$, endowed with the inner product $\langle M, N\rangle$. Hence there exists a limiting variable M in L^2. If we show that M is a martingale the proof is concluded; we will prove that $M_s = \mathbb{E}(M_t|\mathcal{F}_s)$ for $s < t$. We have to show that $\mathbb{E}(M_s\mathbf{1}_A) = \mathbb{E}(M_t\mathbf{1}_A)\ \forall A \in \mathcal{F}_s$, but we have that[13] $\mathbb{E}(M_s\mathbf{1}_A) = \lim_{n\to\infty}\mathbb{E}(M_s^{(n)}\mathbf{1}_A)$ and $\mathbb{E}(M_t\mathbf{1}_A) = \lim_{n\to\infty}\mathbb{E}(M_t^{(n)}\mathbf{1}_A)$ and $\mathbb{E}(M_s^{(n)}\mathbf{1}_A) = \mathbb{E}(M_t^{(n)}\mathbf{1}_A)$ by martingale property for each $A \in \mathcal{F}_s$. Clearly the limit is an $\mathcal{F}_t$-adapted process and by Jensen's inequality

$$\mathbb{E}(M_s^2) = \mathbb{E}(\mathbb{E}(M_t|\mathcal{F}_s)^2) \leq \mathbb{E}(\mathbb{E}(M_t^2|\mathcal{F}_s)) = \mathbb{E}(M_t^2) < \infty.$$

If we take the right continuous version[14] of this martingale we can assert that $\mathcal{M}^2_{\tilde{t}}$ is a Hilbert space. □

Our aim is constructing a sequence of right continuous, zero expectation and square integrable martingales converging to a process having as Lévy symbol η^3. Suppose we have for $n \in \mathbb{N}$ a sequence of independent Poisson processes $N^{(n)} = (N_t^{(n)})_{t\in T}$ with rate $\lambda_n \geq 0$ and the i.i.d. sequences of $(J_k^{(n)})_{k\in\mathbb{N}}$, which are themselves mutually independent, with common law μ_{J_n} which does not assign mass to the origin and has finite second moment. Associated with each pair (λ_n, μ_{J_n}) is the square integrable

[13] We used the proposition according to which the convergence in L^p, $p > 1$, implies the convergence of the means.

[14] We use here the hypothesis of standard filtration, that guarantees the existence of a version càdlàg martingale.

martingale[15] described by the compensated compound Poisson process, we denote it by $M^{(n)}$ and its natural filtration by $(\mathcal{F}_t^{(n)})_{t\in T}$. We put all processes on the same probability space and take them as martingales with respect to the common filtration

$$\mathcal{F}_t := \sigma\left(\bigcup_{n\in\mathbb{N}} \mathcal{F}_t^{(n)}\right).$$

Theorem 1.2.7. *If $\sum_{n\geq 1}\lambda_n \int_{\mathbb{R}^d} |x|^2 \mu_{J_n}(\mathrm{d}x) < \infty$ then there exists a Lévy process X defined on the same space as the processes $\left\{M^{(n)} : n \in \mathbb{N}\right\}$ which is a square integrable martingale with Lévy exponent given by*

$$\eta(u) = \int_{\mathbb{R}^d} (\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle) \sum_{n\geq 1} \lambda_n \mu_{J_n}(\mathrm{d}x)$$

$\forall u \in \mathbb{R}^d$. Moreover for each $\tilde{t} > 0$

$$\lim_{k\to\infty} \mathbb{E}\left[\sup_{t\leq \tilde{t}} \left(X_t - \sum_{n=1}^{k} M_t^{(n)}\right)^2\right] = 0.$$

Proof. First of all $\sum_{n=1}^{k} M_t^{(n)}$ is a martingale, since it is sum of martingales and it is square integrable as

$$\mathbb{E}\left[\left(\sum_{n=1}^{k} M_t^{(n)}\right)^2\right] = \sum_{n=1}^{k} \mathbb{E}((M_t^{(n)})^2) = t\sum_{n=1}^{k} \lambda_n \int_{\mathbb{R}^d} |x|^2 \mu_{J_n}(\mathrm{d}x) < \infty$$

by independence and $\mathbb{E}(M_t^{(i)} M_t^{j}) = \mathbb{E}(M_t^{(i)})\mathbb{E}(M_t^{(j)}) = 0$ for $i \neq j$. Fixed $\tilde{t} > 0$, it easy to prove that $X_t^{(k)} = \sum_{n=1}^{k} M_t^{(n)}$, $0 \leq t \leq \tilde{t}$, is a Cauchy sequence with respect to the seminorm L^2, in effect

$$\begin{aligned}||X^{(k)} - X^{(l)}||^2 &= t\sum_{n=l}^{k} \lambda_n \int_{\mathbb{R}^d} |x|^2 \mu_{J_n}(\mathrm{d}x) = \\ &= t\sum_{n=1}^{k} \lambda_n \int_{\mathbb{R}^d} |x|^2 \mu_{J_n}(\mathrm{d}x) - t\sum_{n=1}^{l-1} \lambda_n \int_{\mathbb{R}^d} |x|^2 \mu_{J_n}(\mathrm{d}x) \to 0\end{aligned}$$

[15] If J_n has finite second moment for $n \in \mathbb{N}$ the process $M_t = \sum_{n=1}^{N_t} J_n - \lambda t\mathbb{E}(J_1)$ is a square integrable martingale with respect to its natural filtration. In effect $\mathbb{E}(M_t|\mathcal{F}_s) = M_s + \mathbb{E}(M_t - M_s|\mathcal{F}_s) = M_s + \mathbb{E}(M_{t-s}) = M_s$ $\forall\ 0 \leq s < t$ and

$$\begin{aligned}\mathbb{E}(M_t^2) &= \mathbb{E}\left[\left(\sum_{n=1}^{N_t} J_n\right)^2\right] - \lambda^2 t^2 \mathbb{E}^2(J_1) \\ &= \mathbb{E}\left(\sum_{n=1}^{N_t} J_n^2\right) + \mathbb{E}\left(\sum_{n=1}^{N_t}\sum_{m=1}^{N_t} J_n J_m \mathbf{1}_{\{n\neq m\}}\right) - \lambda^2 t^2 \mathbb{E}^2(J_1) \\ &= \lambda t\mathbb{E}(J_1^2) + \mathbb{E}(N_t^2 - N_t)\mathbb{E}^2(J_1) - \lambda^2 t^2 \mathbb{E}^2(J_1) \\ &= \lambda t\mathbb{E}(J_1^2) < \infty.\end{aligned}$$

as $k, l \to \infty$. We saw that the space $\mathcal{M}^2$ is a complete space, hence there exists a martingale $X = (X_t)_{0 \le t \le \tilde{t}}$ such that $||X^{(k)} - X||^2 \to 0$ as $k \to \infty$. Thanks to this convergence we can claim that the law of $X_t^{(k)}$ converges to the X law, and consequently since the processes $X^{(k)}$ are Lévy processes

$$\mathbb{E}(\mathrm{e}^{i\langle u, X_t - X_s\rangle}) = \lim_{k\to\infty} \mathbb{E}(\mathrm{e}^{i\langle u, X_t^{(k)} - X_s^{(k)}\rangle}) = \lim_{k\to\infty} \mathbb{E}(\mathrm{e}^{i\langle u, X_{t-s}^{(k)}\rangle}) = \mathbb{E}(\mathrm{e}^{i\langle u, X_{t-s}\rangle})$$

and we conclude that X has stationary increments. X has also independent increments

$$\begin{aligned} P(X_t - X_s \in A, X_r \in B) &= \lim_{k\to\infty} P(X_t^{(k)} - X_s^{(k)} \in A, X_r^{(k)} \in B) = \\ &= P(X_t - X_s \in B) P(X_r \in B) \end{aligned}$$

for all $r \le s < t$ by convergence of finite dimensional distributions. Furthermore thanks to Doob's inequality

$$\lim_{k\to\infty} \mathbb{E}(\sup_{0\le t\le \tilde{t}} (X_t - X_t^{(k)})^2) \le \lim_{k\to\infty} 4\mathbb{E}[(X_{\tilde{t}} - X_{\tilde{t}}^{(k)})^2] = 0.$$

To prove stochastic continuity[16]

$$\begin{aligned} \lim_{t\to 0} P(|X_t| > \epsilon) &\le \lim_{t\to 0}\lim_{k\to\infty} P(|X_t^{(k)}| > \epsilon/2) + \lim_{t\to 0}\lim_{k\to\infty} P(|X_t - X_t^{(k)}| > \epsilon/2) \\ &= \lim_{t\to 0}\lim_{k\to\infty} P(|X_t^{(k)}| > \epsilon/2) \\ &\le \lim_{t\to 0}\lim_{k\to\infty} 2\mathbb{E}(|X_t^{(k)}|)/\epsilon \\ &\le \lim_{t\to 0}\lim_{k\to\infty} 2\mathbb{E}\left(\sum_{n=1}^{k} |M_t^{(n)}|\right)/\epsilon \\ &\le \lim_{t\to 0}\lim_{k\to\infty} 2\mathbb{E}^{1/2}\left(\sum_{n=1}^{k} |M_t^{(n)}|\right)^2/\epsilon \\ &= \frac{2}{\epsilon}\lim_{t\to 0}\sqrt{t\sum_{n\ge 1}\lambda_n \int_{\mathbb{R}^d} |x|^2 \mu_{J_n}(\mathrm{d}x)} = 0. \end{aligned}$$

Then the limiting process is a Lévy process and

$$\lim_{k\to\infty} \mathbb{E}(\mathrm{e}^{i\langle u, X_t^{(k)}\rangle}) = \lim_{k\to\infty} \exp\left[\int_{\mathbb{R}^d} (\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle)\sum_{n=1}^{k}\lambda_n \mu_{J_n}(\mathrm{d}x)\right] = \mathrm{e}^{\eta(u)}.$$

The limiting process X depends on $\tilde{t}$ (say $X^{\tilde{t}}$), and we want to deal this issue. It is well known that if a_n and b_n are two sequences of real numbers, $\sup_n a_n^2 = (\sup_n |a_n|)^2$ and $\sup_n |a_n + b_n| \le \sup_n |a_n| + \sup_n |b_n|$, hence, by using also Minkowski's inequality,

[16] We use Markov and Cauchy-Schwarz inequalities and hypothesis on the convergence of the series $\sum_{n\ge 1}\lambda_n \int_{\mathbb{R}^d} |x|^2 \mu_{J_n}(\mathrm{d}x)$.

for $\tilde{t}_1 \leq \tilde{t}_2$

$$\begin{aligned}
\mathbb{E}^{1/2}\left[\sup_{t\leq\tilde{t}_1}(X_t^{\tilde{t}_1} - X_t^{\tilde{t}_2})^2\right] &= \mathbb{E}^{1/2}\left[\sup_{t\leq\tilde{t}_1}(X_t^{\tilde{t}_1} - X_t^{(k)} + X_t^{(k)} - X_t^{\tilde{t}_2})^2\right] \\
&= \mathbb{E}^{1/2}\left[\left(\sup_{t\leq\tilde{t}_1}|X_t^{\tilde{t}_1} - X_t^{(k)} + X_t^{(k)} - X_t^{\tilde{t}_2}|\right)^2\right] \\
&\leq \mathbb{E}^{1/2}\left[\left(\sup_{t\leq\tilde{t}_1}|X_t^{\tilde{t}_1} - X_t^{(k)}| + \sup_{t\leq\tilde{t}_1}|X_t^{(k)} - X_t^{\tilde{t}_2}|\right)^2\right] \\
&\leq \mathbb{E}^{1/2}\left[\left(\sup_{t\leq\tilde{t}_1}|X_t^{\tilde{t}_1} - X_t^{(k)}|\right)^2\right] + \mathbb{E}^{1/2}\left[\left(\sup_{t\leq\tilde{t}_1}|X_t^{\tilde{t}_2} - X_t^{(k)}|\right)^2\right] \\
&= \mathbb{E}^{1/2}\left[\sup_{t\leq\tilde{t}_1}(X_t^{\tilde{t}_1} - X_t^{(k)})^2\right] + \mathbb{E}^{1/2}\left[\sup_{t\leq\tilde{t}_1}(X_t^{\tilde{t}_2} - X_t^{(k)})^2\right] \to 0
\end{aligned}$$

as $k \to \infty$. Consequently the two processes agree a.s. on the time horizon $[0, \tilde{t}_1]$ and we may say that the limit does not depend on time $\tilde{t}$. □

We are then ready to prove the decomposition theorem.

Proof. (Theorem 1.2.5) The component η^3 of the Lévy symbol can be written as follows

$$\eta^3 = \sum_{n\geq 0}\left[\lambda_n \int_{\mathbb{R}^d}(\mathrm{e}^{i\langle u,x\rangle} - 1)\tilde{\nu}_n(\mathrm{d}x) + \lambda_n i \int_{\mathbb{R}^d}\langle u, x\rangle\tilde{\nu}_n(\mathrm{d}x)\right]$$

where $\lambda_n = \nu(\{x : 2^{-(n+1)} \leq |x| < 2^{-n}\})$ and $\tilde{\nu}_n(\mathrm{d}x) = \mathbf{1}_{\{2^{-(n+1)}\leq|x|<2^{-n}\}}\frac{\nu(\mathrm{d}x)}{\lambda_n}$ with the understanding that the n-th integral is absent if $\lambda_n = 0$. Since $\sum_{n\geq 0}\lambda_n \int_{\mathbb{R}^d}|x|^2\tilde{\nu}_n(\mathrm{d}x) = \int_B |x|^2\nu(\mathrm{d}x) < \infty$ then η^3 is the characteristic exponent of a Lévy process by Theorem 1.2.7, it is enough to take $\lambda_n = \nu(\{x : 2^{-(n+1)} \leq |x| < 2^{-n}\})$ and $\tilde{\nu}_n(\mathrm{d}x) = \mathbf{1}_{\{2^{-(n+1)}\leq|x|<2^{-n}\}}\frac{\nu(\mathrm{d}x)}{\lambda_n}$.
About the fact that the three processes are on the same space we can construct a product space which supports all these independent processes. □

Remark 10. If the Lévy measure is finite from the symbol we get that L is sum of a linear Brownian motion and a compound Poisson process,

$$\eta(u) = i\langle u, \int_{\mathbb{R}^d} x\mathbf{1}_B(x)\nu(\mathrm{d}x)\rangle - \frac{1}{2}\langle u, Cu\rangle + \nu(\mathbb{R}^d)\int_{B^c}(\mathrm{e}^{i\langle u,x\rangle} - 1)\frac{\nu(\mathrm{d}x)}{\nu(\mathbb{R}^d)}.$$

Then almost all paths of L have a finite number of jumps on every compact interval; we say that L has *finite activity.* If ν is not finite the compound Poisson martingale associated with $(\lambda_n, \tilde{\nu}_n)$ is such that the rate of arrival of the jumps increases and the size of the jumps decreases as n increases. Then the process has a countable infinity of small jumps and L has *infinite activity.*

1.2.3 Right continuous version, filtrations and strong Markov property

We know that the Brownian motion has an a.s. continuous version, we can formulate a similar theorem about Lévy processes and the proof can be read in Applebaum [2].

Theorem 1.2.8. *Every Lévy process has a càdlàg modification that is itself a Lévy process.*

We are going to introduce important results concerning filtrations in the Lévy processes theory.

Proposition 1.2.9. *Let $L = (L_t)_{t\in T}$ an $\mathcal{F}_t$-Lévy Process, then L is independent of $\mathcal{F}_0$.*

Proof. We need to prove independence of $\sigma(L)$ and $\mathcal{F}_0$, but we can just prove the independence of J and $\mathcal{F}_0$, where

$$J := \left\{ \{L_{t_1} \in A_1, \dots, L_{t_k} \in A_k\}, k \in \mathbb{N}, t_i \in T, A_i \in \mathcal{B}(\mathbb{R}^d) \right\}$$

is a base for $\sigma(L)$. Let introduce the increments $Y_i := L_{t_i} - L_{t_i - 1}$, for $1 \le i \le k$ and $t_0 = 0$ and prove the independence among $\mathcal{F}_0, \sigma(Y_1), \dots, \sigma(Y_k)$. For each $D \in \mathcal{F}_0$ and $A_1, \dots, A_k \in \mathcal{B}(\mathbb{R}^d)$ we need to show that

$$P\left(D \bigcap_{i=1}^{k} \{Y_i \in A_i\}\right) = P(D) \prod_{i=1}^{k} P(Y_i \in A_i) \tag{1.3}$$

and we are proceeding by induction on k. For $k = 1$ it is clear as $Y_1 = L_{t_1} - L_0$ is independent of $\mathcal{F}_0$. Similarly $Y_k = L_{t_k} - L_{t_{k-1}}$ is independent of $\mathcal{F}_{k-1}$ and $D \bigcap_{i=1}^{k-1} \{Y_i \in A_i\} \in \mathcal{F}_{t_{k-1}}$ by Doob measurability criterion. If (1.3) is true for $k-1$, then

$$\begin{aligned} P\left(D \bigcap_{i=1}^{k} \{Y_i \in A_i\}\right) &= P\left(D \bigcap_{i=1}^{k-1} \{Y_i \in A_i\}\right) P(Y_k \in A_k) \\ &= P(D) \prod_{i=1}^{k} P(Y_i \in A_i). \end{aligned}$$

Then the vector $(Y_1 \dots, Y_k)$ is independent of $\mathcal{F}$, and since $(L_{t_1}, \dots, L_{t_k})$ is measurable function of $(Y_1, \dots, Y_k)$ by Doob criterion is itself independent of $\mathcal{F}_0$. □

Proposition 1.2.10. *If L is an $\mathcal{F}_t$-Lévy process with càdlàg paths, then it is at the same time an $\mathcal{F}_{t+}$-Lévy process.*

Proof. We need to prove only the independent increments condition. We have to show that $L_t - L_s$ is independent of $\mathcal{F}_{s+} = \bigcap_{\epsilon>0} \mathcal{F}_{s+\epsilon}$. By independence of the increments for each $\epsilon > 0$ $L_{t+\epsilon} - L_{s+\epsilon}$ is independent of $\mathcal{F}_{s+\epsilon}$, and it is also independent of $\mathcal{F}_{s+}$ as $\mathcal{F}_{s+} \subseteq \mathcal{F}_{s+\epsilon}$. In order to conclude the proof it is enough to show that $\mathbb{E}(h(L_t - L_s)|A) =$

$\mathbb{E}(h(L_t - L_s))$ for each $A \in \mathcal{F}_{s+}$ and for any $h : \mathbb{R}^d \to \mathbb{R}$ bounded and continuous[17]. We have proved so far

$$\mathbb{E}(h(L_{t+1/n} - L_{s+1/n})|A) = \mathbb{E}(h(L_{t+1/n} - L_{s+1/n})).$$

L is right continuous and by continuity of the test functions h

$$\lim_{n\to\infty} h(L_{t+1/n} - L_{s+1/n}) = h(L_t - L_s) \quad a.s.;$$

then, by dominated convergence,

$$\mathbb{E}(h(L_t - L_s)|A) = \mathbb{E}(h(L_t - L_s))$$

for $A \in \mathcal{F}_{s+}$ and h. $\square$

Proposition 1.2.11. *Every right continuous $\mathcal{F}_t$- Lévy process L defined on a complete probability space is an $\bar{\mathcal{F}}_{t+}$-Lévy process, where $\bar{\mathcal{F}}_{t+} := \sigma(\mathcal{F}_{t+}, \mathcal{N})$, with $\mathcal{N} := \{N \in \mathcal{F} : P(N) = 0\}$.*

Proof. As in the previous theorem we just prove independence of the increments, i.e. independence between $L_t - L_s$ and $\bar{\mathcal{F}}_{s+} = \sigma(\mathcal{F}_{s+}, \mathcal{N})$, where $\mathcal{N} = \{N \in \mathcal{F} : P(N) = 0\}$. We need to remember that if a r.v. X is independent of $\mathcal{G} \subset \mathcal{F}$, then it is independent of $\bar{\mathcal{G}} := \sigma(\mathcal{G}, \mathcal{N})$ too. Thanks to this fact we can end the proof. $\square$

Theorem 1.2.12. *(0-1 Blumenthal law) Let $L = (L_t)_{t\in T}$ be a right continuous Lévy process and let $\mathcal{F}_t$ be its natural filtration, then $\mathcal{F}_{0+}$ is trivial: for each $A \in \mathcal{F}_{0+}$ $P(A) = 0$ or $P(A) = 1$.*

Proof. By Proposition 1.2.10 L is also an $\mathcal{F}_{t+}$-Lévy process and by Proposition 1.2.9 it is independent of $\mathcal{F}_{0+}$, hence $\sigma(L) := \sigma((L_t)_{t\in T})$ is independent of $\mathcal{F}_{0+}$. As $\mathcal{F}_{0+} \subseteq \sigma(L)$, then $\mathcal{F}_{0+}$ is independent of itself. Consequently $P(A) = P(A \cap A) = P(A)P(A)$ for $A \in \mathcal{F}_{0+}$ and $P(A) = 0$ or $P(A) = 1$. $\square$

Theorem 1.2.13. *(Strong Markov property) Let $L = (L_t)_{t\in T}$ and τ be a right continuous $\mathcal{F}_t$-Lévy process and an a.s. finite stopping time. Then $Z_t := L_{\tau+t} - L_\tau$ is a $\mathcal{G}_t$-Lévy process, where $\mathcal{G}_t := \mathcal{F}_{\tau+t}$, and it has the same law of L.*

Proof. $\tau + t$ is a stopping time for each $t \in T$, hence $\mathcal{G}_t = \mathcal{F}_{\tau+t}$ is well defined, and as $\tau + s \leq \tau + t$ for $s < t$ $\mathcal{G}_t$ is a filtration. Z_t is $\mathcal{F}_{\tau+t}$-measurable[18], then Z is an adapted process.
$Z_0 = L_\tau - L_\tau = 0$ gets right continuity by L. In order to prove independence and stationarity of the increments we show first of all that

$$\mathbb{E}(h(L_{\tau+t} - L_{\tau+s})|G) = \mathbb{E}(h(L_{\tau+t} - L_{\tau+s})) \tag{1.4}$$

[17] An $\mathbb{R}^d$-valued random variable X is independent of $\mathcal{G}$ if and only if $\mathbb{E}(h(X)|G) = \mathbb{E}(h(X))$ for any $G \in \mathcal{G}$ with $P(G) > 0$ and for each $h : \mathbb{R}^d \to \mathbb{R}$ bounded and continuous. We call this function h also test function.

[18] If X is a progressively measurable process and τ is an a.s. finite stopping time, then X_τ is $\mathcal{F}_\tau$-measurable. Our process L is progressively measurable: it is right continuous and adapted.

for $s < t$ and $h : \mathbb{R}^d \to \mathbb{R}$ bounded and continuous and $G \in \mathcal{G}_s$ such that $P(G) > 0$. Start with discrete stopping times taking the values $(t_m)_{m\in I}$, then for any $m \in I$

$$\mathbb{E}(h(L_{\tau+t} - L_{\tau+s})\mathbf{1}_G\mathbf{1}_{\{\tau=t_m\}}) = \mathbb{E}(h(L_{\tau+t} - L_{\tau+s})\mathbf{1}_{G\cap\{\tau=t_m\}}).$$

$G \in \mathcal{G}_s = \mathcal{F}_{\tau+s}$, consequently $G \cap \{\tau = t_m\} = G \cap \{\tau + s = t_m + s\} \in \mathcal{F}_{t_m+s}$ and by hyphotesis, since L is a Lévy process and $L_{t_m+t} - L_{t_m+s}$ is independent of $\mathcal{F}_{t_m+s}$, $h(L_{\tau+t} - L_{\tau+s})$ and $\mathbf{1}_{G\cap\{\tau=t_m\}}$ are independent. Hence

$$\mathbb{E}(h(L_{\tau+t} - L_{\tau+s})\mathbf{1}_G\mathbf{1}_{\{\tau=t_m\}}) = \mathbb{E}(h(L_{\tau+t} - L_{\tau+s}))P(G \cap \{\tau = t_m\}).$$

As the increments of L are stationary

$$\begin{aligned}\mathbb{E}(h(L_{\tau+t} - L_{\tau+s})\mathbf{1}_G) &= \sum_{m\in I} \mathbb{E}(h(L_{\tau+t} - L_{\tau+s})\mathbf{1}_G\mathbf{1}_{\tau=t_m})\\ &= \mathbb{E}(h(L_{t-s})) \sum_{m\in I} P(G \cap \{\tau = t_m\})\\ &= \mathbb{E}(h(L_{t-s}))P(G)\end{aligned}$$

and $\frac{\mathbb{E}(h(L_{\tau+t}-L_{\tau+s})\mathbf{1}_G)}{P(G)} = \mathbb{E}(h(L_{\tau+t} - L_{\tau+s})|G) = \mathbb{E}(h(L_{t-s}))$.
If we have an arbitrary stopping time τ we can consider a sequence τ_n of discrete stopping times converging[19] to τ. We have

$$\mathbb{E}(h(L_{\tau_n+t} - L_{\tau_n+s})|G) = \mathbb{E}(h(L_{t-s}))$$

and by continuity of the test function h and right continuity of the process L $h(L_{\tau_n+t} - L_{\tau_n+s}) \to h(L_{\tau+t} - L_{\tau+s})$ a.s. for $n \to \infty$. By dominated convergence theorem (h is bounded) we obtain (1.4).
$Z_t - Z_s = L_{\tau+t} - L_{\tau+s}$, then $\mathbb{E}(h(Z_t - Z_s)|G) = \mathbb{E}(h(L_{t-s}))$ for each $0 \leq s < t$, for every function $h : \mathbb{R}^d \to \mathbb{R}$ continuous and bounded and for every event $G \in \mathcal{G}_s$ such that $P(G) > 0$. If $G = \Omega$, then $\mathbb{E}(h(Z_t - Z_s)) = \mathbb{E}(h(L_{t-s}))$, and the law of $Z_t - Z_s$ depends only on $t - s$[20] and it follows that the increments are homogeneous. Moreover $\mathbb{E}(h(Z_t - Z_s)|G) = \mathbb{E}(h(Z_t - Z_s))$ and we can assert that $Z_t - Z_s$ is independent of $\mathcal{G}_s$ and that the process Z is a Lévy process.
About the law we just prove that Z_t and L_t have the same distribution for $t \in T$ and the fact follows by $\mathbb{E}(h(Z_t - Z_s)) = \mathbb{E}(h(L_{t-s}))$ with $s = 0$[21]. □

Remark 11. Theorem 1.2.13 is a generalization of the strong Markov property for the Brownian motion, according to which $Z_t := B_{\tau+t} - B_\tau$ is an $\mathcal{F}_{\tau+t}$-Brownian motion independent of $\mathcal{F}_\tau$ if τ is an a.s. finite stopping time and $B = (B_t)_{t\in T}$ is a $\mathcal{F}_t$-Brownian motion.

---

[19] If τ is a stopping time, then there exists a decreasing sequence τ_n of discrete stopping times such that for every $\omega \in \Omega$ $\tau_n(\omega) \downarrow \tau(\omega)$ for $n \to \infty$.

[20] A probability measure on $\mathbb{R}^d$ is defined via integrals of continuous and bounded functions. Let μ and ν be two probability measures on $\mathbb{R}^d$ such that $\int h d\mu = \int h d\nu$ for every function $h : \mathbb{R}^d \to \mathbb{R}$ continuous and bounded, then $\mu = \nu$.

[21] Two Lévy processes X and Y have the same law if and only if X_t and Y_t have the same law for every $t \in T$.

1.2.4 Jump processes and Poisson random measures

We introduce the jump process $\Delta L = (L_t - L_{t^-})_{t\in T}$, where $L_{t^-} = \lim_{s\uparrow t} L_s$, associated with a Lévy process L. It is important to check the jump process cannot have independent increments.

Lemma 1.2.14. *If L is a Lévy process, then $\Delta L_t = 0$ a.s. for fixed $t \in T$.*

Proof. Let $(t_n)_{n\in\mathbb{N}}$ be a monotone sequence in $\mathbb{R}^+$ such that $t_n \uparrow t$ as $n \to \infty$. L has càdlàg paths, then $\lim_{n\to\infty} L_{t_n} = L_{t^-}$. By stochastic continuity $\lim_{t_n\to t} P(|L_{t_n} - L_t| > \epsilon) = 0 \; \forall \; \epsilon > 0$, hence there exists a subsequence which converges almost surely to L_t. Consequently we have

$$\lim_{n\to\infty} L_{t_n} = L_{t^-} \quad \text{and} \quad \lim_{j\to\infty} L_{t_{n_j}} = L_t$$

and by uniqueness of limit $\Delta L_t = L_t - L_{t^-} = 0$ a.s. □

Remark 12. This lemma says that a Lévy process L has no fixed times of discontinuity. In general, the sum of the jumps of a Lévy process does not converge, it is possible to have

$$\sum_{0\le s\le t} |\Delta L_s| = \infty.$$

But we always have

$$\sum_{0\le s\le t} |\Delta L_s|^2 < \infty.$$

Definition 1.3. (Random measure) Given $(E, \mathcal{E})$ a measurable space a random measure on $(E, \mathcal{E})$ is a transition kernel from $(\Omega, \mathcal{F})$ to $(E, \mathcal{E})$; explicitly a mapping $M : \Omega \times \mathcal{E} \to \bar{\mathbb{R}}^+$ such that $\omega \mapsto M(\omega, A)$ is a random variable for every $A \in \mathcal{E}$ and $A \mapsto M(\omega, A)$ is a measure on $(E, \mathcal{E})$ for each ω in Ω.

A convenient tool for analyzing the jumps of a Lévy process is the random counting measure of the jumps. Consider a set $A \in \mathcal{B}(\mathbb{R}^d\backslash 0)$ such that $0 \notin \bar{A}$ (in this case we will say also that A is *bounded below*), for fixed $t \in T$ define the following random measure

$$N(t, A) = \#\{0 \le s \le t : \Delta L_s \in A\} = \sum_{0<s\le t} \mathbf{1}_A(\Delta L_s) = \sum_{0<s\le t} \delta_{\Delta L_s}(A)$$

that counts the jumps of the process L of size in A up to time t.

Proposition 1.2.15. *The set function $A \mapsto N(t, A)$ defines a $\sigma-$finite measure on $\mathbb{R}^d\backslash \{0\}$ for each (ω, t). Moreover the set function $\lambda(A) = \mathbb{E}(N(1, A))$ defines a $\sigma-$finite measure on $\mathbb{R}^d\backslash \{0\}$.*

Proof. The set function $A \mapsto N(t, A)$ is a counting measure for every (ω, t), hence it is $\sigma-$finite on $\mathbb{R}^d\backslash \{0\}$. $\mathbb{E}(N(t, A)) = \int_\Omega N(\omega, t, A)P(\mathrm{d}\omega)$, and since $\mathbb{E}(N(t, \emptyset)) = 0$ and for disjoint sets A_i, $i \in \mathbb{N}$

$$\mathbb{E}\left(N\left(t, \bigcup_i A_i\right)\right) = \int_\Omega N\left(\omega, t, \bigcup_i A_i\right) P(\mathrm{d}\omega) = \int_\Omega \sum_i N(\omega, t, A_i)P(\mathrm{d}\omega) =$$
$$= \sum_i \mathbb{E}(N(t, A_i))$$

it is a measure on $\mathbb{R}^d \setminus \{0\}$. It is also $\sigma-$finite by Radon-Nikodym theorem. □

The following theorem gives some properties about the random measure $N(t,\cdot)$, in particular in terms of kind of process. The proof is in Applebaum [2].

Theorem 1.2.16.

1. *If A is bounded below, then $(N(t,A))_{t\in T}$ is a Poisson process with intensity $\lambda(A) = \mathbb{E}(N(1,A))$.*

2. *If $A_1,\dots,A_m \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ are disjoint, then $N(t,A_1),\dots,N(t,A_m)$ are independent random variables.*

Definition 1.4. (Poisson random measure) Let $(E,\mathcal{E})$ be a measurable space and let λ be a σ-finite measure on it. A random measure N on $(E,\mathcal{E})$ is called Poisson random measure if

1. for every $A \in \mathcal{E}$ the random variable $N(A)$ is Poisson distributed with expectation $\lambda(A)$;

2. whenever $A_1,\dots,A_n$ are in $\mathcal{E}$ and disjoint the r.v. $N(A_1),\dots,N(A_n)$ are independent for $n \geq 2$.

Given a σ-finite measure λ and a measurable space $(E,\mathcal{E})$ we can prove the existence of a Poisson random measure on $(E,\mathcal{E})$ with intensity λ. It holds the following lemma.

Lemma 1.2.17. *Given a σ-finite measure λ on a measurable space $(E,\mathcal{E})$ there exists a Poisson random measure N on a probability space such that $\lambda(A) = \mathbb{E}(N(A))$ for all $A \in \mathcal{E}$.*

Proof. We assume first λ is finite and let $\pi = \lambda(E)$ and $\tilde{\lambda} = \lambda/\pi$. $\tilde{\lambda}$ is a probability measure and let $(\xi_n)_{n\mathbb{N}}$ be a sequence of i.i.d. r.v. with common law $\tilde{\lambda}$ and X a Poisson r.v. with parameter π independent of the sequence. Then the random measure

$$N = \sum_{j=1}^{X} \delta_{\xi_j}$$

is a Poisson random measure with intensity λ. If λ is σ-finite, then there exists a sequence $(B_n)_{n\in\mathbb{N}}$ of subsets of E such that $E = \bigcup_n B_n$ and $\lambda(B_n) < \infty$ for every n. The restriction λ_n of the measure λ on B_n is finite, hence we can construct independent Poisson random measures N_n with intensity λ_n. It follows that $N = \sum_n N_n$ is a Poisson random measure. □

Remark 13. Suppose that $E = \mathbb{R}^+ \times U$, where U is a space on which is defined a σ-field $\mathcal{C}$ and $\mathcal{E} = \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{C}$. Let $X = (X_t)_{t\in T}$ be an adapted process taking value in U such that N is a Poisson random measure on $(E,\mathcal{E})$, where $N([0,t]\times A) = \#\{0 \leq s \leq t : X_s \in A\}$ for each $t \in T, A \in \mathcal{C}$. We will call X Poisson point process and N its associated Poisson random measure[22]. If $U = \mathbb{R}^d \setminus \{0\}$, $\mathcal{C}$ is its Borel σ-field and L is a Lévy process then ΔL is a Poisson point process and N is its associated Poisson random measure.

[22]Some authors call Poisson point process the random measure N.

Let $N(t, A) = \#\{0 \leq s \leq t : \Delta L_s \in A\}$ with A bounded below and $f : \mathbb{R}^d \to \mathbb{R}^d$ a Borel measurable function. For each $t \in T$, $\omega \in \Omega$ and A bounded below we define the *Poisson integral* of f by

$$\int_A f(x) N(t, \mathrm{d}x) = \sum_{x \in A} f(x) N(t, \{x\}).$$

Since N depends on ω $\int_A f(x) N(t, \mathrm{d}x)$ is an $\mathbb{R}^d$-valued random variable and generates a càdlàg stochastic process. We have $N(t, \{x\}) \neq 0 \Leftrightarrow \Delta L_s = x$ for at least one $0 \leq s \leq t$, then

$$\int_A f(x) N(t, \mathrm{d}x) = \sum_{s \leq t} f(\Delta L_s) \mathbf{1}_A(\Delta L_s). \tag{1.5}$$

It follows finally a useful theorem concerning Poisson integration (see [2] for the proof).

Theorem 1.2.18. *Let A be bounded below, then*

1. $\left(\int_A f(x) N(t, \mathrm{d}x)\right)_{t \in T}$ *is a compound Poisson process such that for each $u \in \mathbb{R}^d$*

$$\mathbb{E}\left(\exp\left[i\langle u, \int_A f(x) N(t, \mathrm{d}x)\rangle\right]\right) = \exp\left[t \int_A (\mathrm{e}^{i\langle u, x\rangle} - 1)\lambda_f(\mathrm{d}x)\right] =$$
$$= \exp\left[t \int_A (\mathrm{e}^{i\langle u, f(x)\rangle} - 1)\nu(\mathrm{d}x)\right]$$

where $\lambda_f = \lambda \circ f^{-1}$;

2. *if $f \in L^1(A, \lambda_A)$*

$$\mathbb{E}\left(\int_A f(x) N(t, \mathrm{d}x)\right) = t \int_A f(x) \lambda(\mathrm{d}x);$$

3. *if $f \in L^2(A, \lambda_A)$*

$$\mathbb{V}\mathrm{ar}\left(\left|\int_A f(x) N(t, \mathrm{d}x)\right|\right) = t \int_A |f(x)|^2 \lambda(\mathrm{d}x).$$

1.2.5 Path variation and moments

Remark 14. By Theorem 1.2.18 the Lévy-Itô decomposition can be written as

$$L_t = bt + \sqrt{C} B_t + \int_{|x|>1} x N(t, \mathrm{d}x) + \int_{|x|\leq 1} x \tilde{N}(t, \mathrm{d}x)$$

where $\tilde{N}(t, A) = N(t, A) - t\lambda(A)$. By the same theorem we can assert that λ is the Lévy measure of the process L. The L^2-martingale $(\int_{|x|\leq 1} x\tilde{N}(t, \mathrm{d}x))_{t \in T}$ describes the small jumps, while the process $\int_{|x|>1} x N(t, \mathrm{d}x)$ describes the large jumps and it is a compound Poisson process as we have already seen.

The Lévy measure is responsible for other interesting properties, summarized in the following remark and propositions about paths variation and moments.

Remark 15. From the Lévy-Îto decomposition is clear that the presence of the linear Brownian motion would imply that paths of the Lévy process have unbounded variation. But if there is no Brownian component the process may or may not have unbounded variation. The term L^2, being a compound Poisson process, has only bounded variation. Hence, in the case there is no diffusion component, understanding whether the Lévy process has unbounded variation is an issue determined by the process L^3. $\int_B xN(t, \mathrm{d}x) < \infty$ if and only if $\int_B |x|\nu(\mathrm{d}x) < \infty$. In that case we can identify L^3 directly via

$$L_t^3 = \int_B xN(t, \mathrm{d}x) - t \int_B x\nu(\mathrm{d}x), \quad t \in T.$$

This also tells us that L^3 will be of bounded varition if and only if $\int_B |x|\nu(\mathrm{d}x) < \infty$. Consequently we have the following proposition.

Proposition 1.2.19. *Let L be a Lévy process with triplet (b, C, ν).*

1. *If $C = 0$ and $\int_{|x|\leq 1} |x|\nu(\mathrm{d}x) < \infty$, then L has finite variation.*
2. *If $C \neq 0$ or $\int_{|x|\leq 1} |x|\nu(\mathrm{d}x) = \infty$, then L has infinite variation.*

Proof. It follows immediately from Proposition 1.2.19. □

It also holds an important result abount moments of the process for fixed $t \in T$.

Proposition 1.2.20. *If L is a Lévy process with triplet (b, C, ν), then*

1. *L_t has r-th moment for $r \in \mathbb{R}^+$ if and only if $\int_{|x|>1} |x|^r \nu(\mathrm{d}x) < \infty$.*
2. *L_t has moment generation function $\mathbb{E}(\mathrm{e}^{\langle u, L_t\rangle}) < \infty$ for $u \in \mathbb{R}^d$ if and only if $\int_{|x|>1} \mathrm{e}^{\langle u,x\rangle}\nu(\mathrm{d}x) < \infty$.*

Proof. We prove the second statement without loss of generality for $d = 1$. First suppose that $\mathbb{E}(e^{uL_t}) < \infty$ for some $t > 0$. Recall L^1, L^2 and L^3 given in the Lévy-Îto decomposition. Note that L^2 is a compound Poisson process with arrival rate $\lambda := \nu(\mathbb{R} \setminus (-1, 1))$ and jump distribution $F(\mathrm{d}x) := \mathbf{1}_{\{|x|\geq 1\}}\nu(\mathrm{d}x)/\nu(\mathbb{R} \setminus (-1, 1))$ and $L^1 + L^3$ is a Lévy process with Lévy measure $\mathbf{1}_{\{|x|\leq 1\}}\nu(\mathrm{d}x)$. Since

$$\mathbb{E}(\mathrm{e}^{uL_t}) = \mathbb{E}(\mathrm{e}^{uL_t^2})\mathbb{E}(\mathrm{e}^{u(L_t^1+L_t^3)})$$

it follows that

$$\mathbb{E}(\mathrm{e}^{uL_t^2}) < \infty, \tag{1.6}$$

and, as L^2 is a compound Poisson process,

$$\begin{aligned} \mathbb{E}(\mathrm{e}^{uL_t^2}) &= \mathrm{e}^{-\lambda t} \sum_{k\geq 0} \frac{(\lambda t)^k}{k!} \int_{\mathbb{R}} \mathrm{e}^{ux} F^{*k}(\mathrm{d}x) \\ &= \mathrm{e}^{\nu(\mathbb{R}\setminus(-1,1))t} \sum_{k\geq 0} \frac{t^k}{k!} \int_{\mathbb{R}} \mathrm{e}^{ux} (\nu|_{\mathbb{R}\setminus(-1,1)})^{*k}(\mathrm{d}x) < \infty, \end{aligned} \tag{1.7}$$

where F^{*n} and $(\nu|_{\mathbb{R}\setminus(-1,1)})^{*n}$ are the n-fold convolution of F and $\nu|_{\mathbb{R}\setminus(-1,1)}$ and $\nu|_{\mathbb{R}\setminus(-1,1)}$ is the restriction of ν to $\mathbb{R}\setminus(-1,1)$. For $k=1$ we have

$$\int_{|x|\geq 1} \mathrm{e}^{ux}\nu(\mathrm{d}x) < \infty.$$

Now suppose that $\int_{\mathbb{R}} \mathrm{e}^{ux}\mathbf{1}_{\{|x|\geq 1\}}\nu(\mathrm{d}x) < \infty$ for some $u\in\mathbb{R}$. Since $(\nu|_{\mathbb{R}\setminus(-1,1)})^{*n}(\mathrm{d}x)$ is a finite measure, we have

$$\int_{\mathbb{R}} \mathrm{e}^{ux}(\nu|_{\mathbb{R}\setminus(-1,1)})^{*n}(\mathrm{d}x) = \left(\int_{|x|\geq 1} \mathrm{e}^{ux}\nu(\mathrm{d}x)\right)^n,$$

and hence (1.6) and (1.7) hold for all $t>0$. The proof is completed once we show that for $t>0$

$$\mathbb{E}(\mathrm{e}^{u(L_t^1+L_t^3)}) < \infty.$$

Since L^1+L^3 has a Lévy measure with finite support, it follows that its characteristic exponent can be extended to an entire function analytic on the whole $\mathbb{C}$. To see this, note that

$$\begin{aligned}\int_{(-1,1)} (1-\mathrm{e}^{iux}+iux)\nu(\mathrm{d}x) &= \int_{(-1,1)} \sum_{k\geq 0} \frac{(iux)^{k2}}{(k+2)!}\nu(\mathrm{d}x)\\ &\leq \sum_{k\geq 0} \frac{(|u|)^{k2}}{(k+2)!}\int_{(-1,1)} x^2\nu(\mathrm{d}x) < \infty.\end{aligned}$$

Hence the characteristic symbol can be written as a power series for all $u\in\mathbb{C}$ and is thus entire. The proof of the first part can be found in [48]. □

Remark 16. The variation of a Lévy process depends on the small jumps and the diffusion component, the moment properties depend on the big jumps, while the activity depends on all the jumps of the process.

Remark 17. As already said if X and Y are independent Lévy processes with triplets (b,C,ν) and (b',C',ν'), then the sum is also a Lévy process. In the light of the decomposition theorem as

$$\phi_{X_t+Y_t}(u) = \exp\left\{t\left[i\langle\hat{b},u\rangle - \frac{1}{2}\langle u,\hat{C}u\rangle + \int_{\mathbb{R}^d}[\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u,x\rangle\mathbf{1}_B(x)]\hat{\nu}(\mathrm{d}x)\right]\right\}$$

where $\hat{b}=b+b'$, $\hat{C}=C+C'$ and $\hat{\nu}=\nu+\nu'$.

It will be useful the following theorem too (see [48] for details).

Theorem 1.2.21. *Let L be a Lévy process on $\mathbb{R}^d$. Define $L_t^* = \sup_{s\in[0,t]}|L_s|$ and let $g(r)$ be a non-negative continuous submultiplicative*[23] *function on $[0,\infty)$, increasing to ∞ as $r\to\infty$. Then the following statements are equivalent.*

1. *$\mathbb{E}(g(L_t^*)) < \infty$ for some $t\in T$*

[23] A function $g:\mathbb{R}^d\to[0,\infty)$ is called submultiplicative if there exists a constant $K>0$ such that $g(x+y)\leq Kg(x)g(y)$ for all $x,y\in\mathbb{R}^d$

2. $\mathbb{E}(g(L_t^*)) < \infty$ *for every* $t \in T$
3. $\mathbb{E}(|g(L_t^*)|) < \infty$ *for some* $t \in T$
4. $\mathbb{E}(|g(L_t^*)|) < \infty$ *for every* $t \in T$.

We conclude this paragraph with an important result which we will use a lot in the next chapters (Kyprianou [30] shows the proof).

Theorem 1.2.22. *(Compensation formula) Suppose* $f : T \times \mathbb{R} \times \Omega \to [0, \infty)$ *is a random time-space function such that*

1. *as a trivariate function* $f = f(t,x)(\omega)$ *is measurable,*
2. *for each* $t \in T$ $f(t,x)(\omega)$ *is* $\mathcal{F}_t \times \mathcal{B}(\mathbb{R})$*-measurable,*
3. *for each* $x \in \mathbb{R}$*, with probability one,* $\{f(t,x)(\omega) : t \in T\}$ *is a left continuous process.*

Then for all $t \in T$

$$\mathbb{E}\left(\int_{[0,t]}\int_{\mathbb{R}} f(s,x)N(\mathrm{d}s,\mathrm{d}x)\right) = \mathbb{E}\left(\int_0^t\int_{\mathbb{R}} f(s,x)\mathrm{d}s\nu(\mathrm{d}x)\right). \tag{1.8}$$

1.2.6 Subordinators

A subordinator $S = (S_t)_{t\in T}$ is a one-dimensional Lévy process such that $t \mapsto S_t$ is a.s. nondecreasing. By Lévy process definition this is equivalent to ask that $S_t \geq 0$ a.s. for every $t \in T$.

Example 1.9. If S_t is a Brownian motion we have $P(S_t \leq 0) = 1/2$, then it is clear that such a process cannot be a subordinator.

Theorem 1.2.23. *A Lévy process* S *is a subordinator if and only if its characteristic triplet has the form* $(d, 0, \nu)$*, where* $d \geq 0$*,* $\nu((-\infty, 0)) = 0$ *and* $\int_{\mathbb{R}+}(x \wedge 1)\nu(\mathrm{d}x) < \infty$.

Proof. S is a subordinator if and only if its paths are monotone (nondecreasing) and its jumps are nonnegative. The monotonicity of the paths is equivalent to their bounded variation, consequently $C = 0$ and $\int_{\mathbb{R}}(x\wedge 1)\nu(\mathrm{d}x) < \infty$, and the jumps are nonnegative if and only if $J_t := \sum_{0\leq s\leq t} \mathbf{1}_{(-\infty,0)}(\Delta S_s) = 0$ for every $t \in T$. Moreover $J_t = 0$ if and only if $\mathbb{E}(J_t) = t\nu((-\infty, 0)) = 0$. Finally it is clear that S is a subordinator if and only if $d \geq 0$. ☐

Remark 18. We saw that a subordinator is such that $\nu((-\infty, 0)) = 0$, we call such a process *spectrally positive*. On the other hand a process X is *spectrally negative* if $\nu((0, \infty)) = 0$.

Corollary 1.2.24. *S is a Lévy subordinator if and only if its symbol takes the form*

$$\eta(u) = idu + \int_{(0,\infty)} (\mathrm{e}^{iux} - 1)\nu(\mathrm{d}x)$$

where $d = b - \int_{(0,1)} x\nu(\mathrm{d}x)$.

Proof. It follows from Theorem 1.2.23 and from the characteristic function formula for Lévy processes. □

We will call the pair (d, ν) the characteristics of the subordinator S.

Example 1.10. (Gamma process) For $\vartheta, \alpha > 0$ define tha gamma law

$$\mu_X(\mathrm{d}x) = \frac{\vartheta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \mathrm{e}^{-\vartheta x} \mathbf{1}_{(0,\infty)}(x)\mathrm{d}x.$$

It is known that

$$\mathbb{E}(\mathrm{e}^{iuX}) = \frac{1}{(1 - iu/\vartheta)^\alpha} = \left[\frac{1}{(1 - iu/\vartheta)^{\alpha/n}}\right]^n$$

and that such a distribution is infinitely divisible. For the Lévy-Khintchine theorem we have $a = \int_{(0,1)} x\nu(\mathrm{d}x)$, $C = 0$ and $\nu(\mathrm{d}x) = \alpha x^{-1}\mathrm{e}^{-\vartheta x}\mathbf{1}_{(0,\infty)}(x)\mathrm{d}x$. This result follows from Frullani integral

$$\frac{1}{(1 - z/\vartheta)^\alpha} = \exp\left[\int_0^\infty (\mathrm{e}^{zx} - 1)\alpha x^{-1}\mathrm{e}^{-\vartheta x}\mathrm{d}x\right]$$

for all $\alpha, \vartheta > 0$ and $z \in \mathbb{C}$ such that $\Re(z) \leq 0$. The choice of b in the formula is the necessary quantity to cancel the term $\mathbf{1}_{\{|x|\leq 1\}}$ in the integral with respect to ν. Then there exists a Lévy process L such that L_1 has Lévy-Khintchine formula given by Frullani integral; this process is called gamma process. Hence the gamma process is a subordinator with $d = 0$ and $\nu(\mathrm{d}x) = \alpha x^{-1}\mathrm{e}^{-\vartheta x}\mathbf{1}_{(0,\infty)}(x)\mathrm{d}x$. The Lévy measure is such that $\int_{\mathbb{R}^+}(x \wedge 1)\nu(\mathrm{d}x) < \infty$. In fact

$$\alpha \int_1^\infty \frac{\mathrm{e}^{-\vartheta x}}{x}\mathrm{d}x + \alpha \int_0^1 \mathrm{e}^{-\vartheta x}\mathrm{d}x < \infty.$$

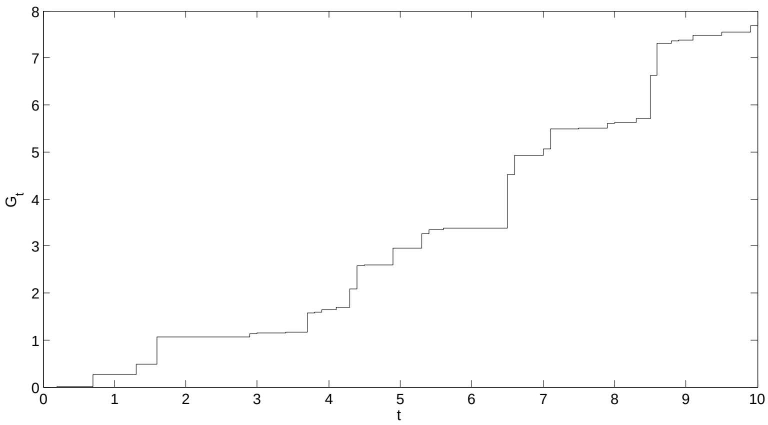

Figure 1.6: Gamma process sample path with $\alpha = \vartheta = 1$

Example 1.11. (Poisson process) Poisson processes are clearly subordinators. Compound Poisson processes are subordinators if and only if the random variable J_n are all non-negative-valued.

Remark 19. Clearly every subordinator is of finite variation, since its paths are nondecreasing.

Theorem 1.2.25. *(Weak law of large numbers for subordinators) If S is a subordinator with drift d and Lévy measure ν, then $S_t/t \xrightarrow{\mathrm{p}} d$ as $t \downarrow 0$.*

Proof. We can write $S_t = dt + \tilde{S}_t$, where $\tilde{S}_t$ is a subordinator with characteristic symbol $t\int_{(0,\infty)}(\mathrm{e}^{iux} - 1)\nu(\mathrm{d}x)$.

$$\lim_{t\to 0}\mathbb{E}(\mathrm{e}^{-iu\tilde{S}_t/t}) = \lim_{t\to 0}\exp\left[t\int_{(0,\infty)}(\mathrm{e}^{-iux/t} - 1)\nu(\mathrm{d}x)\right] = 1$$

for every $u \in \mathbb{R}$ thanks to the dominated convergence. This proves that $\tilde{S}_t/t$ converges in distribution to 0, and hence in probability. Consequently S_t converges in probability to d. □

A law of large numbers holds for a general Lévy process.

Theorem 1.2.26. *Let L be a Lévy process on $\mathbb{R}^d$. If $\mathbb{E}(|L_1|)$ is finite, then a.s. $\lim_{t\to\infty} L_t/t = \mathbb{E}(L_1)$.*

Proof. Let L_n be a random walk on $\mathbb{R}^d$ and suppose that $\mathbb{E}(|L_1|)$ is finite, then by Kolmogorov strong law of large numbers a.s. $n^{-1}L_n \to \mathbb{E}(L_1)$. Moreover, since $t^{-1}L_t = (t^{-1}n)(n^{-1}L_n + n^{-1}(L_t - L_n))$ it is enough to show

$$n^{-1}\sup_{t\in[n,n+1]}|L_t - L_n| \to 0 \quad \text{a.s.}$$

for $n \to \infty$. Let $Y_n = \sup_{t\in[n,n+1]}|L_t - L_n|$, then it is a sequence of i.i.d. random variables and by Theorem 1.2.21 $\mathbb{E}(Y_1)$ is finite. Hence

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} Y_i = 0 \quad \text{a.s.}.$$

It is clear that also $n^{-1}\sum_{i=1}^{n-1} Y_i = 0 \to 0$ a.s. as $n \to \infty$ since $(n-1)/n \to 1$. It follows that a.s. $n^{-1}Y_n \to 0$. □

We now introduce the time changing theorem. Let L a Lévy process and let S be a subordinator defined on the same space as L such that L and S are independent. We define a new process $Z = (Z_t)_{t\in T}$, where $Z_t := X_{S_t}$ for each $t \in T$.

Theorem 1.2.27. *Z is a Lévy process.*

Proof. Z_0 is clearly 0. To prove stationarity let $0 \le t_1 < t_2 < \infty$, $A \in \mathcal{B}(\mathbb{R})$ and denote as μ_{t_1,t_2} the joint probability of S_{t_1} and S_{t_2}.

$$\begin{aligned}
P(Z_{t_2} - Z_{t_1} \in A) &= P(L_{S_{t_2}} - L_{S_{t_1}} \in A)\\
&= \int_{(0,\infty)}\int_{(0,\infty)} P(L_{s_2} - L_{s_1} \in A)\mu_{t_1,t_2}(\mathrm{d}s_1, \mathrm{d}s_2)\\
&= \int_{(0,\infty)}\int_{(0,\infty)} P(L_{s_2 - s_1} \in A)\mu_{t_1,t_2}(\mathrm{d}s_1, \mathrm{d}s_2)\\
&= P(Z_{t_2 - t_1} \in A)
\end{aligned}$$

by independence of L and S. About the independence of the increments, let $0 \leq t_1 < t_2 < t_3 < \infty$ and write μ_{t_1,t_2,t_3} for the joint law of S_{t_1}, S_{t_2} and S_{t_3}. We define, for $x \in \mathbb{R}$, $h_x : \mathbb{R}^+ \to \mathbb{C}$ by $h_x(s) := \mathbb{E}(\mathrm{e}^{ixL_s})$, and, for $x_1, x_2 \in \mathbb{R}$, define $f_{x_1,x_2} : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{C}$ by

$$f_{x_1,x_2}(u_1, u_2, u_3) := \mathbb{E}(\mathrm{e}^{ix_1(L_{u_2}-L_{u_1})})\mathbb{E}(\mathrm{e}^{ix_2(L_{u_3}-Lu_2)})$$

for $0 \leq u_1 < u_2 < u_3 < \infty$. By conditioning and using independence of X and S and of the increments of L

$$\mathbb{E}(\mathrm{e}^{ix_1(Z_{t_2}-Z_{t_1})+ix_2(Z_{t_3}-Z_{t_2})}) = \mathbb{E}(f_{x_1,x_2}(S_{t_1}, S_{t_2}, S_{t_3})).$$

By stationary increments of L we also have for $0 \leq u_1 < u_2 < u_3 < \infty$

$$f_{x_1,x_2}(u_1, u_2, u_3) = h_{x_1}(u_2 - u_1)h_{x_2}(u_3 - u_2).$$

Hence

$$\begin{aligned}
\mathbb{E}(\mathrm{e}^{ix_1(Z_{t_2}-Z_{t_1})+ix_2(Z_{t_3}-Z_{t_2})}) &= \mathbb{E}(h_{x_1}(S_{t_2} - S_{t_1})h_{x_2}(S_{t_3} - S_{t_2})) \\
&= \mathbb{E}(h_{x_1}(S_{t_2} - S_{t_1}))\mathbb{E}(h_{x_2}(S_{t_3} - S_{t_2})) \\
&= \mathbb{E}(\mathrm{e}^{ix_1 Z_{t_2-t_1}})\mathbb{E}(\mathrm{e}^{ix_2 Z_{t_3-t_2}}).
\end{aligned}$$

By Kac's theorem we have independence of $Z_{t_2} - Z_{t-1}$ and $Z_{t_3} - Z_{t_1}$. Analogously we can extend to n time intervals.
About stochastic continuity we know that L and S are stochastically continuous, then for any $a > 0$, fixed $\epsilon > 0$, there exists $\delta > 0$ such that $0 < h < \delta$ implies $P(|L_h| > a) < \epsilon/2$, and there exists also $\delta' > 0$ such that $0 < h < \delta'$ implies $P(S_h > \delta) < \epsilon/2$. Hence for $t \in T$ and $h < \min(\delta, \delta')$

$$\begin{aligned}
P(|Z_h| > a) = P(|L_{S_h}| > a) &= \int_{[0,\infty)} P(|L_u| > a)\mu_{S_h}(\mathrm{d}u) \\
&= \int_{[0,+\delta)} P(|L_u| > a)\mu_{S_h}(\mathrm{d}u) + \int_{(\delta,\infty)} P(|L_u| > a)\mu_{S_h}(\mathrm{d}u) \\
&\leq \sup_{0\leq u<\delta} P(|L_u| > a) + P(S_h \geq \delta) \\
&< \epsilon/2 + \epsilon/2 = \epsilon.
\end{aligned}$$

□

Example 1.12. (Variance gamma process) Consider $Z_t = B_{S_t}$ for $t \in T$, where B is a standard Brownian motion and S is an independent gamma subordinator. In this case the name derives from the fact that each Z_t arises by replacing the variance of the normal random variable by a gamma law. We also are calling variance gamma the following process $Z_t = \mu S_t + \sigma B_{S_t}$ for $\sigma > 0$ and $\mu \in \mathbb{R}$. Usually we take $\theta = \alpha = 1/\tau =: D > 0$, with $\tau > 0$.
The characteristic function of Z_t for $t \in T$ and $u \in \mathbb{R}$ is given by

$$\mathbb{E}(\mathrm{e}^{iuZ_t}) = \left(1 - iu\mu\tau + \frac{1}{2}\sigma^2\tau u^2\right)^{-t/\tau}.$$

In this case the Lévy measure is defined as

$$\nu(\mathrm{d}x) = \begin{cases} \frac{a_n^2}{b_n^2} \frac{\exp\left(-\frac{a_n}{b_n}|x|\right)}{|x|}\mathrm{d}x & x < 0 \\ \frac{a_p^2}{b_p^2} \frac{\exp\left(-\frac{a_p}{b_p}|x|\right)}{|x|}\mathrm{d}x & x > 0 \end{cases}$$

where

$$a_p = \frac{1}{2}\sqrt{\mu^2 + \frac{2\sigma^2}{\tau}} + \frac{\mu}{2}, \quad b_p = a_p^2\tau$$
$$a_n = \frac{1}{2}\sqrt{\mu^2 + \frac{2\sigma^2}{\tau}} - \frac{\mu}{2}, \quad b_n = a_n^2\tau.$$

Then for $x < 0$

$$\begin{aligned} \nu(\mathrm{d}x) &= \frac{1}{\tau|x|}\exp\left(\frac{x}{\left(\frac{1}{2}\sqrt{\mu^2 + \frac{2\sigma^2}{\tau}} - \frac{\mu}{2}\right)\tau}\right)\mathrm{d}x \\ &= -D\frac{1}{x}\exp\left(x\underbrace{\left(\sqrt{\frac{1}{4}\mu^2\tau^2 + \frac{1}{2}\sigma^2\tau} - \frac{\mu\tau}{2}\right)^{-1}}_{:=G}\right)\mathrm{d}x, \end{aligned}$$

and for $x > 0$

$$\begin{aligned} \nu(\mathrm{d}x) &= \frac{1}{\tau|x|}\exp\left(-\frac{x}{\left(\frac{1}{2}\sqrt{\mu^2 + \frac{2\sigma^2}{\tau}} + \frac{\mu}{2}\right)\tau}\right)\mathrm{d}x \\ &= D\frac{1}{x}\exp\left(-x\underbrace{\left(\sqrt{\frac{1}{4}\mu^2\tau^2 + \frac{1}{2}\sigma^2\tau} + \frac{\mu\tau}{2}\right)^{-1}}_{:=M}\right)\mathrm{d}x. \end{aligned}$$

Thus,

$$\nu(\mathrm{d}x) = \begin{cases} -D\exp(Gx)x^{-1}\mathrm{d}x & x < 0 \\ D\exp(-Mx)x^{-1}\mathrm{d}x & x > 0, \end{cases}$$

i.e.

$$\nu(\mathrm{d}x) = D\exp(-G|x|)|x|^{-1}\mathbf{1}_{(0,\infty)}(x)\mathrm{d}x + D\exp(-M|x|)|x|^{-1}\mathbf{1}_{(-\infty,0)}(x)\mathrm{d}x.$$

Furthermore this is a pure jump process because it has no gaussian component. About its moments one can prove that for $t \in T$

$$\mathbb{E}(Z_t) = \mu t \qquad \mathbb{V}\mathrm{ar}(Z_t) = (\mu^2\tau + \sigma^2)t. \tag{1.9}$$

The variance gamma process is an infinite activity pure jump Lévy process which has been used to model log-returns, see [33] for a first application.

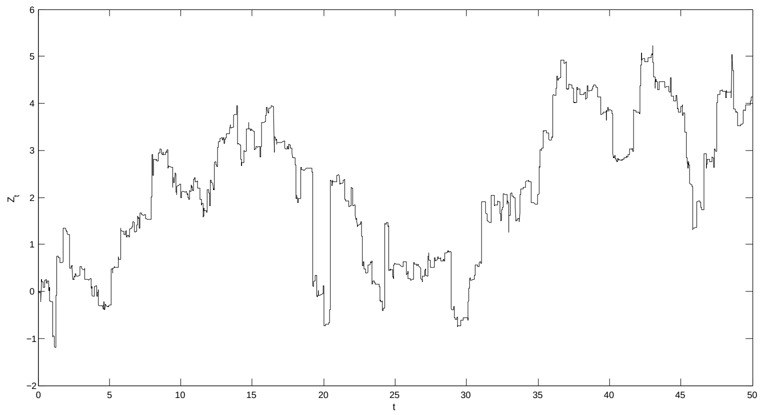

Figure 1.7: Variance gamma process sample path with $\tau = 0.5$, $\mu = 0$ and $\sigma = 1$

1.2.7 Martingales, semimartingales and Lévy processes

Proposition 1.2.28. *Let N, defined by $N(t, A) = \#\{0 \le s \le t : \Delta L_s \in A\}$, be the Poisson random measure and let $f : \mathbb{R}^d \to \mathbb{R}^d$ an integrable function. If A is bounded below, then the process*

$$M_t := \int_A f(x)\tilde{N}(t, \mathrm{d}x) = \int_A f(x)N(t, \mathrm{d}x) - t\int_A f(x)\lambda(\mathrm{d}x)$$

is a martingale. Moreover if $f \in L^2$ M is a square integrable martingale and

$$\mathbb{E}(|M_t|^2) = t\int_A |f(x)|^2\lambda(\mathrm{d}x).$$

Proof. M is a compensated compound Poisson process by Theorem 1.2.18, then it is a martingale. From the same theorem we have that it also is a L^2-martingale. □

It is interesting to get when a Lévy process L is a martingale, for this reason we prove an interesting sufficient and necessary condition.

Proposition 1.2.29. *Let L a Lévy process with triplet (b, C, ν) and assume that $\mathbb{E}(|L_t|) < \infty$. Then L is a martingale if and only if $\tilde{b} = b + \int_{|x|>1} |x|\nu(\mathrm{d}x) = 0$. Similarly L is a submartingale if $\tilde{b} > 0$ and a supermartingale if $\tilde{b} < 0$.*

Proof. By Lévy-Itô decomposition

$$\begin{aligned} L_t &= bt + \sqrt{C}B_t + \int_{|x|>1} xN(t, \mathrm{d}x) + \int_{|x|\le 1} x\tilde{N}(t, \mathrm{d}x) \\ &= bt + \sqrt{C}B_t + \int_{\mathbb{R}^d} x\tilde{N}(t, \mathrm{d}x) + t\int_{|x|>1} x\nu(\mathrm{d}x) \\ &= \tilde{b}t + \sqrt{C}B_t + \int_{\mathbb{R}^d} x\tilde{N}(t, \mathrm{d}x) \end{aligned}$$

as L has finite first moment if and only if $\int_{|x|>1} |x|\nu(\mathrm{d}x) < \infty$. Then L is a martingale if and only if $\tilde{b} = 0$ since the processes $\sqrt{C}B_t$ and $\int_{\mathbb{R}^d} x\tilde{N}(t, \mathrm{d}x)$ are martingales. □

Remark 20. The decomposition $L_t = \tilde{b}t + \sqrt{C}B_t + \int_{\mathbb{R}^d} x\tilde{N}(t, \mathrm{d}x)$ is called canonical, and one observes that $\int_{\mathbb{R}^d} xN(t, \mathrm{d}x) = \sum_{0\le s\le t} \Delta L_s$ by (1.5). Then

$$\begin{aligned} L_t &= \tilde{b}t + \sqrt{C}B_t + \sum_{0\le s\le t} \Delta L_s - t\int_{\mathbb{R}^d} x\nu(\mathrm{d}x) \\ &= bt + \sqrt{C}B_t + \sum_{0\le s\le t} \Delta L_s - t\int_{|x|\le 1} x\nu(\mathrm{d}x). \end{aligned}$$

One can obtain a martingale in the following way too.

Proposition 1.2.30. *Let L be a Lévy process with triplet (b, C, ν), then the process $M_t = \frac{e^{i\langle u, L_t\rangle}}{e^{t\eta(u)}}$ for all $u \in \mathbb{R}^d$ is a complex martingale.*

Proof. $\mathbb{E}(M_t) = 1$ for $t \in T$. And

$$\mathbb{E}(M_t|\mathcal{F}_s) = \mathbb{E}\left(\frac{\mathrm{e}^{i\langle u, L_s\rangle}\mathrm{e}^{i\langle u,(L_t-L_s)\rangle}}{\mathrm{e}^{s\eta(u)}\mathrm{e}^{(t-s)\eta(u)}} \middle| \mathcal{F}_s \right) = M_s\mathbb{E}\left(\frac{\mathrm{e}^{i\langle u,(L_t-L_s)\rangle}}{\mathrm{e}^{(t-s)\eta(u)}} \middle| \mathcal{F}_s \right) = M_s$$

for $0 \le s < t$. □

We now recall briefly the semimartingale definition in order to show that any Lévy proces is a semimartingale.

Definition 1.5. X is a semimartingale if it is an adapted process such that for each $t \in T$

$$X_t = X_0 + M_t + D_t$$

where $M = (M_t)_{t\in T}$ is a local martingale and $D = (D_t)_{t\in T}$ is an adapted process of finite variation.

An important class of semimartingales is given by the following result.

Proposition 1.2.31. *Every Lévy process L is a semimartingale.*

Proof. By the Lévy-Itô decomposition we have for each $t \in T$

$$L_t = M_t + D_t$$

where $M_t = \sqrt{C}B_t + \int_{|x|\le 1} x\tilde{N}(t, \mathrm{d}x)$ is a martingale (it is linear combination of martingales) and $D_t = bt + \int_{|x|>1} xN(t, \mathrm{d}x)$ a process with bounded variation. □

1.2.8 Semimartingales and stochastic calculus

Denote with $\mathbb{D}$ the space of adapted processes with càdlàg paths, with $\mathbb{L}$ we denote the space of adapted processes with càglàd paths and with $\mathbb{S}$ the collection of the simple predictable processes.
We briefly recall the most important results about stochastic integration with semimartingales as integrators. All proofs can be found in Protter [46].

Definition 1.6. (Stochastic integral) For $H \in \mathbb{S}$ and X càdlàg semimartingale, define the linear mapping $J_X : \mathbb{S} \to \mathbb{D}$ by

$$J_X(H) = H_0 X_0 + \sum_{i=1}^{n} H_i (X^{T_{i+1}} - X^{T_i})$$

for $H = H_0 \mathbf{1}_{\{0\}} + \sum_{i=1}^{n} H_i \mathbf{1}_{(T_i, T_{i+1}]}$, $H_i \in \mathcal{F}_{T_i}$ and $0 = T_0 \leq T_1 \leq \cdots \leq T_{n+1} < \infty$ stopping times. We call $J_X(H)$ the stochastic integral of H with respect to X.

Remark 21. One can use different notations. We will use the following ones

$$J_X(H) = \int H_s \mathrm{d}X_s = H \cdot X.$$

In order to enlarge the space of processes we can consider as integrands we introduce the uniform convergence on compacts in probability.

Definition 1.7. A sequence of processes $(H^n)_{n \in \mathbb{N}}$ converges to a process H *uniformly on compacts in probability* (*ucp*) if for each $t > 0$ $\sup_{0 \leq s \leq t} |H_s^n - H_s| \xrightarrow{\mathrm{P}} 0$. If $H_t^* = \sup_{0 \leq s \leq t} |H_s|$, then if $Y^n \in \mathbb{D}$, then $Y^n \to Y$ in *ucp* if $(Y^n - Y)_t^* \xrightarrow{\mathrm{P}} 0$ for every $t > 0$.

We denote with $\mathbb{S}_{ucp}$, $\mathbb{D}_{ucp}$ and $\mathbb{L}_{ucp}$ the respective spaces endowed with the *ucp*-topology. $\mathbb{D}_{ucp}$ is a metrizable space and it is complete under the following metric

$$\mathrm{d}(X, Y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \mathbb{E}[\min(1, (X - Y)_n^*)].$$

To generalize our definition of stochastic integral we also introduce the following theorems.

Theorem 1.2.32. *The space $\mathbb{S}$ is dense in $\mathbb{L}$ under the ucp-topology.*

Theorem 1.2.33. *Let X be a semimartingale. Then the mapping $J_X : \mathbb{L}_{ucp} \to \mathbb{D}_{ucp}$ is continuous.*

The integration operator J_X is continuous on $\mathbb{S}_{ucp}$, and also $\mathbb{S}_{ucp}$ is dense in $\mathbb{L}_{ucp}$. Hence we can extend the integration operator from $\mathbb{S}$ to $\mathbb{L}$ by continuity, since $\mathbb{D}_{ucp}$ is a complete metric space. Therefore we have the following new definition.

Definition 1.8. Let X be a semimartingale. The continuous linear mapping $J_X : \mathbb{L}_{ucp} \to \mathbb{D}_{ucp}$ obtained as the extension of $J_X : \mathbb{S} \to \mathbb{D}$ is called stochastic integral.

Theorem 1.2.34. *(Associativity) The stochastic integral process $Y = H \cdot X$ is itself a semimartingale, and for $G \in \mathbb{L}_{ucp}$ we have*

$$G \cdot Y = G \cdot (H \cdot X) = (GH) \cdot X.$$

It is important to understand the behavior of the jump process of the integral process too.

Theorem 1.2.35. *The jump process $\Delta(H \cdot X)_t$ is indistinguishible from $H_t \Delta X_t$.*

We recall here the definition of the quadratic covariation (or *bracket process*) between two semimartingales which plays a fundamental role.

Definition 1.9. Let X and Y be two semimartingales. The *quadratic variation* process of X, $[X,X] = ([X,X]_t)_{t\geq 0}$, or $[X] = ([X]_t)_{t\geq 0}$, is defined by

$$[X,X] := X^2 - 2\int X_- \mathrm{d}X$$

and the *quadratic covariation* of X and Y is defined by

$$[X,Y] := XY - \int X_- \mathrm{d}Y - \int Y_- \mathrm{d}X$$

where we set $X_{0-} = 0$ and $Y_{0-} = 0$.

Remark 22. The operation $(X,Y) \mapsto [X,Y]$ is bilinear and symmetric, so we have the *polarization identity*

$$[X,Y] = \frac{1}{2}([X+Y, X+Y] - [X,X] - [Y,Y]).$$

Theorem 1.2.36. *The bracket process $[X,Y]$ of two semimartingales has paths of finite variation on compacts, and it is also a semimartingale.*

Theorem 1.2.37. *Let X and Y be two semimartingales, then*

$[X,X]_0 = X_0^2$ *and* $\Delta[X,X] = (\Delta X)^2$

$[X,Y]_0 = X_0 Y_0$ *and* $\Delta[X,Y] = \Delta X \Delta Y$.

Theorem 1.2.38. *(Integration by parts) Let X and Y be two semimartingales. Then XY is a semimartingale and*

$$XY = \int X_- \mathrm{d}Y + \int Y_- \mathrm{d}X + [X,Y].$$

Theorem 1.2.39. *Let X and Y be two semimartingales, and let H and $K \in \mathbb{L}_{ucp}$. Then*

$$[H\cdot X, K\cdot Y] = \int_0^t H_s K_s \mathrm{d}[X,Y]_s$$

and in particular

$$[H\cdot X, H\cdot X]_t = \int_0^t H_s^2 \mathrm{d}[X,X]_s.$$

Definition 1.10. For a semimartingale X, the process $[X,X]^c$ denotes the path-by-path continuous part of $[X,X]$. Analogously, $[X,Y]^c$ denotes the path-by-path continuous part of $[X,Y]$.

Remark 23. We can write $[X,X]_t = [X,X]_t^c + X_0^2 + \sum_{0<s\leq t}(\Delta X_s)^2$.

Definition 1.11. A semimartingale X is called *quadratic pure jump* if $[X,X]^c = 0$.

Example 1.13. The Poisson process is an example of a quadratic pure jump semimartingale.

Theorem 1.2.40. *If X is adapted, càdlàg, with paths of finite variation on compacts, then X is a quadratic pure jump semimartingale.*

Theorem 1.2.41. *If X is a càdlàg semimartingale and Y is a process of bounded variation, then $[X,Y]_t = \sum_{s\le t} \Delta X_s \Delta Y_s$.*

Theorem 1.2.42. *(Itô's formula - I) Let X be a semimartingale and let f be a C^2 real function. Then $f(X)$ is a semimartingale and the following formula holds:*

$$
\begin{aligned}
f(X_t) - f(X_0) = &\int_{0+}^{t} f'(X_{s-})\mathrm{d}X_s + \frac{1}{2}\int_{0+}^{t} f''(X_{s-})\mathrm{d}[X,X]^c_s + \\
&+ \sum_{0<s\le t} (f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s).
\end{aligned}
\tag{1.10}
$$

Theorem 1.2.43. *(Itô's formula - II) Let $X = (X^1, \dots, X^n)$ be a n-tuple of semimartingales and let $f : \mathbb{R}^n \to \mathbb{R}$ be a function with continuous second order partial derivatives. Then $f(X)$ is a semimartingale and*

$$
\begin{aligned}
f(X_t) - f(X_0) = &\sum_{i=1}^{n} \int_{0+}^{t} \frac{\partial f}{\partial x_i}(X_{s-})\mathrm{d}X^i_s + \\
&+ \frac{1}{2} \sum_{1\le i,j\le n} \int_{0+}^{t} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{s-})\mathrm{d}[X^i, X^j]^c_s + \\
&+ \sum_{0<s\le t} \left[f(X_s) - f(X_{s-}) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X_{s-})\Delta X^i_s \right].
\end{aligned}
\tag{1.11}
$$

Chapter 2

Continuous time GARCH(1,1) processes

Considering data of financial time series, such as log-returns on indexes, a collection of empirical observations, known as *stylized facts*, can be noticed (see [52]).

- Returns series show little serial correlation, but they are not independent;
- series of absolute or squared returns show serial correlation;
- conditional expected returns are close to zero;
- volatility seems to vary over time and has jumps;
- extreme returns appear in clusters (*volatility clustering*);
- volatility is a stochastic process with long-range dependence effect;
- returns and volatility have heavy-tailed (higher moments do not exist) and skewed marginals.

ARCH (autoregressive conditionally heteroscedastic) and GARCH (Generalised ARCH) processes, introduced by Engle [15] and Bollerslev [11], are popular in financial econometrics to capture some of the distinctive features of financial data listed above. Autoregressive conditional heteroscedaticity means that past observations and past volatilities have an impact on the present value of the volatility and therefore on the present observation. Then, we have a time-varying, non constant, conditional volatility. With GARCH processes we can analyse financial time series data discretely, but thanks to the fast development of higher and higher memory capacities of computers it is possible to record more and more data (*high-frequency data*). In order to analyse this huge amount of data an extension from discrete time models to continuous time models is necessary. Continuous time processes are also crucial to model irregularly spaced data too. Various attempts have been made to capture the stylized facts in a continuous time model. A first extension goes back to Nelson [43], who proposed to extend the discrete time model by making diffusion approximations. This leads to stochastic volatility models of the type

$$\mathrm{d}Y_t = \sigma_t \mathrm{d}B_t^{(1)}, \quad \mathrm{d}\sigma_t^2 = (\gamma - \theta\sigma_t^2)\mathrm{d}t + \rho\sigma_t^2 \mathrm{d}B_t^{(2)}, \quad t > 0,$$

where $B^{(1)} = (B_t^{(1)})_{t\geq 0}$ and $B^{(2)} = (B_t^{(2)})_{t\geq 0}$ are independent Brownian motions, and $\gamma > 0$, $\theta \geq 0$ and $\rho \geq 0$. Clearly Y_t models the log-price and σ_t the volatility. The main difference between this model and the original GARCH is the fact that in the GARCH modelling there is one single source of randomness. Moreover such diffusion limits to discrete time GARCH can model heavy tails, but obviously they are not able to model volatility jumps. Related models have been suggested and investigated, many generalisations are based on Lévy processes replacing the Brownian motions and on relaxing the independence property. We refer here to Barndorff-Nielsen and Shephard [4] (see [1] for another Lévy driven model). This stochastic volatility model specifies the volatility as an Ornstein-Uhlenbeck process, driven by a subordinator. Precisely, let $L = (L_t)_{t\geq 0}$ be a subordinator and $\alpha > 0$. Then the squared volatility process $(\tilde{\sigma}_t^2)$ is defined by the stochastic differential equation (SDE)

$$\mathrm{d}\tilde{\sigma}_t^2 = -\alpha\tilde{\sigma}_t^2 \mathrm{d}t + \mathrm{d}L_{\alpha t}, \quad t > 0, \tag{2.1}$$

where $\tilde{\sigma}_0^2$ is a finite random variable independent of L and $\tilde{\sigma}_t := \sqrt{\tilde{\sigma}_t^2}$. The solution of (2.1) is the Ornstein-Uhlenbeck type process

$$\tilde{\sigma}_t^2 = \left(\int_0^t \mathrm{e}^{\alpha s}\mathrm{d}L_{\alpha s} + \tilde{\sigma}_0^2\right)\mathrm{e}^{-\alpha t}, \quad t > 0.$$

The logarithmic price process $\tilde{G} = (\tilde{G}_t)_{t\geq 0}$ is such that

$$\mathrm{d}\tilde{G}_t = (\mu + b\tilde{\sigma}_t^2)\mathrm{d}t + \tilde{\sigma}_t \mathrm{d}B_t, \quad t > 0,$$

with $\tilde{G}_0 = 0$, μ and b real constants and $B = (B_t)_{t\geq 0}$ a standard Brownian motion, independent of $\tilde{\sigma}_0^2$ and L. As the empirical volatility has jumps, it seems to make sense to choose a model driven by a Lévy process. However, this model still has two sources of randomness. In [27] Klüppelberg, Lindner and Maller adopt this idea about jumps and a single noise process and suggest a new continuous time GARCH model, which captures all the stylized facts as the discrete time GARCH does. As noise process they choose a Lévy process and its increments in order to replace the innovations in the discrete time GARCH model. Another characteristic is that stock returns are negative correlated with changes in the volatility; the volatility tends to increase after negative shocks and to fall after positive ones. This effect is called *leverage effect.* In order to capture this effect other models have been proposed. Haug and Czado [22] introduced a continuous time exponential GARCH. Lee [31] proposed a continuous time asymmetric power GARCH process, which contains the continuous time GJR-GARCH model recently analysed in [5].

We show both symmetric and asymmetric models with more details in the following sections. We also like to refer to [12] for COGARCH(p,q) models and to [51] for multivariate COGARCH(1,1) processes.

2.1 From discrete to continuous time GARCH process

We start by defining, for $n \in \mathbb{N}$, the discrete time GARCH(1,1) process, which is given by the following equations

$$Y_n = \epsilon_n \sigma_n \tag{2.2}$$

$$\sigma_n^2 = \beta + \lambda Y_{n-1}^2 + \delta \sigma_{n-1}^2, \tag{2.3}$$

where σ_n is the positive square root of σ_n^2, $(\epsilon_n)_{n\in\mathbb{N}}$ is a sequence of i.i.d. non-degenerate random variables such that $P(\epsilon_1 = 0) = 0$ and the parameters are such that $\beta > 0$, $\lambda \geq 0$ and $\delta \geq 0$. We assume also some initial almost surely finite, and random in general, values for ϵ_0 and σ_0, independent of each other and independent of $(\epsilon_n)_{n\in\mathbb{N}}$. It is clear that when $\delta = 0$ one obtains the ARCH(1) model, instead when $\delta = \lambda = 0$ $(Y_n)_{n\in\mathbb{N}}$ is just a sequence of i.i.d. random variables. For this reason $\lambda + \delta > 0$ is assumed through all the thesis. Details can be found in [15] and [11].

Remark 24. Equation (2.2) specifies the *mean level* process and Equation (2.3) models the conditional volatility process, which is time dependent and randomly fluctuating.

Remark 25. Iterating (2.3)

$$\begin{aligned}
\sigma_n^2 &= \beta + (\delta + \lambda\epsilon_{n-1}^2)\sigma_{n-1}^2 \\
&= \beta + (\delta + \lambda\epsilon_{n-1}^2)(\beta + (\delta + \lambda\epsilon_{n-2}^2)\sigma_{n-2}^2) \\
&= \beta(1 + \delta + \lambda\epsilon_{n-1}^2) + (\delta + \lambda\epsilon_{n-1}^2)(\delta + \lambda\epsilon_{n-2}^2)(\beta + (\delta + \lambda\epsilon_{n-3}^2)\sigma_{n-3}^2) \\
&= \beta[1 + \delta + (\lambda\epsilon_{n-1}^2) + (\delta + \lambda\epsilon_{n-2}^2)(\delta + \lambda\epsilon_{n-3}^2)] + (\delta + \lambda\epsilon_{n-1}^2)(\delta + \lambda\epsilon_{n-2}^2)\cdot \\
&\quad \cdot (\delta + \lambda\epsilon_{n-3}^2)\sigma_{n-3}^2 \\
&= \cdots \\
&= \beta \sum_{i=0}^{n-1} \prod_{j=i+1}^{n-1} (\delta + \lambda\epsilon_j^2) + \sigma_0^2 \prod_{j=0}^{n-1} (\delta + \lambda\epsilon_j^2), \quad n \in \mathbb{N},
\end{aligned}$$

where $\prod_{j=a}^{b} = 1$ if $a > b$. Furthermore

$$\begin{aligned}
\sigma_n^2 &= \beta \sum_{i=0}^{n-1} \prod_{j=i+1}^{n-1} (\delta + \lambda\epsilon_j^2) + \sigma_0^2 \prod_{j=0}^{n-1} (\delta + \lambda\epsilon_j^2) \\
&\cong \beta \int_0^n \exp\left(\sum_{j=\lfloor s\rfloor+1}^{n-1} \log(\delta + \lambda\epsilon_j^2) \right) \mathrm{d}s + \sigma_0^2 \exp\left(\sum_{j=0}^{n-1} \log(\delta + \lambda\epsilon_j^2) \right) \\
&= \beta \int_0^n \exp\left(\sum_{j=0}^{n-1} \log(\delta + \lambda\epsilon_j^2) - \sum_{j=0}^{\lfloor s\rfloor} \log(\delta + \lambda\epsilon_j^2) \right) \mathrm{d}s + \sigma_0^2 \cdot \exp\left(\sum_{j=0}^{n-1} \log(\delta + \lambda\epsilon_j^2) \right) \\
&= \left(\beta \int_0^n \exp\left(- \sum_{j=0}^{\lfloor s\rfloor} \log(\delta + \lambda\epsilon_j^2) \right) \mathrm{d}s + \sigma_0^2 \right) \exp\left(\sum_{j=0}^{n-1} \log(\delta + \lambda\epsilon_j^2) \right).
\end{aligned} \tag{2.4}$$

We write the summation in the following way

$$\sum_{j=0}^{n-1} \log(\delta + \lambda\epsilon_j^2) = n \log\delta + \sum_{j=0}^{n-1} \log\left(1 + \frac{\lambda}{\delta}\epsilon_j^2\right)$$

in order to define, for every $0 < \delta < 1$, the auxiliary càdlàg Lévy process $X = (X_t)_{t\geq 0}$

$$X_t := -t\log\delta - \sum_{0<s\leq t} \log\left(1 + \frac{\lambda}{\delta}(\Delta L_s)^2\right) \tag{2.5}$$

which represents the continuous time equivalent of the summation above taken with the opposite sign. We have replaced the noise variables ϵ_j with the increments of a càdlàg Lévy process $L = (L_t)_{t\geq 0}$. X is clearly right continuous since the summation includes t.
Now, using this process and looking at Equation (2.4), we define the left continuous square volatility process as

$$\sigma_t^2 := \left(\beta \int_0^t \mathrm{e}^{X_s} \mathrm{d}s + \sigma_0^2\right) \mathrm{e}^{-X_{t-}}, \quad t \geq 0 \tag{2.6}$$

with σ_0 finite random variable independent of $(L_t)_{t\geq 0}$.

Remark 26. The volatility process is a generalized Ornstein-Uhlenbeck process.

Once constructed these processes we can define the càdlàg integrated COGARCH(1,1) $G = (G_t)_{t\geq 0}$ as solution of the following SDE

$$\begin{cases} \mathrm{d}G_t = \sigma_t \mathrm{d}L_t \\ G_0 = 0. \end{cases}$$

Remark 27. Surely by Theorems 1.2.34 and 1.2.35 G is a semimartingale jumping at the same times as L does, and $\Delta G_t = \sigma_t \Delta L_t$.

2.1.1 Properties of the auxiliary process

We can begin by investigating the process X, which has a special structure. For this reason we state and prove the following proposition.

Proposition 2.1.1. *X is a spectrally negative Lévy process with bounded variation. Moreover the characteristic triplet is such that $C_X = 0$,*

$$\nu_X([0,\infty)) = 0,$$

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

and

$$\gamma_{X,0} := \gamma_X - \int_{[-1,1]} x\nu_X(\mathrm{d}x) = -\log\delta.$$

Proof. We first prove that X is a Lévy process. It is clear that a.s. $X_0 = 0$. About the stationarity and independence of the increments for $0 < s < t$

$$\begin{aligned} X_t - X_s &= -(t-s)\log\delta - \sum_{s<k\leq t} \log\left(1 + \frac{\lambda}{\delta}(\Delta L_k)^2\right) \\ &\stackrel{\mathrm{d}}{=} -(t-s)\log\delta - \sum_{0<k\leq t-s} \log\left(1 + \frac{\lambda}{\delta}(\Delta L_k)^2\right) = X_{t-s} \end{aligned}$$

and

$$\begin{aligned}\phi_{X_t - X_s, X_s}(u_1, u_2) &= \mathbb{E}(\mathrm{e}^{iu_1 X_{t-s} + iu_2 X_s}) \\ &= \mathrm{e}^{-iu_1(t-s)\log\delta}\mathrm{e}^{-iu_2 s\log\delta} \cdot \\ &\quad \cdot \mathbb{E}\left(\mathrm{e}^{-iu_1\sum_{s<k\le t}\log\left(1+\frac{\lambda}{\delta}(\Delta L_k)^2\right)}\mathrm{e}^{-iu_2\sum_{0<k\le s}\log\left(1+\frac{\lambda}{\delta}(\Delta L_s)^2\right)}\right) \\ &= \mathrm{e}^{-iu_1(t-s)\log\delta}\mathrm{e}^{-iu_2 s\log\delta}\mathbb{E}\left(\mathrm{e}^{-iu_1\sum_{s<k\le t}\log\left(1+\frac{\lambda}{\delta}(\Delta L_k)^2\right)}\right) \cdot \\ &\quad \cdot \mathbb{E}\left(e^{-iu_2\sum_{0<k\le s}\log\left(1+\frac{\lambda}{\delta}(\Delta L_k)^2\right)}\right) \\ &= \phi_{X_{t-s}}(u_1)\phi_{X_s}(u_2).\end{aligned}$$

Now, we just have to prove stochastic continuity.

$$\begin{aligned}\lim_{t\downarrow 0} P(|X_t| > \epsilon) &= \lim_{t\downarrow 0} P\left(\left|-t\log\delta - \sum_{0<s\le t}\log\left(1+\frac{\lambda}{\delta}(\Delta L_s)^2\right)\right| > \epsilon\right) \\ &\le \lim_{t\downarrow 0} P(t|\log\delta| > \epsilon/2) + \lim_{t\downarrow 0}\frac{2}{\epsilon}\sum_{0<s\le t}\mathbb{E}\left(\log\left(1+\frac{\lambda}{\delta}(\Delta L_s)^2\right)\right) \\ &= \lim_{t\downarrow 0} P(\epsilon < -2t\log\delta) \\ &= \lim_{t\downarrow 0}\mathbf{1}_{(-\infty,-2t\log\delta)}(\epsilon) = 0.\end{aligned}$$

Then X is a Lévy process. Surely it has no positive jumps since

$$\Delta X_t = X_t - \lim_{s\uparrow t} X_s = -\log\left(1+\frac{\lambda}{\delta}(\Delta L_t)^2\right) < 0 \quad \forall t > 0.$$

This means that for $A \subseteq \mathbb{R}_+$

$$\begin{aligned}\nu_X(A) &= \mathbb{E}\left(\sum_{0<s\le 1}\mathbf{1}_A(\Delta X_s)\right) = \\ &= \mathbb{E}\left[\sum_{0<s\le 1}\mathbf{1}_A\left(-\log\left(1+\frac{\lambda}{\delta}(\Delta L_s)^2\right)\right)\right] = 0.\end{aligned}$$

Instead, for $x > 0$

$$\begin{aligned}\nu_X((-\infty,-x]) &= \mathbb{E}\left[\sum_{0<s\le 1}\mathbf{1}_{(-\infty,-x]}\left(-\log\left(1+\frac{\lambda}{\delta}(\Delta L_s)^2\right)\right)\right] \\ &= \mathbb{E}\left[\sum_{0<s\le 1}\mathbf{1}_{[\mathrm{e}^x,\infty)}\left(1+\frac{\lambda}{\delta}(\Delta L_s)^2\right)\right] \\ &= \mathbb{E}\left[\sum_{0<s\le 1}\mathbf{1}_{[\sqrt{(\mathrm{e}^x-1)\delta/\lambda},\infty)}(|\Delta L_s|)\right] \\ &= \nu_L\left(\left\{y\in\mathbb{R} : |y| \ge \sqrt{(\mathrm{e}^x-1)\delta/\lambda}\right\}\right).\end{aligned}$$

Hence ν_X is the image measure of ν_L under the transformation $T : \mathbb{R} \to (-\infty, 0]$ $x \mapsto -\log\left(1 + \frac{\lambda}{\delta}x^2\right)$. Moreover

$$\int_{|x|\le 1} |x|\nu_X(\mathrm{d}x) = \int_{|y|\le\sqrt{(\mathrm{e}-1)\delta/\lambda}} \log\left(1+\frac{\lambda}{\delta}y^2\right)\nu_L(\mathrm{d}y)$$
$$\le \frac{\lambda}{\delta}\int_{|y|\le\sqrt{(\mathrm{e}-1)\delta/\lambda}} y^2\nu_L(\mathrm{d}y).$$

We observe that if $\sqrt{(\mathrm{e}-1)\delta/\lambda} \le 1$

$$\int_{|y|\le\sqrt{(\mathrm{e}-1)\delta/\lambda}} y^2\nu_L(\mathrm{d}x) \le \int_{|y|\le 1} y^2\nu_L(\mathrm{d}y) < \infty,$$

and if $\sqrt{(\mathrm{e}-1)\delta/\lambda} > 1$

$$\begin{aligned}\int_{|y|\le\sqrt{(\mathrm{e}-1)\delta/\lambda}} y^2\nu_L(\mathrm{d}x) &= \int_{|y|\le 1} y^2\nu_L(\mathrm{d}y) + \int_{[-\sqrt{(\mathrm{e}-1)\delta/\lambda},-1)} y^2\nu_L(\mathrm{d}y)+\\ &+\int_{(1,\sqrt{(\mathrm{e}-1)\delta/\lambda}]} y^2\nu_L(\mathrm{d}y)\\ &\le c + \frac{\delta}{\lambda}(\mathrm{e}-1)\int_{[-\sqrt{(\mathrm{e}-1)\delta/\lambda},-1)} \nu_L(\mathrm{d}y)\\ &+\frac{\delta}{\lambda}(\mathrm{e}-1)\int_{(1,\sqrt{(\mathrm{e}-1)\delta/\lambda}]} \nu_L(\mathrm{d}y)\\ &< \infty.\end{aligned}$$

Moreover by the Lévy-Itô decomposition we know that for every Lévy process L such that $\int_{|x|\le 1}|x|\nu_L(\mathrm{d}x)$ is finite

$$L_t = \gamma_L t + \sqrt{C_L}B_t + \sum_{0<s\le t}\Delta L_s - t\int_{|x|\le 1} x\nu_L(\mathrm{d}x),$$

and since in our case $X_t = -t\log\delta - \sum_{0<s\le t}\Delta X_s$, then $C_X = 0$ and $\gamma_X - \int_{[-1,1]} x\nu_X(\mathrm{d}x) = -\log\delta$.
As $C_X = 0$ and $\int_{[-1,1]}|x|\nu_X(\mathrm{d}x)$ is finite, Theroem 1.2.19 implies that X is a process of bounded variation (in effect for fixed $\omega\in\Omega$ it is difference of two non decreasing processes: $-t\log\delta$ and $\sum_{0<s\le t}\log\left(1+\frac{\lambda}{\delta}(\Delta L_s)^2\right)$).
By Lévy-Khintchine formula (Theorem 1.2.3) we finally get that

$$\mathbb{E}(\mathrm{e}^{iuX_t}) = \exp\left(-itu\log\delta + t\int_{(-\infty,0)}(\mathrm{e}^{iux}-1)\nu_X(\mathrm{d}x)\right),$$

which shows that X is the negative of a subordinator with positive drift. □

2.2 The volatility process: stability, markovianity and moments

Firstly we prove that the square volatility process is solution, like G, of a stochastic differential equation.

Proposition 2.2.1. *The process* $(\sigma_t^2)_{t\geq 0}$ *satisfies the following stochastic differential equation*

$$\mathrm{d}\sigma_{t+}^2 = \beta \mathrm{d}t + \sigma_t^2 e^{X_{t-}} \mathrm{d}(\mathrm{e}^{-X_t}), \quad t > 0$$

and we have

$$\sigma_t^2 = \sigma_0^2 + \beta t + \log\delta \int_0^t \sigma_s^2 \mathrm{d}s + \frac{\lambda}{\delta} \sum_{0<s<t} \sigma_s^2 (\Delta L_s)^2. \tag{2.7}$$

Proof. Set $K_t := t\log\delta$ and $S_t := \prod_{0<s\leq t}\left(1+\frac{\lambda}{\delta}(\Delta L_s)^2\right)$; by Definition 1.5 they are semimartingales since they are processes of finite variation. Let $f(k,s) = \mathrm{e}^k s$ and apply Ito's lemma in two variables for semimartingales (Theorem 1.2.43)

$$f(K_t, S_t) = \mathrm{e}^{K_t} S_t = \mathrm{e}^{-X_t},$$

$$\begin{aligned}\mathrm{e}^{-X_t} &= 1 + \int_{0+}^t \mathrm{e}^{K_{s-}} \mathrm{d}S_s + \int_{0+}^t \mathrm{e}^{-X_{s-}} \mathrm{d}K_s + \frac{1}{2}\int_{0+}^t \mathrm{e}^{-X_{s-}} \mathrm{d}[K,K]_s^c + \\ &+ \int_{0+}^t \mathrm{e}^{K_{s-}} \mathrm{d}[K,S]_s^c + \sum_{0<s\leq t} (\mathrm{e}^{-X_s} - \mathrm{e}^{-X_{s-}} - \mathrm{e}^{K_{s-}}\Delta S_s - \mathrm{e}^{X_{s-}}\Delta K_s).\end{aligned}$$

Clearly $\Delta K_S = 0$ for $s > 0$ and $\mathrm{e}^{-X_s} - \mathrm{e}^{-X_{s-}} = \mathrm{e}^{K_s} S_s - \mathrm{e}^{K_s} S_{s-} = \mathrm{e}^{K_s}\Delta S_s$ so that the summation vanishes. Then

$$\begin{aligned}\mathrm{e}^{-X_t} &= 1 + \int_{0+}^t \mathrm{e}^{K_{s-}} \mathrm{d}S_s + \log\delta \int_0^t \mathrm{e}^{-X_{s-}} \mathrm{d}s + \frac{1}{2}\int_{0+}^t \mathrm{e}^{-X_{s-}} \mathrm{d}[K,K]_s^c + \\ &+ \int_{0+}^t \mathrm{e}^{K_{s-}} \mathrm{d}[K,S]_s^c \\ &= 1 + \int_0^t \mathrm{e}^{K_s} \mathrm{d}S_s + \log\delta \int_0^t \mathrm{e}^{-X_{s-}} \mathrm{d}s\end{aligned}$$

because $[K,K]_t = C_K t + \sum_{0<s\leq t} (\Delta K_s)^2 = 0$ and by Theorem 1.2.41

$$[K,S]_t = \sum_{0<s\leq t} (\Delta K_s \Delta S_s) = 0.$$

S ia a pure jump process, hence $\int_0^t \mathrm{e}^{K_s} \mathrm{d}S_s = \frac{\lambda}{\delta}\sum_{0<s\leq t} \mathrm{e}^{-X_{s-}} (\Delta L_s)^2$.
Integration by parts (Theorem 1.2.38) gives

$$\begin{aligned}\mathrm{e}^{-X_t}\int_0^t \mathrm{e}^{X_s}\mathrm{d}s &= \int_{0+}^t \mathrm{e}^{-X_{s-}} \mathrm{d}\left(\int_0^s \mathrm{e}^{X_u}\mathrm{d}u\right) + \int_{0+}^t \left(\int_0^{s-} \mathrm{e}^{X_u}\mathrm{d}u\right)\mathrm{d}(\mathrm{e}^{-X_s}) + \\ &+ [\mathrm{e}^{-X_\cdot}, \int_0^\cdot \mathrm{e}^{X_S}\mathrm{d}s]_t \\ &= \int_{0+}^t \mathrm{e}^{-X_{s-}} \mathrm{d}\left(\int_0^s \mathrm{e}^{X_u}\mathrm{d}u\right) + \int_{0+}^t \left(\int_0^{s} \mathrm{e}^{X_u}\mathrm{d}u\right)\mathrm{d}(\mathrm{e}^{-X_s}) + \\ &+ [\mathrm{e}^{-X_\cdot}, \int_0^\cdot \mathrm{e}^{X_s}\mathrm{d}s]_t.\end{aligned}$$

By the associativity of the stochastic integral (Theorem 1.2.34) and as $X_s = X_{s-}$ for fixed s almost surely

$$\int_{0+}^{t} \mathrm{e}^{-X_{s-}} \mathrm{d}\left(\int_0^s \mathrm{e}^{X_u} \mathrm{d}u\right) = \int_{0+}^{t} \mathrm{e}^{-X_{s-}} \mathrm{e}^{X_s} \mathrm{d}s = t.$$

Moreover

$$[\mathrm{e}^{-X_\cdot}, \int_0^\cdot \mathrm{e}^{X_S} \mathrm{d}s]_t = [\int_0^t \mathrm{e}^{-X_s} \mathrm{d}\mathbf{1}_{[t,\infty)}(s), \int_0^t \mathrm{e}^{X_s} \mathrm{d}s] = \int \mathrm{d}[\mathbf{1}_{[t,\infty)}(s), s].$$

$[\mathbf{1}_{[t,\infty)}(s), s] = s\mathbf{1}_{[t,\infty)}(s) - \int_0^s u \mathrm{d}\mathbf{1}_{[t,\infty)}(u) - \int_0^s \mathbf{1}_{[t,\infty)}(u)\mathrm{d}u$, and if $s \geq t$ $[\mathbf{1}_{[t,\infty)}(s), s] = s - t - (s-t) = 0$ as like if $s < t$. Hence $[\mathrm{e}^{-X_\cdot}, \int_0^\cdot \mathrm{e}^{X_S}\mathrm{d}s]_t = 0$ and

$$\mathrm{e}^{-X_t} \int_0^t \mathrm{e}^{X_s} \mathrm{d}s = t + \int_{0+}^{t} \left(\int_0^{s-} \mathrm{e}^{X_u} \mathrm{d}u\right) \mathrm{d}(\mathrm{e}^{-X_s}).$$

Then

$$\begin{aligned}
\sigma_{t+}^2 &= \beta \int_0^t \mathrm{e}^{X_s - X_t} \mathrm{d}s + \sigma_0^2 \mathrm{e}^{-X_t} \\
&= \beta \left(t + \int_{0+}^{t} \left(\int_0^s \mathrm{e}^{X_u} \mathrm{d}u\right) \mathrm{d}(\mathrm{e}^{-X_s})\right) + \sigma_0^2 \mathrm{e}^{-X_t} \\
&= \beta \left(t + \int_{0+}^{t} \left(\int_0^s \mathrm{e}^{X_u - X_{s-}}\right) \mathrm{e}^{X_{s-}} \mathrm{d}(\mathrm{e}^{-X_s})\right) + \sigma_0^2 \mathrm{e}^{-X_t} \\
&= \beta t + \int_{0+}^{t} (\sigma_s^2 - \mathrm{e}^{-X_{s-}} \sigma_0^2) \mathrm{e}^{X_{s-}} \mathrm{d}(\mathrm{e}^{-X_s}) + \sigma_0^2 \mathrm{e}^{-X_t} \\
&= \beta t + \int_{0+}^{t} \sigma_s^2 \mathrm{e}^{X_s} \mathrm{d}(\mathrm{e}^{-X_s}) - \sigma_0^2 \int_{0+}^{t} \mathrm{d}(\mathrm{e}^{-X_s}) + \sigma_0^2 \mathrm{e}^{-X_t} \\
&= \beta t + \int_{0+}^{t} \sigma_s^2 \mathrm{e}^{X_s} \mathrm{d}(\mathrm{e}^{-X_s}) - \sigma_0^2 (\mathrm{e}^{-X_t} - 1) + \sigma_0^2 \mathrm{e}^{-X_t} \\
&= \beta t + \int_{0+}^{t} \sigma_s^2 \mathrm{e}^{X_s} \mathrm{d}(\mathrm{e}^{-X_s}) + \sigma_{0+}^2
\end{aligned}$$

because $\sigma_t^2 = \sigma_{t+}^2$ almost surely. Then

$$\mathrm{d}\sigma_{t+}^2 = \beta \mathrm{d}t + \sigma_t^2 \mathrm{e}^{X_{t-}} \mathrm{d}(\mathrm{e}^{-X_t}), \quad t > 0.$$

We also can observe that

$$\begin{aligned}
\sigma_{t+}^2 &= \beta t + \int_{0+}^{t} \sigma_s^2 \mathrm{e}^{X_{s-}} \mathrm{d}(\mathrm{e}^{-X_s}) + \sigma_{0+}^2 \\
&= \sigma_{0+}^2 + \beta t + \log \delta \int_0^t \sigma_s^2 \mathrm{d}s + \int_0^t \mathrm{e}^{X_{s-}} \sigma_s^2 \mathrm{d}\left(\frac{\lambda}{\delta} \sum_{0<u\leq s} \mathrm{e}^{-X_{u-}} (\Delta L_u)^2\right) \\
&= \sigma_{0+}^2 + \beta t + \log \delta \int_0^t \sigma_s^2 \mathrm{d}s + \sum_{0<s\leq t} \sigma_s^2 (\Delta L_s)^2,
\end{aligned}$$

or equivalently

$$\sigma_t^2 = \sigma_0^2 + \beta t + \log\delta \int_0^t \sigma_s^2 \mathrm{d}s + \frac{\lambda}{\delta} \sum_{0<s<t} \sigma_s^2 (\Delta L_s)^2. \tag{2.8}$$

Hence, remembering that $[L,L]_t = C_L t + \sum_{0<s\le t} (\Delta L_s)^2 = C_L t + [L,L]_t^d$, we get that

$$\mathrm{d}\sigma_{t+}^2 = (\beta + \log\delta\sigma_t^2)\mathrm{d}t + \frac{\lambda}{\delta}\sigma_t^2 \mathrm{d}[L,L]_t^d. \tag{2.9}$$

□

Remark 28. In GARCH(1,1) model $\sigma_{n+1}^2 = \beta + \lambda\sigma_n^2\epsilon_n^2 + \delta\sigma_n^2$, so

$$\sigma_{n+1}^2 - \sigma_n^2 = \beta + \delta\sigma_n^2 - \sigma_n^2 + \lambda\sigma_n^2\epsilon_n^2 = \beta - (1-\delta)\sigma_n^2 + \lambda\sigma_n^2\epsilon_n^2$$

similarly to (2.9). Or alternatively by summation

$$\sigma_n^2 = \beta n - (1-\delta)\sum_{i=0}^{n-1} \sigma_i^2 + \lambda \sum_{i=0}^{n-1} \sigma_i^2\epsilon_i^2 + \sigma_0^2$$

which is similar to (2.8). These representations (both in discrete and in continuous time) show feedback and autoregressive aspects.

To study the stability it is interesting to understand when the volatility process converges as $t \to \infty$. For this reason we introduce and prove the following theorem.

Theorem 2.2.2. *Suppose that*

$$\int_{\mathbb{R}} \log\left(1 + \frac{\lambda}{\delta} y^2\right) \nu_L(\mathrm{d}y) < -\log\delta, \tag{2.10}$$

then the process $(\sigma_t^2)_{t\ge0}$ converges weakly, as $t \to \infty$, to a finite random variable σ_∞^2 such that

$$\sigma_\infty^2 \overset{\mathrm{d}}{=} \beta \int_0^\infty \mathrm{e}^{-X_t} \mathrm{d}t.$$

Conversely, if (2.10) does not hold, then $\sigma_t^2 \xrightarrow{\mathrm{P}} \infty$ as $t \to \infty$.

Proof. By a continuous time analogue of the Goldie-Maller theorem (see Erickson and Maller [16] or [8]) $\int_0^\infty \mathrm{e}^{-X_s}\mathrm{d}s$ converges almost surely to a finite random variable if $X_t \to \infty$ a.s. and $\sigma_t^2 \xrightarrow{\mathrm{P}} \infty$ as $t \to \infty$ otherwise. By definition

$$\sigma_t^2 = \beta \int_0^t \mathrm{e}^{X_s - X_{t-}} \mathrm{d}s + \sigma_0^2 \mathrm{e}^{-X_{t-}},$$

but, as we are interested in its behavior when $t \to \infty$, we will not write X_{t-}, but X_t. If $X_t \to \infty$ we have that a.s. $\sigma_0^2\mathrm{e}^{-X_t} \to 0$. Moreover if we set $t - s = u$

$$\beta \int_0^t \mathrm{e}^{X_s - X_t}\mathrm{d}s \overset{\mathrm{d}}{=} \beta \int_0^t \mathrm{e}^{-X_{t-s}}\mathrm{d}s = \beta \int_0^t \mathrm{e}^{-X_s}\mathrm{d}s.$$

So, we only need to show that

$$X_t \overset{a.s.}{\to} \infty \Leftrightarrow \int_{\mathbb{R}} \log\left(1+\frac{\lambda}{\delta}y^2\right)\nu_L(\mathrm{d}y) < -\log\delta.$$

$$\begin{aligned}\mathbb{E}(X_1) &= -i\phi'_{X_1}(0)\\ &= -i\left[\exp\left(-iu\log\delta + \int_{(-\infty,0)}(\mathrm{e}^{iux}-1)\nu_X(\mathrm{d}x)\right)\right]_{u=0}\cdot\\ &\quad\cdot\left(-i\log\delta + i\int_{(-\infty,0)}(\mathrm{e}^{iux}x)\nu_X(\mathrm{d}x)\right)_{u=0}\\ &= -\log\delta + \int_{(-\infty,0)} x\nu_X(\mathrm{d}x),\end{aligned}$$

then $\mathbb{E}(X_1)$ always exists ($\nu_X([0,\infty))=0$), possibly $\mathbb{E}(X_1)=-\infty$ and $X_t/t \to \mathbb{E}(X_1)$ almost surely as $t\to\infty$ by Theorem 1.2.26. If $\mathbb{E}(X_1)\le 0$, then $X_t\to-\infty$ or oscillates. So $X_t\to\infty$ a.s. if and only if $\mathbb{E}(X_1)>0$ (see Theorem 1 in Bertoin and Yor [8]). Indeed if $X_t/t\to\infty$ a.s. X_t cannot converge for its special structure and cannot diverge to $-\infty$ because $\lim_{t\to\infty} X_t = 0\cdot\lim_{t\to\infty} t\ge 0$. So it remains to prove that X_t cannot go to ∞ if $\mathbb{E}(X_1)=0$. Suppose that $X_t\to\infty$, which is equivalent to

$$\lim_{t\to\infty} -t\log\delta - \lim_{t\to\infty}\sum_{0<s\le t}\log\left(1+\frac{\lambda}{\delta}(\Delta L_s)^2\right) = \infty.$$

Then $A_t := \sum_{0<s\le t}\log\left(1+\frac{\lambda}{\delta}(\Delta L_s)^2\right)$ converges to $c\in[0,\infty)$ or diverges to ∞ but slower than $-t\log\delta$. If $A_t\to c$, $X_t/t\to-\log\delta>0$ and we have the first contradiction. If $A_t\to\infty$ (slower than t), then $X_t/t\to\log\delta>0$ and we have the same contradiction. In the last possible case, i.e. $A_t\asymp t$ ($A_t/t\to k>0$), $X_t/t\to-\log\delta-k$. The last quantity is greater than zero because $\lim_{t\infty} X_t = \lim_{t\to\infty}(-t\log\delta - kt)=\infty$ if and only if $-\log\delta>k$ and since $X_t/t\to\mathbb{E}(X_1)$ $A_t/t\to k=\int_{\mathbb{R}}\log\left(1+\frac{\lambda}{\delta}y^2\right)\nu_L(\mathrm{d}y)$. By our assumption $k<-\log\delta$ so that $\mathbb{E}(X_1)>0$ if and only if

$$-\log\delta + \int_{(-\infty,0)} x\nu_X(\mathrm{d}x) = -\log\delta - \int_{\mathbb{R}}\log\left(1+\frac{\lambda}{\delta}y^2\nu_L(\mathrm{d}y)\right) > 0.$$

$\square$

Another property we are going to prove is the markovianity of $(\sigma_t^2)_{t\ge0}$.

Theorem 2.2.3. *$(\sigma_t^2)_{t\ge0}$ is a Markov process, moreover if σ_∞^2 exists and $\sigma_0^2 \overset{\mathrm{d}}{=} \sigma_\infty^2$ (independent of $(L_t)_{t\ge0}$), then $(\sigma_t^2)_{t\ge0}$ is strictly stationary*

Proof. Let $(\mathcal{F}_t)_{t\geq 0}$ be the filtration generated by $(\sigma_t^2)_{t\geq 0}$ and let $y \in [0,t)$,

$$\begin{aligned}
\sigma_t^2 &= \beta \int_0^y \mathrm{e}^{X_s} \mathrm{d}s \mathrm{e}^{-X_{t-}} + \beta \int_y^t \mathrm{e}^{X_s} \mathrm{d}s \mathrm{e}^{-X_{t-}} + \sigma_0^2 \mathrm{e}^{-X_{t-}} \\
&= \beta \int_0^y \mathrm{e}^{X_s} \mathrm{d}s \mathrm{e}^{-X_{y-}} \mathrm{e}^{X_{y-}-X_{t-}} + \beta \int_0^y \mathrm{e}^{X_s} \mathrm{d}s \mathrm{e}^{-X_{t-}} + \sigma_0^2 \mathrm{e}^{-X_{t-}} \\
&= (\sigma_y^2 - \sigma_0^2 \mathrm{e}^{-X_{y-}}) \mathrm{e}^{-(X_{t-}-X_{y-})} + \beta \int_0^y \mathrm{e}^{X_s} \mathrm{d}s \mathrm{e}^{-X_{t-}} + \sigma_0^2 \mathrm{e}^{-X_{t-}} \\
&= \sigma_y^2 \mathrm{e}^{-(X_{t-}-X_{y-})} + \beta \int_0^y \mathrm{e}^{X_s} \mathrm{d}s \mathrm{e}^{-X_{t-}} \\
&= \sigma_y^2 A_{y,t} + B_{y,t}
\end{aligned}$$

where $A_{y,t} := \mathrm{e}^{-(X_{t-}-X_{y-})}$ and $B_{y,t} := \beta \int_y^t \mathrm{e}^{(X_s - X_{y-})} \mathrm{d}s \mathrm{e}^{-(X_{t-}-X_{y-})}$, independent of $(\mathcal{F}_t)_{t\geq 0}$. Then conditioning to $\mathcal{F}_y$ σ_t^2 depends only on σ_y^2. Now set $\sigma_0^2 \overset{\mathrm{d}}{=} \sigma_\infty^2$ in order to prove that $\sigma_t^2 \overset{\mathrm{d}}{=} \sigma_\infty^2$ for $t > 0$. We can take $\sigma_0^2 := \beta \int_0^\infty \mathrm{e}^{-(X_{s+t}-X_t)} \mathrm{d}s$, then

$$\begin{aligned}
\sigma_{t+}^2 &= \beta \int_0^t \mathrm{e}^{X_{s-}-X_t} \mathrm{d}s + \beta \int_0^\infty \mathrm{e}^{-X_{s+t}} \mathrm{d}s \\
&= \beta \int_0^t \mathrm{e}^{X_{(t-u)-}-X_t} \mathrm{d}u + \beta \int_0^\infty \mathrm{e}^{-X_{s+t}} \mathrm{d}s
\end{aligned}$$

where in the first integral we set $s = t - u$. By the time-reversal property $X_{(t-u)-} - X_t \overset{\mathrm{d}}{=} -X_u$ for $0 \leq u \leq t$. Hence

$$\sigma_{t+}^2 \overset{\mathrm{d}}{=} \beta \int_0^t \mathrm{e}^{-X_s} \mathrm{d}s + \beta \int_t^\infty \mathrm{e}^{-X_s} \mathrm{d}s = \beta \int_0^\infty \mathrm{e}^{-X_s} \mathrm{d}s \overset{\mathrm{d}}{=} \sigma_0^2.$$

As σ_t has no fixed point of discontinuity $\sigma_t^2 = \sigma_{t+}^2$ almost surely. Then $\sigma_t^2 \overset{\mathrm{d}}{=} \sigma_0^2$ for $t > 0$. □

Remark 29.

$$\begin{aligned}
\mathbb{E}(\sigma_t^2 | \mathcal{F}_y) &= \mathbb{E}(\sigma_y^2 A_{y,t} | \mathcal{F}_y) + \mathbb{E}(B_{y,t} | \mathcal{F}_y) \\
&= \sigma_y^2 \mathbb{E}(A_{y,t}) + \mathbb{E}(B_{y,t}).
\end{aligned}$$

Theorem 2.2.4. $(\sigma_t^2)_{t\geq 0}$ *is also time-homogeneous.*

Proof. Let $\mathbb{D}[0,\infty)$ be the Skorohod space of càdlàg functions on $[0,\infty)$ and define the function

$$g_{y,t} : \mathbb{D}[0,\infty) \to \mathbb{R}^2 \quad x \mapsto (\mathrm{e}^{-(x_{t-}-x_{y-})}, \beta \int_y^t \mathrm{e}^{-(x_{t-}-x_s)} \mathrm{d}s).$$

X is a Lévy process so $(X_s)_{s\geq 0} \overset{d}{=} (X_{s+h} - X_h)_{s\geq 0}$ for every positive h. Furthermore $(A_{y,t}, B_{y,t}) = g_{y,t}((X_s)_{s\geq 0})$ and $(A_{y+h,t+h}, B_{y+h,t+h}) = g_{y,t}((X_{s+h} - X_h)_{s\geq 0}) \overset{d}{=} g_{y,t}((X_s)_{s\geq 0})$, then the joint distribution of $(A_{y,t}, B_{y,t})$ depends only on $t - y$ and by independence of σ_y^2 and $(A_{y,t}, B_{y,t})$ the transition functions are time homogeneous. □

Corollary 2.2.5. *If* $(\sigma_t^2)_{t\geq 0}$ *is the stationary version of the process with* $\sigma_0^2 \overset{\mathrm{d}}{=} \sigma_\infty^2$*, then* $(G_t)_{t\geq 0}$ *has stationary increments.*

Proof. For $t \geq 0$ $G_t = \int_0^t \sigma_s \mathrm{d}L_s$. We can write G_t in the following way too

$$G_t = G_y + \int_{y^+}^t \sigma_s \mathrm{d}L_s, \quad y \in [0,t).$$

Consequently

$$\begin{aligned} G_t - G_y &= \int_{y^+}^t \sigma_s \mathrm{d}L_s = \sigma_0 \int_{y^+}^t \mathrm{d}L_s = \sigma_0(L_t - L_{y^+}) = \\ &= \sigma_0(L_t - L_y) \overset{\mathrm{d}}{=} \sigma_0 L_{t-y} = \int_0^{t-y} \sigma_0 \mathrm{d}L_s = \int_0^{t-y} \sigma_s \mathrm{d}L_s = G_{t-y}. \end{aligned}$$

□

Since the integrand $(\sigma_s)_{y<s\leq t}$ depends on the past until y only through σ_y and the integrator L_s is independent of this past the following result holds.

Corollary 2.2.6. *The bivariate process* $(\sigma_t, G_t)_{t\geq 0}$ *is markovian.*

Remark 30. We have for $t \geq 0$ $X_t = -t\log\delta - \sum_{0<s\leq t} \log\left(1 + \frac{\lambda}{\delta}(\Delta L_s)^2\right)$.
If $\lambda = 0$ $X_t = -t\log\delta$ and

$$\begin{aligned} \sigma_t^2 &= \left(\beta \int_0^t \delta^{-s} \mathrm{d}s + \sigma_0^2\right)\delta^t = \left(\frac{\beta}{\log\delta}\int_1^{\delta^t} x^{-2}\mathrm{d}x + \sigma_0^2\right)\delta^t = \\ &= -\frac{\beta}{\log\delta}(1-\delta^t) + \sigma_0^2\delta^t. \end{aligned}$$

For the discrete time GARCH model if $\lambda = 0$

$$\begin{aligned} \sigma_n^2 &= \beta\sum_{i=0}^{n-1}\delta^{n-i-1} + \sigma_0^2\delta^n = \beta\delta^{n-1}\frac{1-\delta^{-n}}{1-\delta^{-1}} + \sigma_0^2\delta^n = \\ &= \beta\delta^n\frac{1-\delta^{-n}}{\delta-1} + \sigma_0^2\delta^n = \beta\frac{1-\delta^n}{1-\delta} + \sigma_0^2\delta^n, \quad n \in \mathbb{N} \end{aligned}$$

which proves the analogy between these two models. We also notice that only $\delta > 0$ is allowed so far. This continuous time model does not contain a COARCH submodel. If we want that $\delta = 0$ we have to go back to

$$\sum_{j=0}^{n-1}\log(\delta + \lambda\epsilon_j^2) = \sum_{j=0}^{n-1}\log(\lambda\epsilon_j^2) = n\log\lambda + \sum_{j=0}^{n-1}\log\epsilon_j^2$$

in order to define

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

which is a Lévy process only if L is a compound Poisson process.

Let introduce the following lemma concerning the relationship between the Laplace transform of X_t and the stability of the volatility process.

Lemma 2.2.7. *Keep* $c > 0$

(a) Let $\lambda > 0$*, then the Laplace transform* $\mathbb{E}(\mathrm{e}^{-cX_t})$ *of* X_t *at* c *is finite for some* $t > 0$*, or equivalently for every* $t > 0$*, if and only if* $\mathbb{E}(L_1^{2c}) < \infty$*.*

(b) If $\mathbb{E}(\mathrm{e}^{-cX_1}) < \infty$*, then* $|\Psi(c)| < \infty$*, where* $\Psi(c) = \log \mathbb{E}(\mathrm{e}^{-cX_1})$ *and* $\mathbb{E}(\mathrm{e}^{-cX_t}) = \mathrm{e}^{t\Psi(c)}$*, with*

$$\Psi(c) = c \log \delta + \int_{\mathbb{R}} \left[\left(1 + \frac{\lambda}{\delta} y^2\right)^c - 1 \right] \nu_L(\mathrm{d}y).$$

(c) If $\mathbb{E}(L_1^2) < \infty$ *and* $\Psi(1) < 0$*, then* $\sigma_t^2 \xrightarrow{\mathrm{d}} \sigma_\infty^2$*, where* σ_∞^2 *is a finite random variable.*

(d) If $\Psi(c) < 0$ *for some* $c > 0$*, then* $\Psi(d) < 0$ *for* $d \in (0, c)$*.*

Proof. (a) The Laplace transform $\mathbb{E}(\mathrm{e}^{-cX_t})$ is finite for some, and hence all, $t \geq 0$ if and only if

$$\begin{aligned}\int_{|x|>1} \mathrm{e}^{-cx} \nu_X(\mathrm{d}x) &= \int_{|y|>\sqrt{(\mathrm{e}-1)\delta/\lambda}} \left(1 + \frac{\lambda}{\delta} y^2\right)^c \nu_L(\mathrm{d}y) \\ &= \int_{(-\infty,-\sqrt{(\mathrm{e}-1)\delta/\lambda})} \left(1 + \frac{\lambda}{\delta} y^2\right)^c \nu_L(\mathrm{d}y) + \\ &\quad + \int_{(\sqrt{(\mathrm{e}-1)\delta/\lambda},\infty)} \left(1 + \frac{\lambda}{\delta} y^2\right)^c \nu_L(\mathrm{d}y) < \infty.\end{aligned}$$

For $y \to \pm\infty$

$$\left(1 + \frac{\lambda}{\delta} y^2\right)^c \sim \left(\frac{\lambda}{\delta}\right)^c y^{2c},$$

then

$$A := \left(\frac{\lambda}{\delta}\right)^c \int_{(-\infty,-\sqrt{(\mathrm{e}-1)\delta/\lambda})} y^{2c} \nu_L(\mathrm{d}y) < \infty.$$

In effect since $\mathbb{E}(L_1^{2c}) < \infty \Leftrightarrow \int_{|y|>1} y^{2c} \nu_L(\mathrm{d}y) < \infty$, if $-\sqrt{(\mathrm{e}-1)\delta/\lambda} < -1$, then $A \leq \left(\frac{\lambda}{\delta}\right)^c \int_{y<-1} y^{2c} \nu_L(\mathrm{d}y) < \infty$, instead, if $-\sqrt{(\mathrm{e}-1)\delta/\lambda} > -1$ $A = \int_{(-\infty,-1)} y^{2c} \nu_L(\mathrm{d}y) + \int_{[-1,-\sqrt{(\mathrm{e}-1)\delta/\lambda})} y^{2c} \nu_L(\mathrm{d}y)$ which is finite as well. Analogously for

$$B := \left(\frac{\lambda^c}{\delta}\right) \int_{(\sqrt{(\mathrm{e}-1)\delta/\lambda},\infty)} y^{2c} \nu_L(\mathrm{d}y),$$

if $\sqrt{(\mathrm{e}-1)\delta/\lambda} > 1$ $B \leq \int_{y>1} y^{2c} \nu_L(\mathrm{d}y) < \infty$ and if $\sqrt{(\mathrm{e}-1)\delta/\lambda} < 1$ $B = \int_{(\sqrt{(\mathrm{e}-1)\delta/\lambda},1]} y^{2c} \nu_L(\mathrm{d}y) + \int_{(1,\infty)} y^{2c} \nu_L(\mathrm{d}y)$ is still finite. Hence

$$\int_{|y|>\sqrt{(\mathrm{e}-1)\delta/\lambda}} \left(1 + \frac{\lambda}{\delta} y^2\right)^c \nu_L(\mathrm{d}y) < \infty.$$

(b) $\Psi(c) = \log \mathbb{E}(\mathrm{e}^{-cX_1}) < \infty$ as $\mathbb{E}(\mathrm{e}^{-cX_1}) < \infty$ (the Laplace transform of L_t never vanishes if L is a Lévy process). By the expression of the characteristic function of X_t

$$\Psi(c) = c \log \delta + \int_{\mathbb{R}} \left[\left(1 + \frac{\lambda}{\delta} y^2 \right)^c - 1 \right] \nu_L(\mathrm{d}y). \tag{2.11}$$

(c) $\mathbb{E}(L_1^2) < \infty$ and $\Psi(1) < 0$, then

$$\int_{\mathbb{R}} \frac{\lambda}{\delta} y^2 \nu_L(\mathrm{d}y) < -\log \delta.$$

Since

$$\int_{\mathbb{R}} \log \left(1 + \frac{\lambda}{\delta} y^2 \right) \nu_L(\mathrm{d}y) < \frac{\lambda}{\delta} \int_{\mathbb{R}} y^2 \nu_L(\mathrm{d}y) < -\log \delta,$$

and by Theorem 2.2.2 σ_t^2 converges weakly to a finite random variable σ_∞^2 such that $\sigma_\infty^2 \stackrel{\mathrm{d}}{=} \beta \int_0^\infty \mathrm{e}^{-X_t} \mathrm{d}t$.

(d) $\Psi(c) < 0$ for some $c > 0$. $\mathbb{E}(L_1^{2c}) < \infty$ implies $\mathbb{E}(L_1^{2d}) < \infty$ for $0 < d \leq c$ so that $\mathbb{E}(\mathrm{e}^{-dX_t}) < \infty$. Then $\Psi(d)$ is definable for $0 < d \leq c$. $\Psi(d) < 0$ if and only if

$$\frac{1}{d} \int_{\mathbb{R}} \left[\left(1 + \frac{\lambda}{\delta} y^2 \right)^d - 1 \right] \nu_L(\mathrm{d}y) < -\log \delta.$$

The function $d \mapsto \frac{1}{d} \left[\left(1 + \frac{\lambda}{\delta} y^2 \right)^d - 1 \right]$ is increasing for fixed y, hence

$$\begin{aligned} \frac{1}{d} \int_{\mathbb{R}} \left[\left(1 + \frac{\lambda}{\delta} y^2 \right)^d - 1 \right] \nu_L(\mathrm{d}y) &\leq \frac{1}{c} \int_{\mathbb{R}} \left[\left(1 + \frac{\lambda}{\delta} y^2 \right)^c - 1 \right] \nu_L(\mathrm{d}y) \\ &< -\log \delta. \end{aligned}$$

□

Remark 31. (c) can be extended. If $\mathbb{E}(L_1^{2c}) < \infty$ and $\Psi(c) < 0$ for some $c > 0$, then (2.10) holds, and a stationary version of $(\sigma_t^2)_{t \geq 0}$ exists. All that follows from the fact that $\Psi(c) < 0$ is equivalent to

$$\frac{1}{c} \int_{\mathbb{R}} \left(\left(1 + \frac{\lambda}{\delta} y^2 \right)^c - 1 \right) \nu_L(\mathrm{d}y) < -\log \delta.$$

Since $\log(1 + (\lambda/\delta) y^2) < (1/c)((1 + (\lambda/\delta) y^2)^c - 1)$ for $y \neq 0$, this implies (2.10).

Once proved such important proporties we calculate the moments of the processes and we show some related results in order to show the tails heaviness of the COGARCH process.

Proposition 2.2.8. *Let $\lambda > 0$, $t > 0$ and $h \geq 0$.*

(a) $\mathbb{E}(\sigma_t^2) < \infty$ if and only if $\mathbb{E}(L_1^2) < \infty$ and $\mathbb{E}(\sigma_0^2) < \infty$. If this is so

$$\mathbb{E}(\sigma_t^2) = -\frac{\beta}{\Psi(1)} + \left(\mathbb{E}(\sigma_0^2) + \frac{\beta}{\Psi(1)} \right) \mathrm{e}^{t\Psi(1)}.$$

(If $\Psi(1) = 0$ the right hand side has to be interpreted as its limit as $\Psi(1) \to 0$, and in this case $\mathbb{E}(\sigma_t^2) = \beta t + \mathbb{E}(\sigma_0^2)$.)

(b) $\mathbb{E}(\sigma_t^4) < \infty$ *if and only if* $\mathbb{E}(L_1^4) < \infty$ *and* $\mathbb{E}(\sigma_0^4) < \infty$*. In this case*

$$\mathbb{E}(\sigma_t^4) = \frac{2\beta^2}{\Psi(1)\Psi(2)} + \frac{2\beta^2}{\Psi(2)-\Psi(1)}\left(\frac{\mathrm{e}^{t\Psi(2)}}{\Psi(2)} - \frac{\mathrm{e}^{t\Psi(1)}}{\Psi(1)}\right) + \\ + 2\beta\mathbb{E}(\sigma_0^2)\left(\frac{\mathrm{e}^{t\Psi(2)} - \mathrm{e}^{t\Psi(1)}}{\Psi(2)-\Psi(1)}\right) + \mathbb{E}(\sigma_0^4)\mathrm{e}^{t\Psi(2)}.$$

(Also here the right hand side has to be interpreted as its limit if some of the denominators are zero.) Moreover

$$\mathbb{C}\mathrm{ov}(\sigma_t^2, \sigma_{t+h}^2) = \mathbb{V}\mathrm{ar}(\sigma_t^2)\mathrm{e}^{h\Psi(1)}.$$

Proof. (a) Since $X_t = X_{t-}$ a.s. and by Fubini theorem

$$\begin{aligned}
\mathbb{E}(\sigma_t^2) &= \beta\mathbb{E}\left(\int_0^t \mathrm{e}^{X_s - X_t}\mathrm{d}s\right) + \mathbb{E}(\sigma_0^2\mathrm{e}^{-X_t}) \\
&= \beta\mathbb{E}\left(\int_0^t \mathrm{e}^{-X_{t-s}}\mathrm{d}s\right) + \mathbb{E}(\sigma_0^2)\mathbb{E}(\mathrm{e}^{-X_t}) \\
&= \beta\mathbb{E}\left(\int_0^t \mathrm{e}^{-X_u}\mathrm{d}u\right) + \mathbb{E}(\sigma_0^2)\mathrm{e}^{t\Psi(1)} \\
&= \beta\int_0^t \mathrm{e}^{u\Psi(1)}\mathrm{d}u + \mathbb{E}(\sigma_0^2)\mathrm{e}^{t\Psi(1)} \\
&= -\frac{\beta}{\Psi(1)} + \left(\mathbb{E}(\sigma_0^2) + \frac{\beta}{\Psi(1)}\right)\mathrm{e}^{t\Psi(1)}
\end{aligned}$$

which is finite if and only if $\mathbb{E}(\sigma_0^2)$ and $\Psi(1)$ are finite, i.e. if and only if $\mathbb{E}(\sigma_0^2)$ and $\mathbb{E}(L_1^2)$ are finite quantities. If $\Psi(1) = 0$

$$\lim_{\Psi(1)\to 0}\mathbb{E}(\sigma_t^2) = \lim_{\Psi(1)\to 0}\frac{\beta}{\Psi(1)}(\mathrm{e}^{t\Psi(1)} - 1) + \mathbb{E}(\sigma_0^2) = \beta t + \mathbb{E}(\sigma_0^2).$$

(b)

$$\begin{aligned}
\mathbb{E}(\sigma_t^4) &= \mathbb{E}\left[\beta^2\left(\int_0^t \mathrm{e}^{X_s - X_t}\mathrm{d}s\right)^2\right] + \mathbb{E}(\sigma_0^4)\mathbb{E}(\mathrm{e}^{-2X_t}) + \\
&\quad + 2\beta\mathbb{E}(\sigma_0^2)\mathbb{E}\left(\int_0^t \mathrm{e}^{X_s - 2X_t}\mathrm{d}s\right) \\
&:= \beta^2\mathbb{E}(I_1) + \mathbb{E}(\sigma_0^4)\mathrm{e}^{t\Psi(2)} + 2\beta\mathbb{E}(\sigma_0^2)\mathbb{E}(I_2).
\end{aligned}$$

By Bertoin and Yor theorem (see Theorem 2 in [8])

$$\begin{aligned}
I_1 &= \left(\int_0^t \mathrm{e}^{X_s - X_t}\mathrm{d}s\right)^2 \overset{\mathrm{d}}{=} \left(\int_0^t \mathrm{e}^{-X_s}\mathrm{d}s\right)^2 \\
&= \int_0^t\int_0^t \mathrm{e}^{-X_s}\mathrm{e}^{-X_u}\mathrm{d}u\mathrm{d}s \\
&= \int_0^t\int_0^t \mathrm{e}^{-(X_s - X_u)}\mathrm{e}^{-2X_u}\mathrm{d}u\mathrm{d}s \\
&= 2\int_0^t\int_0^s \mathrm{e}^{-(X_s - X_u)}\mathrm{e}^{-2X_u}\mathrm{d}u\mathrm{d}s,
\end{aligned}$$

then

$$
\begin{aligned}
\mathbb{E}(I_1) &= 2\int_0^t \int_0^s \mathbb{E}(\mathrm{e}^{-(X_s-X_u)})\mathbb{E}(\mathrm{e}^{-2X_u})\mathrm{d}u\mathrm{d}s \\
&= 2\int_0^t \mathrm{e}^{s\Psi(1)} \left(\int_0^s \mathrm{e}^{u(\Psi(2)-\Psi(1))}\mathrm{d}u\right)\mathrm{d}s \\
&= \frac{2}{\Psi(1)\Psi(2)} + \frac{2}{\Psi(2)-\Psi(1)}\left(\frac{\mathrm{e}^{t\Psi(2)}}{\Psi(2)} - \frac{\mathrm{e}^{t\Psi(1)}}{\Psi(1)}\right).
\end{aligned}
$$

$$
\mathbb{E}(I_2) = \mathbb{E}\left(\int_0^t \mathrm{e}^{X_s-2X_t}\mathrm{d}s\right) = \int_0^t \mathrm{e}^{(t-s)\Psi(2)}\mathrm{e}^{s\Psi(1)}\mathrm{d}s = \frac{\mathrm{e}^{t\Psi(2)}-\mathrm{e}^{t\Psi(1)}}{\Psi(2)-\Psi(1)},
$$

hence

$$
\begin{aligned}
\mathbb{E}(\sigma_t^4) = &\frac{2\beta^2}{\Psi(1)\Psi(2)} + \frac{2\beta^2}{\Psi(2)-\Psi(1)}\left(\frac{\mathrm{e}^{t\Psi(2)}}{\Psi(2)} - \frac{\mathrm{e}^{t\Psi(1)}}{\Psi(1)}\right) + \\
&+ 2\beta\mathbb{E}(\sigma_0^2)\left(\frac{\mathrm{e}^{t\Psi(2)}-\mathrm{e}^{t\Psi(1)}}{\Psi(2)-\Psi(1)}\right) + \mathbb{E}(\sigma_0^4)\mathrm{e}^{t\Psi(2)}
\end{aligned}
$$

and we get that it is finite if and only if are finite $\Psi(2)$ and $\mathbb{E}(\sigma_0^4)$, i.e. if $\mathbb{E}(L_1^4)$ and $\mathbb{E}(\sigma_0^4)$ are finite. We calculate now the mixed moment $\mathbb{E}(\sigma_t^2\sigma_{t+h}^2)$ in order to calculate the autocovariance function. We know that

$$
\mathbb{E}(\sigma_{t+h}^2|\mathcal{F}_t) = \sigma_t^2\mathbb{E}(\mathrm{e}^{-(X_{t+h}-X_t)}) + \beta\int_t^{t+h}\mathbb{E}(\mathrm{e}^{-X_{t+h-s}})\mathrm{d}s,
$$

so

$$
\begin{aligned}
\mathbb{E}(\sigma_{t+h}^2|\mathcal{F}_t) &= \sigma_t^2\mathrm{e}^{h\Psi(1)} + \beta\int_t^{t+h}\mathrm{e}^{(t+h-s)\Psi(1)}\mathrm{d}s \\
&= \sigma_t^2\mathrm{e}^{h\Psi(1)} + \beta\int_0^h \mathrm{e}^{u\Psi(1)}\mathrm{d}s \\
&= \sigma_t^2\mathrm{e}^{h\Psi(1)} + \beta\frac{\mathrm{e}^{h\Psi(1)}-1}{\Psi(1)} \\
&= (\sigma_t^2 - \mathbb{E}(\sigma_0^2))\mathrm{e}^{h\Psi(1)} + \mathbb{E}(\sigma_h^2).
\end{aligned}
$$

Consequently

$$
\begin{aligned}
\mathbb{E}(\sigma_t^2\sigma_{t+h}^2) &= \mathbb{E}[\mathbb{E}(\sigma_t^2\sigma_{t+h}^2|\mathcal{F}_t)] \\
&= \mathbb{E}[\sigma_t^2\mathbb{E}(\sigma_{t+h}^2|\mathcal{F}_t)] \\
&= (\mathbb{E}(\sigma_t^4) - \mathbb{E}(\sigma_t^2)\mathbb{E}(\sigma_0^2))\mathrm{e}^{h\Psi(1)} + \mathbb{E}(\sigma_t^2)\mathbb{E}(\sigma_h^2).
\end{aligned}
\tag{2.12}
$$

As

$$
\begin{aligned}
\mathbb{E}(\sigma_t^2)\mathbb{E}(\sigma_h^2) - \mathbb{E}(\sigma_t^2)\mathbb{E}(\sigma_{t+h}^2) &= \mathbb{E}(\sigma_t^2)(\mathbb{E}(\sigma_h^2) - \mathbb{E}(\sigma_{t+h}^2)) \\
&= \mathbb{E}(\sigma_t^2)\left[\left(\mathbb{E}(\sigma_0^2) + \frac{\beta}{\Psi(1)}\right)\mathrm{e}^{h\Psi(1)}(1-\mathrm{e}^{t\Psi(1)})\right] \\
&= \mathrm{e}^{h\Psi(1)}(\mathbb{E}(\sigma_t^2)\mathbb{E}(\sigma_0^2) - \mathbb{E}^2(\sigma_t^2)),
\end{aligned}
$$

therefore

$$\mathbb{C}\mathrm{ov}\left(\sigma_t^2, \sigma_{t+h}^2\right) = (\mathbb{E}(\sigma_t^4) - \mathbb{E}^2(\sigma_t^2))\mathrm{e}^{h\Psi(1)} = \mathbb{V}\mathrm{ar}(\sigma_t^2)\mathrm{e}^{h\Psi(1)}.$$

□

Proposition 2.2.9. *Let $\lambda > 0$. Then the k-th moment of σ_∞^2 (and of σ_t^2 if the volatility process is strictly stationary) is finite if and only if $\mathbb{E}(L_1^{2k}) < \infty$ and $\Psi(k) < 0$ for $k \in \mathbb{N}$ and we have*

$$\mathbb{E}(\sigma_\infty^2) = k!\beta^k \prod_{l=1}^{k} \frac{1}{-\Psi(l)}.$$

Proof.

$$\begin{aligned}
\mathbb{E}(\sigma_\infty^{2k}) &= \beta^k \mathbb{E}\left[\left(\int_0^\infty \mathrm{e}^{-X_t}\mathrm{d}t\right)^k\right] \\
&= \beta^k \mathbb{E}\left[\int_0^\infty \cdots \int_0^\infty \mathrm{e}^{-X_{t_1}} \cdots \mathrm{e}^{-X_{t_k}} \mathrm{d}t_k \cdots \mathrm{d}t_1\right] \\
&= k!\beta^k \mathbb{E}\Big[\int_0^\infty \int_0^{t_1} \cdots \int_0^{t_{k-1}} \mathrm{e}^{-(X_{t_1}-X_{t_2})} \cdots \mathrm{e}^{-(k-1)(X_{t_{k-1}}-X_{t_k})} \cdot \\
&\quad \cdot \mathrm{e}^{-kX_{t_k}} \mathrm{d}t_k \cdots \mathrm{d}t_1\Big] \\
&= k!\beta^k \int_0^\infty \int_0^{t_1} \cdots \int_0^{t_{k-1}} \mathrm{e}^{t_1\Psi(1)} \mathrm{e}^{t_2\Psi(2)-\Psi(1)} \cdots \mathrm{e}^{t_k(\Psi(k)-\Psi(k-1))} \mathrm{d}t_k \cdots \mathrm{d}t_1 \\
&= k!\beta^k \prod_{l=1}^{k} \frac{1}{-\Psi(l)}
\end{aligned}$$

provided that $\Psi(1), \cdots \Psi(k)$ are all negative and defined. If $j \in \{1, \cdots, k\}$ is the first index for which $\Psi(j) \geq 0$ or $\mathbb{E}(\mathrm{e}^{-jX_1}) = \infty$, then $\mathbb{E}(\sigma_\infty^{2k}) = \infty$. $\mathbb{E}(\sigma_\infty^{2k}) < \infty$ if and only if $\Psi(k)$ is defined and negative, namely if and only if $\Psi(k) < 0$ (which implies $\Psi(j) < 0$ for $j < k$) and $\mathbb{E}(L_1^{2k}) < \infty$ (that implies $\mathbb{E}(\mathrm{e}^{-kX_t}) = \mathrm{e}^{t\Psi(k)} < \infty$). □

Corollary 2.2.10. *If $(\sigma_t^2)_{t\geq 0}$ is strictly stationary with $\sigma_0^2 \overset{\mathrm{d}}{=} \sigma_\infty^2$, then*

$$\mathbb{E}(\sigma_\infty^2) = -\frac{\beta}{\Psi(1)}$$

$$\mathbb{E}(\sigma_\infty^4) = \frac{2\beta^2}{\Psi(1)\Psi(2)}$$

$$\mathbb{C}\mathrm{ov}\left(\sigma_t^2, \sigma_{t+h}^2\right) = \beta^2 \left(\frac{2}{\Psi(1)\Psi(2)} - \frac{1}{\Psi^2(1)}\right)\mathrm{e}^{h\Psi(1)}$$

provided that $\mathbb{E}(L_1^{2k}) < \infty$ and $\Psi(k) < 0$ with $k = 1$ for $\mathbb{E}(\sigma_\infty^2)$ and $k = 2$ for $\mathbb{E}(\sigma_\infty^4)$ and $\mathbb{C}\mathrm{ov}\left(\sigma_t^2, \sigma_{t+h}^2\right)$.

Proof. From Propositions 2.2.8 and 2.2.9. □

Theorem 2.2.11. *Let $k \in \mathbb{N}$, $\delta \in (0,1)$ and $\lambda \geq 0$. Then σ_∞^2 exists and has finite k-th moment if and only if*

$$\frac{1}{k}\int_{\mathbb{R}}\left[\left(1+\frac{\lambda}{\delta}y^2\right)^k - 1\right]\nu_L(\mathrm{d}y) < -\log\delta.$$

Proof.

$$\frac{1}{k}\int_{\mathbb{R}}\left[\left(1+\frac{\lambda}{\delta}y^2\right)^k - 1\right]\nu_L(\mathrm{d}y) < -\log\delta$$

if and only if $\Psi(k) < 0$ and $\mathbb{E}(L_1^{2k}) < \infty$. Obviously if $\mathbb{E}(L_1^{2k}) < \infty$ and if $\Psi(k) < 0$, then $\mathbb{E}(L_1^2) < \infty$ and $\Psi(1) < 0$ so that σ_t^2 converges weakly. And σ_∞^2 has finite k-th moment since $\mathbb{E}(L_1^{2k}) < \infty$ and $\Psi(k) < 0$. □

Proposition 2.2.12. *(a) For any Lévy process L with nonzero Lévy measure such that*

$$\int_{\mathbb{R}}\log(1+y^2)\nu_L(\mathrm{d}y) < \infty$$

there exist $\delta, \lambda \in (0,1)$ such that σ_∞^2 exists, but $\mathbb{E}(\sigma_\infty^2) = \infty$.

(b) $k \in \mathbb{N}$, for any Lévy process L such that $\mathbb{E}(L_1^{2k}) < \infty$ and $\forall\delta \in (0,1)$ there exists $\lambda_\delta > 0$ such that σ_∞^2 exists with $\mathbb{E}(\sigma_\infty^{2k}) < \infty$ for every (λ, δ) such that $0 \leq \lambda \leq \lambda_\delta$.

(c) $\lambda > 0$. For no Lévy process L with nonzero Lévy measure do the moments of all orders of σ_∞^2 exist.

Proof. (a) Set $\delta_0 := \exp(-\int_{\mathbb{R}}\log(1+y^2)\nu_L(\mathrm{d}y))$ and $\delta_1 := \exp(-\int_{\mathbb{R}} y^2\nu_L(\mathrm{d}y))$. Since

$$\int_{\mathbb{R}}\log(1+y^2)\nu_L(\mathrm{d}y) < \int_{\mathbb{R}} y^2\nu_L(\mathrm{d}y)$$

$0 \leq \delta_1 < \delta_0 < 1$. For any $\lambda = \delta \in (\delta_1, \delta_0)$

$$\int_{\mathbb{R}}\log\left(1+\frac{\lambda}{\delta}y^2\right)\nu_L(\mathrm{d}y) = \int_{\mathbb{R}}\log(1+y^2)\nu_L(\mathrm{d}y) < -\log\delta$$

since $\exp(-\int_{\mathbb{R}}\log(1+y^2)\nu_L(\mathrm{d}y)) > \delta$ if and only if $\int_{\mathbb{R}}\log(1+y^2)\nu_L(\mathrm{d}y) < -\log\delta$. Then σ_t^2 converges in distribution to a finite random variable σ_∞^2. But $\mathbb{E}(\sigma_\infty^2) = \infty$ because

$$\int_{\mathbb{R}}\frac{\lambda}{\delta}y^2\nu_L(\mathrm{d}y) = \int_{\mathbb{R}} y^2\nu_L(\mathrm{d}y) > -\log\delta$$

as $\exp(-\int_{\mathbb{R}} y^2\nu_L(\mathrm{d}y)) < \delta$ if and only if $\int_{\mathbb{R}} y^2\nu_L(\mathrm{d}y) > -\log\delta$.

(b) If $\mathbb{E}(L_1^{2k}) < \infty$, then $\Psi(k) < \infty$ for $\lambda > 0$. We have

$$\frac{1}{k}\int_{\mathbb{R}}\left[\left(1+\frac{\lambda}{\delta}y^2\right)^k - 1\right]\nu_L(\mathrm{d}y) < \infty$$

and the left hand side goes to 0 if $\lambda \to 0$. So, choosing λ sufficienlty small there exists σ_∞^2 such that $\mathbb{E}(\sigma_\infty^{2k})$ is finite.

(c) Let $\eta > 0$ and $q := \nu_L(\{y : |y| \geq \eta\}) = \int_{|y|\geq\eta} \nu_L(\mathrm{d}y) > 0$. For $k \in \mathbb{N}$

$$\int_{\mathbb{R}} \left[\left(1+\frac{\lambda}{\delta}y^2\right)^k - 1\right] \nu_L(\mathrm{d}y) \geq \int_{|y|\geq\eta} \left[\left(1+\frac{\lambda}{\delta}y^2\right)^k - 1\right] \nu_L(\mathrm{d}y)$$
$$\geq \left[\left(1+\frac{\lambda}{\delta}\eta^2\right)^k - 1\right] \int_{|y|\geq\eta} \nu_L(\mathrm{d}y) = \left[\left(1+\frac{\lambda}{\delta}\eta^2\right)^k - 1\right] q.$$

If all moments of σ^2_∞ existed by Theorem 2.2.11

$$\frac{1}{k}\int_{\mathbb{R}} \left[\left(1+\frac{\lambda}{\delta}y^2\right)^k - 1\right] \nu_L(\mathrm{d}y) < -\log\delta,$$

but we know that

$$\int_{\mathbb{R}} \left[\left(1+\frac{\lambda}{\delta}y^2\right)^k - 1\right] \nu_L(\mathrm{d}y) \geq \left[\left(1+\frac{\lambda}{\delta}\eta^2\right)^k - 1\right] q.$$

Then

$$-\frac{k\log\delta}{q} > \left(1+\frac{\lambda}{\delta}\eta^2\right)^k - 1,$$

contradiction. If k is equal to 1, then $-\log\delta/q > \lambda\eta^2/\delta \Leftrightarrow -\delta\log\delta - q\lambda\eta^2 > 0$, which is a contradiction for fixed $q, \lambda, \eta > 0$, because the left hand side can be negative for some $\delta \in (0,1)$.

$\square$

Remark 32. The volatility process never has moments of all orders. Therefore COGARCH process turns out to be heavy tailed.

We conclude investigating other pathwise properties of the volatility process.

Proposition 2.2.13. *The volatility σ_t satisfies for all $t \geq 0$*

$$\sigma_t^2 \geq \frac{\beta}{-\log\delta}(1-\delta^t).$$

If $\sigma^2_{t_0} \geq \frac{\beta}{-\log\delta}$ for some t_0, then $\sigma^2_t \geq \frac{\beta}{-\log\delta}$ for every $t \geq t_0$. If $\sigma^2_t \stackrel{\mathrm{d}}{=} \sigma^2_\infty$ is the stationary version, then

$$\sigma^2_\infty \geq \frac{\beta}{-\log\delta}. \tag{2.13}$$

Morevoer $\sigma^2_{t+} - \sigma^2_t = \frac{\lambda}{\delta}\sigma^2_t(\Delta L_t)^2$.

Proof. From Equation (2.5) for $0 \leq s < t$,

$$X_s - X_{t-} = (t-s)\log\delta + \sum_{s<u<t} \log\left(1+\frac{\lambda}{\delta}(\Delta L_u)^2\right) \geq (t-s)\log\delta.$$

In particular

$$\begin{aligned}\sigma_t^2 &= \beta \int_0^t \mathrm{e}^{X_s - X_{t-}} \mathrm{d}s + \sigma_0^2 \mathrm{e}^{-X_{t-}} \\ &\geq \beta \int_0^t \mathrm{e}^{(t-s)\log\delta} \mathrm{d}s = \frac{\beta}{-\log\delta}(1 - \mathrm{e}^{t\log\delta}).\end{aligned}$$

Then (2.13) follows as $t \to \infty$. Now let $t > t_0$ and suppose that $\sigma_{t_0}^2 \geq \frac{\beta}{-\log\delta}$. Since

$$\sigma_t^2 = \mathrm{e}^{(X_{t_0-} - X_{t-})}\sigma_{t_0}^2 + \beta \int_{t_0}^t \mathrm{e}^{X_s - X_{t-}} \mathrm{d}s$$

$$\begin{aligned}\sigma_t^2 &\geq \mathrm{e}^{(t-t_0)\log\delta}\sigma_{t_0}^2 + \beta \int_{t_0}^t \mathrm{e}^{(s-t_0)\log\delta} \mathrm{d}s \\ &\geq \mathrm{e}^{(t-t_0)\log\delta}\left(\frac{\beta}{-\log\delta}\right) + \left(\frac{\beta}{-\log\delta}\right)(1 - \mathrm{e}^{(t-t_0)\log\delta}) = \frac{\beta}{-\log\delta}.\end{aligned}$$

From (2.7) we obtain that $\sigma_{t+}^2 - \sigma_t^2 = \frac{\lambda}{\delta}\sigma_t^2(\Delta L_t)^2$. □

Remark 33. The stationary version of the volatility process is always bounded away from 0 once $t > 0$. Furthermore, if a volatility jump occurs, this jump is necessarily positive.

2.3 Moments of the COGARCH increments

Let introduce the increments of the process G corresponding to logarithmic asset returns over time periods of length r

$$G_t^{(r)} := G_{t+r} - G_t = \int_0^{t+r} \sigma_s \mathrm{d}L_s - \int_0^t \sigma_s \mathrm{d}L_s = \int_{t+}^{t+r} \sigma_s \mathrm{d}L_s$$

for $t \geq 0$ and $r > 0$. We have already proved that if $(\sigma_t^2)_{t\geq 0}$ is stationary, then $(G_t^{(r)})_{t\geq 0}$ is stationary too. We are calculating the moments and the autocovariance function of this new process and speaking of which the following proposition holds.

Proposition 2.3.1. *Suppose that L is a quadratic pure jump process (i.e. $C_L = 0$) with $\mathbb{E}(L_1^2) < \infty$ and $\mathbb{E}(L_1) = 0$, and that $\Psi(1) < 0$. Let $(\sigma_t^2)_{t\geq 0}$ be the stationary volatility process with $\sigma_0^2 \overset{\mathrm{d}}{=} \sigma_\infty^2$, then for $t \geq 0$, $h \geq r > 0$*

$$\mathbb{E}(G_t^{(r)}) = 0$$

$$\mathbb{E}((G_t^{(r)})^2) = -\frac{\beta r}{\Psi(1)}\mathbb{E}(L_1^2)$$

$$\mathbb{C}\mathrm{ov}\left(G_t^{(r)}, G_{t+h}^{(r)}\right) = 0.$$

Assume furthermore that $\mathbb{E}(L_1^4) < \infty$ and $\Psi(2) < 0$, then

$$\mathbb{C}\mathrm{ov}\left((G_t^{(r)})^2, (G_{t+h}^{(r)})^2\right) = \left(\frac{\mathrm{e}^{-r\Psi(1)} - 1}{-\Psi(1)}\right)\mathbb{E}(L_1^2)\mathrm{e}^{h\Psi(1)}\mathbb{C}\mathrm{ov}(G_r^2, \sigma_r^2).$$

If further $\lambda > 0$, $\mathbb{E}(L_1^8) < \infty$, $\Psi(4) < 0$, $\int_{|x|\le 1} |x|\nu_L(\mathrm{d}x) < \infty$ *and* $\int_{\mathbb{R}} x^3\nu_L(\mathrm{d}x) = 0$, *then* $\mathbb{C}\mathrm{ov}\left((G_t^{(r)})^2, (G_{t+h}^{(r)})^2\right) > 0$.

Proof. L is a quadratic pure jump Lévy process with zero mean, so $[L, L]_t = \sum_{0<s\le t}(\Delta L_s)^2$, $t \ge 0$ and moreover it is a martingale because $\mathbb{E}(L_t|\mathcal{F}_s^*) = \mathbb{E}(L_t - L_s|\mathcal{F}_s^*) + L_s = \mathbb{E}(L_{t-s}) + L_s = L_s$ for $0 < s < t$, where $(\mathcal{F}_t^*)_{t\ge 0}$ is the filtration generated by L. We want to calculate the expectation of $G_t^{(r)}$; we have that

$$\mathbb{E}(G_t^{(r)}) = \mathbb{E}(G_{t+r} - G_t) = \mathbb{E}(G_r) = \mathbb{E}\left(\int_0^r \sigma_s \mathrm{d}L_s\right) = \mathbb{E}(L_1)\int_0^r \mathbb{E}(\sigma_s)\mathrm{d}s = 0.$$

Integration by parts implies that

$$G_t^2 = G_0^2 + 2\int_{0+}^t G_{s-}\mathrm{d}G_s + [G, G]_t,$$

and we obtain

$$\begin{aligned}\mathbb{E}(G_r^2) &= \mathbb{E}\left(2\int_{0+}^r G_{s-}\sigma_s \mathrm{d}L_s\right) + \mathbb{E}([G, G]_r)\\ &= \mathbb{E}\left(\int_0^r \sigma_s^2 \mathrm{d}[L, L]_s\right).\end{aligned}$$

The first term vanishes and about the second one we see from the compensation formula (Theorem 1.2.22) that

$$\begin{aligned}\mathbb{E}\left(\sum_{0<s\le r}\sigma_s^2(\Delta L_s)^2\right) &= \mathbb{E}\left(\int_0^r\int_{\mathbb{R}} x^2\sigma_s^2 N(\mathrm{d}s, \mathrm{d}x)\right)\\ &= \mathbb{E}\left(\int_0^r\int_{\mathbb{R}} x^2\sigma_s^2 \nu_L(\mathrm{d}x)\mathrm{d}s\right)\\ &= \int_{\mathbb{R}} x^2\nu_L(\mathrm{d}x)\int_0^r \mathbb{E}(\sigma_s^2)\mathrm{d}s\\ &= \mathbb{E}(L_1^2)\mathbb{E}(\sigma_0^2)r = \frac{-\beta r}{\Psi(1)}\mathbb{E}(L_1^2).\end{aligned}$$

We have $\int_{\mathbb{R}} x^2\nu_L(\mathrm{d}x) = \mathbb{E}(L_1^2)$ because

$$\begin{aligned}\mathbb{E}(L_1^2) &= -\phi''_{L_1}(0)\\ &= -\left[\left(i\gamma_L + i\int_{\mathbb{R}}(x - x\mathbf{1}_{\{|x|\le 1\}})\nu_L(\mathrm{d}x)\right)^2 - \tau_L^2 + i^2\int_{\mathbb{R}} x^2\nu_L(\mathrm{d}x)\right]\\ &= -\left[\left(i\gamma_L + i\int_{|x|>1} x\nu_L(\mathrm{d}x)\right)^2 - \int_{\mathbb{R}} x^2\nu_L(\mathrm{d}x)\right],\end{aligned}$$

but as

$$\begin{aligned}\mathbb{E}(L_1) &= -\phi'_{L_1}(0)\\ &= -i\left[i\gamma_L + i\int_{\mathbb{R}}(x - x\mathbf{1}_{\{|x|\le 1\}})\nu_L(\mathrm{d}x)\right]\\ &= \gamma_L + \int_{|x|>1} x\nu_L(\mathrm{d}x) = 0\end{aligned}$$

$$\mathbb{E}(L_1^2) = -\left(-i\int_{|x|>1} x\nu_L(\mathrm{d}x) + i\int_{|x|>1} x\nu_L(\mathrm{d}x)\right)^2 + \int_{\mathbb{R}} x^2\nu_L(\mathrm{d}x).$$

And we finally see that $\mathbb{E}((G_t^{(r)})^2) = \mathbb{E}(G_r^2)$. In order to calculate the autocovariance

$$\begin{aligned}\mathbb{E}(G_t^{(r)}G_{t+h}^{(r)}) &= \mathbb{E}\left(\int_{t+}^{t+r}\sigma_s\mathrm{d}L_s\int_{(t+h)+}^{t+r+h}\sigma_s\mathrm{d}L_s\right)\\ &= \mathbb{E}\left(\int_0^{t+h+r}\sigma_s^2\mathbf{1}_{(t,t+r]}(s)\mathbf{1}_{(t+h,t+r+h]}(s)\mathrm{d}[L,L]_s\right) = 0\end{aligned}$$

if $h \geq r$. Consequently $\mathbb{C}\mathrm{ov}(G_t^{(r)}, G_{t+h}^{(r)}) = 0$ if $h \geq r > 0$.
Suppose now $\mathbb{E}(L_1^4) < \infty$ and $\Psi(2) < 0$ and let $\mathbb{E}_r(\cdot)$ be the conditional expectation given $\mathcal{F}_r$, the $\sigma-$algebra generted by $(\sigma_s^2)_{0\leq s\leq r}$. By the integration by parts and compensation formula

$$\begin{aligned}\mathbb{E}_r((G_h^{(r)})^2) &= \mathbb{E}\left(2\int_{h+}^{h+r} G_{s-}\mathrm{d}G_s + [G,G]_{h+r} - [G,G]_h\right)\\ &= \mathbb{E}_r\left(2\int_{h+}^{h+r} G_{s-}\sigma_s\mathrm{d}L_s\right) + \mathbb{E}_r\left(\int_h^{h+r}\sigma_s^2\mathrm{d}[L,L]_s\right)\\ &= \int_{\mathbb{R}} x^2\nu_L(\mathrm{d}x)\int_h^{h+r}\mathbb{E}_r(\sigma_s^2)\mathrm{d}s\\ &= \mathbb{E}(L_1^2)\int_h^{h+r}[(\sigma_r^2 - \mathbb{E}(\sigma_0^2))\mathrm{e}^{(s-r)\Psi(1)} + \mathbb{E}(\sigma_{s-r}^2)]\mathrm{d}s\\ &= \mathbb{E}(L_1^2)(\sigma_r^2 - \mathbb{E}(\sigma_0^2))\left(\frac{\mathrm{e}^{-r\Psi(1)}-1}{-\Psi(1)}\right)\mathrm{e}^{h\Psi(1)} + \mathbb{E}(L_1^2)\mathbb{E}(\sigma_0^2)r.\end{aligned}$$

Hence

$$\begin{aligned}\mathbb{E}((G_0^{(r)})^2(G_h^{(r)})^2) &= \mathbb{E}[\mathbb{E}_r((G_0^{(r)})^2(G_h^{(r)})^2)\\ &= \mathbb{E}[G_r^2\mathbb{E}_r((G_h^{(r)})^2)\\ &= \frac{(\mathrm{e}^{-r\Psi(1)}-1)}{-\Psi(1)}\mathrm{e}^{h\Psi(1)}\mathbb{E}(L_1^2)\mathbb{E}(G_r^2\sigma_r^2 - G_r^2\mathbb{E}(\sigma_0^2))+\\ &\quad + \mathbb{E}(G_r^2)\mathbb{E}(L_1^2)\mathbb{E}(\sigma_0^2)r.\end{aligned}$$

It follows that

$$\mathbb{C}\mathrm{ov}((G_0^{(r)})^2, (G_h^{(r)})^2) = \frac{(\mathrm{e}^{-r\Psi(1)}-1)}{-\Psi(1)}\mathrm{e}^{h\Psi(1)}\mathbb{E}(L_1^2)\mathbb{C}\mathrm{ov}(G_r^2, \sigma_r^2) + \mathbb{E}(G_r^2).$$

In order to prove that the covariance is positive we need to prove that $\mathbb{C}\mathrm{ov}(G_r^2, \sigma_r^2) > 0$. We assume $\mathbb{E}(L_1^8) < \infty$, $\Psi(4) < 0$, $\int_{|x|\leq 1}|x|\nu_L(\mathrm{d}x) < \infty$ and $\int_{\mathbb{R}} x^3\nu_L(\mathrm{d}x) = 0$. Integration by parts says that

$$G_t^2 = [G,G]_t + 2\int_{0+}^t G_{s-}\mathrm{d}G_s = \sum_{0<s\leq t}\sigma_s^2(\Delta L_s)^2 + 2\int_{0+}^t G_{s-}\sigma\mathrm{d}L_s,$$

from (2.7)

$$\frac{\lambda}{\delta}G_t^2 = \sigma_{t+}^2 - \beta t - \log\delta \int_0^t \sigma_s^2 \mathrm{d}s - \sigma_0^2 + 2\frac{\lambda}{\delta}\int_{0+}^t G_{s-}\sigma_s \mathrm{d}L_s$$

so that

$$\begin{aligned}\frac{\lambda}{\delta}\mathbb{E}(G_t^2\sigma_t^2) &= \mathbb{E}(\sigma_t^4) - \beta t\mathbb{E}(\sigma_t^2) - \log\delta\mathbb{E}\left(\sigma_t^2\int_0^t \sigma_s^2\mathrm{d}s\right) - \mathbb{E}(\sigma_t^2\sigma_0^2)+ \\ &+ 2\frac{\lambda}{\delta}\mathbb{E}\left(\sigma_t^2\int_{0+}^t G_{s-}\sigma_s\mathrm{d}L_s\right).\end{aligned}$$

$$\begin{aligned}\sigma_t^2\int_{0+}^t G_{s-}\sigma_s\mathrm{d}L_s &= \int_{0+}^t G_{s-}\sigma_s\left(\sigma_s^2\mathrm{e}^{X_{s-}-X_{t-}} + \beta\int_s^t \mathrm{e}^{X_u-X_{t-}}\mathrm{d}u\right)\mathrm{d}L_s \\ &= \mathrm{e}^{-X_{t-}}\int_{0+}^t G_{s-}\sigma_s^3\mathrm{e}^{X_{s-}}\mathrm{d}L_s+ \\ &+ \int_{0+}^t G_{s-}\sigma_s\left(\beta\int_s^t \mathrm{e}^{X_u-X_{t-}}\mathrm{d}u\right)\mathrm{d}L_s.\end{aligned} \tag{2.14}$$

Define $I_t := \int_{0+}^t G_{s-}\sigma_s^3\mathrm{e}^{X_{s-}}\mathrm{d}L_s$, then

$$\mathrm{e}^{-X_t}I_t = \underbrace{\int_{0+}^t \mathrm{e}^{-X_{s-}}dI_s}_{:=A_t} + \underbrace{\int_{0+}^t I_{s-}\mathrm{d}(\mathrm{e}^{-X_s})}_{:=B_t} + \underbrace{[\mathrm{e}^{-X_\cdot}, I]_t}_{:=C_t^*}.$$

$$\begin{aligned}\mathbb{E}(A_t) &= \mathbb{E}\left(\int_{0+}^t \mathrm{e}^{-X_{s-}}\mathrm{d}\left(\int_{0+}^s G_{u-}\sigma_u^3\mathrm{e}^{X_{u-}}\mathrm{d}L_u\right)\right) = \mathbb{E}\left(\int_{0+}^t G_{s-}\sigma_s^3\mathrm{d}L_s\right) \\ &= \mathbb{E}(L_1)\int_0^t \mathbb{E}(G_{s-}\sigma_s^3)\mathrm{d}s = 0.\end{aligned}$$

We observe that

$$\begin{aligned}\mathrm{d}(\mathrm{e}^{-X_t}) &= \mathrm{d}(\mathrm{e}^{-X_t}) - \Psi(1)\mathrm{e}^{-X_t}\mathrm{d}t + \Psi(1)\mathrm{e}^{-X_t}\mathrm{d}t \\ &= \mathrm{e}^{t\Psi(1)}(\mathrm{d}(\mathrm{e}^{-X_t})\mathrm{e}^{-t\Psi(1)} - \Psi(1)\mathrm{e}^{-X_t}\mathrm{e}^{-t\Psi(1)}\mathrm{d}t) + \Psi(1)\mathrm{e}^{-X_t}\mathrm{d}t \\ &= \mathrm{e}^{t\Psi(1)}\mathrm{d}(\mathrm{e}^{-X_t-t\Psi(1)} - 1) + \Psi(1)\mathrm{e}^{-X_t}\mathrm{d}t\end{aligned}$$

and about B_t we have

$$\begin{aligned}\int_{0+}^t I_{s-}\mathrm{d}(\mathrm{e}^{-X_s}) &= \int_{0+}^t I_{s-}(\mathrm{e}^{s\Psi(1)}\mathrm{d}(\mathrm{e}^{-X_s-s\Psi(1)} - 1) + \Psi(1)\mathrm{e}^{-X_s}\mathrm{d}s) \\ &= \int_{0+}^t I_{s-}\mathrm{e}^{s\Psi(1)}\mathrm{d}(\mathrm{e}^{-X_s-s\Psi(1)} - 1) + \Psi(1)\int_{0+}^t I_{s-}\mathrm{e}^{-X_s}\mathrm{d}s.\end{aligned}$$

$R_t := \mathrm{e}^{-X_t-t\Psi(1)} - 1$ is a martingale because $\mathbb{E}(|R_t|) \le \mathbb{E}(\mathrm{e}^{-X_t-t\Psi(1)}) + 1 = 2$ and

$$\begin{aligned}\mathbb{E}(R_t|\mathcal{F}_s^*) &= \mathbb{E}(\mathrm{e}^{-(X_t-X_s)} - (t-s)\Psi(1)\mathrm{e}^{-X_s-s\Psi(1)} - 1|\mathcal{F}_s^*) \\ &= \mathrm{e}^{-X_s-s\Psi(1)}\mathbb{E}(\mathrm{e}^{-(X_t-X_s)} - (t-s)\Psi(1)) - 1 \\ &= \mathrm{e}^{-X_s-s\Psi(1)} - 1 = R_s\end{aligned}$$

for $0 < s < t$. Moreover $\mathbb{E}(R_t) = 0$ and hence

$$\mathbb{E}(B_t) = \mathbb{E}\left(\int_{0+}^t I_{s-}\mathrm{e}^{s\Psi(1)}\mathrm{d}R_s\right) + \Psi(1)\mathbb{E}\left(\int_0^t I_{s-}\mathrm{e}^{-X_s}\mathrm{d}s\right)$$
$$= \Psi(1)\int_0^t \mathbb{E}(I_{s-}\mathrm{e}^{-X_s})\mathrm{d}s.$$

Analyzing C_t^* we see that

$$\Delta C_t^* = \Delta \mathrm{e}^{-X_t}\Delta I_t = \frac{\lambda}{\delta}G_t\sigma_t^3(\Delta L_t)^3$$

since $\Delta I_t = G_{t-}\sigma_t^3\mathrm{e}^{-X_{t-}}\Delta L_t$ and

$$\Delta \mathrm{e}^{-X_t} = \delta^t\left(\prod_{0<s\leq t}\left(1+\frac{\lambda}{\delta}(\Delta(L_s))^2\right) - \prod_{0<s<t}\left(1+\frac{\lambda}{\delta}(\Delta L_s)^2\right)\right)$$
$$= \frac{\lambda}{\delta}\mathrm{e}^{-X_{t-}}(\Delta L_t)^2.$$

By Theorem 1.2.41

$$C_t^* = \sum_{0<s\leq t}\Delta\mathrm{e}^{-X_s}\Delta I_s = \frac{\lambda}{\delta}\sum_{0<s\leq t}G_{s-}\sigma_s^3(\Delta L_s)^3 = \frac{\lambda}{\delta}\int_{0+}^t G_{s-}\sigma_s^3\mathrm{d}M_s$$

with $M_t := \sum_{0<s\leq t}(\Delta L_s)^3$ which is a martingale too, indeed

$$\mathbb{E}(|M_t|) = \mathbb{E}\left(\left|\sum_{0<s\leq t}(\Delta L_s)^3\right|\right) \leq \mathbb{E}\left(\sum_{0<s\leq t}|\Delta L_s|^3\right) < \infty$$

because $\int_{|x|\leq 1}|x|\nu_L(\mathrm{d}x) < \infty$ and

$$\mathbb{E}(M_t|\mathcal{F}_s^*) = \mathbb{E}(M_t - M_s|\mathcal{F}_s^*) + M_s = \mathbb{E}\left(\sum_{s<k\leq t}(\Delta L_s)^3\right) + M_s = M_s, \quad 0 < s < t.$$

Moreover from the compensation formula and hypotesis about $\int_{\mathbb{R}} x^3\nu_L(\mathrm{d}x)$

$$\mathbb{E}(M_t) = \mathbb{E}\left(\int_0^t\int_{\mathbb{R}} x^3\nu_L(\mathrm{d}x)\mathrm{d}s\right) = 0$$

and it follows

$$\mathbb{E}(C_t^*) = \frac{\lambda}{\delta}\mathbb{E}(M_1)\int_0^t \mathbb{E}(G_{s-}\sigma_s^3)\mathrm{d}s = 0.$$

Therefore $\mathbb{E}(\mathrm{e}^{-X_t}I_t) = \Psi(1)\int_0^t \mathbb{E}(\mathrm{e}^{-X_s}I_s)\mathrm{d}s$ and we have the ODE

$$\mathrm{d}(\mathbb{E}(\mathrm{e}^{-X_t}I_t)) = \Psi(1)\mathbb{E}(\mathrm{e}^{-X_t}I_t)\mathrm{d}t$$

whose solution is $\mathbb{E}(\mathrm{e}^{-X_t}I_t) = 0$. Analyze the other term of the (2.14)

$$
\begin{aligned}
&\int_{0+}^{t} G_{s-}\sigma_s \left(\beta \int_{s}^{t} \mathrm{e}^{X_u - X_{t-}} \mathrm{d}u\right) \mathrm{d}L_s = \\
&= \beta \left(\int_{0}^{t} \mathrm{e}^{X_u - X_{t-}} \mathrm{d}u\right) \left(\int_{0+}^{t} G_{s-}\sigma_s \mathrm{d}L_s\right) - \beta \int_{0+}^{t} G_{s-}\sigma_s \left(\int_{0+}^{s} \mathrm{e}^{X_u - X_{t-}} \mathrm{d}u\right) \mathrm{d}L_s \\
&= \beta \int_{0+}^{t} \left(\int_{0}^{s} \mathrm{e}^{X_u - X_{t-}} \mathrm{d}u\right) \mathrm{d}\left(\int_{0+}^{s} G_{u-}\sigma_u \mathrm{d}L_u\right) \\
&\quad + \beta \int_{0+}^{t} \left(\int_{0+}^{s} G_{u-}\sigma_u \mathrm{d}L_u\right) \mathrm{d}\left(\int_{0}^{s} \mathrm{e}^{X_u - X_{t-}} \mathrm{d}u\right) + \beta \tilde{C}_t + \\
&\quad - \beta \int_{0+}^{t} G_{s-}\sigma_s \left(\int_{0+}^{s} \mathrm{e}^{X_u - X_{t-}} \mathrm{d}u\right) \mathrm{d}L_s \\
&= \beta \int_{0+}^{t} \left(\int_{0}^{s} \mathrm{e}^{X_u - X_{t-}} \mathrm{d}u\right) G_{s-}\sigma_s \mathrm{d}L_s + \beta \int_{0+}^{t} \mathrm{e}^{X_s - X_{t-}} \left(\int_{0}^{s} G_{u-}\sigma_u \mathrm{d}L_u\right) \mathrm{d}s + \\
&\quad + \beta \tilde{C}_t - \beta \int_{0+}^{t} G_{s-}\sigma_s \left(\int_{0}^{s} \mathrm{e}^{X_u - X_{t-}} \mathrm{d}u\right) \mathrm{d}L_s \\
&= \beta \int_{0+}^{t} \mathrm{e}^{X_s - X_{t-}} \left(\int_{0}^{s} G_{u-}\sigma_u \mathrm{d}L_u\right) \mathrm{d}s + \beta \tilde{C}_t,
\end{aligned}
$$

with $\tilde{C}_t$ the quadratic covariation between $\int_0^t \mathrm{e}^{X_u - X_{t-}} \mathrm{d}u$ and $\int_{0+}^t G_{s-}\sigma_s \mathrm{d}L_s$.

$$
\begin{aligned}
\Delta \tilde{C}_t &= G_{t-}\sigma_t \Delta L_t \Delta \left(\int_{0}^{t} \mathrm{e}^{X_u - X_{t-}} \mathrm{d}u\right) \\
&= G_{t-}\sigma_t \Delta L_t \left(\mathrm{e}^{-X_t} \int_{0}^{t} \mathrm{e}^{X_u} \mathrm{d}u - \mathrm{e}^{-X_{t-}} \int_{0}^{t} \mathrm{e}^{X_u} \mathrm{d}u\right) \\
&= G_{t-}\sigma_t \Delta L_t \int_{0}^{t} \mathrm{e}^{X_u} \mathrm{d}u \frac{\lambda}{\delta} \mathrm{e}^{-X_{t-}} (\Delta L_t)^2 \\
&= \frac{\lambda}{\delta} \mathrm{e}^{-X_{t-}} \left(\int_{0}^{t} \mathrm{e}^{X_u} \mathrm{d}u\right) G_{t-}\sigma_t (\Delta L_t)^3
\end{aligned}
$$

and

$$
\begin{aligned}
\tilde{C}_t &= \sum_{0<s\le t} \Delta \left(\mathrm{e}^{-X_s} \int_{0}^{s} \mathrm{e}^{X_u} \mathrm{d}u\right) G_{s-}\sigma_s \Delta L_s \\
&= \frac{\lambda}{\delta} \sum_{0<s\le t} \mathrm{e}^{-X_{s-}} \left(\int_{0}^{s} \mathrm{e}^{X_u} \mathrm{d}u\right) G_{s-}\sigma_s (\Delta L_s)^3 \\
&= \frac{\lambda}{\delta} \int_{0+}^{t} \mathrm{e}^{-X_{s-}} \left(\int_{0}^{s} \mathrm{e}^{X_u} \mathrm{d}u\right) G_{s-}\sigma_s \mathrm{d}M_s
\end{aligned}
$$

so that

$$
\mathbb{E}(\tilde{C}_t) = \frac{\lambda}{\delta} \mathbb{E}(M_1) \int_{0+}^{t} \mathbb{E}\left(\mathrm{e}^{-X_{s-}} \left(\int_{0}^{s} \mathrm{e}^{X_u} \mathrm{d}u\right) G_{s-}\sigma_s\right) \mathrm{d}s = 0.
$$

So

$$\begin{aligned}
&\mathbb{E}\left(\beta\int_0^t\left(\int_{0+}^s G_{u-}\sigma_u \mathrm{d}L_u\right)\mathrm{e}^{-(X_{t-}-X_s)}\mathrm{d}s\right)\\
&\quad=\beta\int_0^t \mathbb{E}\left(\int_{0+}^s G_{u-}\sigma_u \mathrm{d}L_u \mathrm{e}^{-(X_t-X_s)}\right)\mathrm{d}s\\
&\quad=\beta\int_0^t \mathbb{E}\left(\int_{0+}^s G_{u-}\sigma_u \mathrm{d}L_u\right)\mathbb{E}\left(\mathrm{e}^{-(X_t-X_s)}\right)\mathrm{d}s=0.
\end{aligned}$$

Therefore

$$\frac{\lambda}{\delta}\mathbb{E}(G_t^2\sigma_t^2)=\mathbb{E}(\sigma_t^4)-\beta t\mathbb{E}(\sigma_t^2)-\log\delta\mathbb{E}\left(\sigma_t^2\int_0^t\sigma_s^2\mathrm{d}s\right)-\mathbb{E}(\sigma_t^2\sigma_0^2).$$

From (2.12) and stationarity of the volatility process

$$\mathbb{E}(\sigma_t^2\sigma_s^2)=\mathbb{V}\mathrm{ar}(\sigma_0^2)\mathrm{e}^{(t-s)\Psi(1)}+\mathbb{E}^2(\sigma_0^2)$$

hence

$$\begin{aligned}
\frac{\lambda}{\delta}\mathbb{E}(G_t^2\sigma_t^2)&=\mathbb{E}(\sigma_t^4)-\beta t\mathbb{E}(\sigma_t^2)-\log\delta\mathbb{E}\left(\int_0^t(\mathbb{V}\mathrm{ar}(\sigma_0^2)\mathrm{e}^{(t-s)\Psi(1)}+\mathbb{E}^2(\sigma_0^2)\mathrm{d}s\right)+\\
&\quad-\mathbb{V}\mathrm{ar}(\sigma_0^2)\mathrm{e}^{t\Psi(1)}-\mathbb{E}^2(\sigma_0^2)\\
&=\mathbb{V}\mathrm{ar}(\sigma_0^2)(1-\mathrm{e}^{t\Psi(1)})-\beta t\mathbb{E}(\sigma_0^2)-\log\delta\mathbb{V}\mathrm{ar}(\sigma_0^2)\frac{1-\mathrm{e}^{t\Psi(1)}}{-\Psi(1)}+\\
&\quad-\log\delta\mathbb{E}^2(\sigma_0^2)t.
\end{aligned}$$

From (2.11) $\frac{\lambda}{\delta}\mathbb{E}(L_1^2)=\Psi(1)-\log\delta$ and we get

$$\begin{aligned}
\frac{\lambda}{\delta}\mathbb{E}(G_t^2)\mathbb{E}(\sigma_t^2)&=\frac{-\beta t}{\Psi(1)}\mathbb{E}(L_1^2)\mathbb{E}(\sigma_0^2)\frac{\lambda}{\delta}\\
&=-\beta t\mathbb{E}(\sigma_0^2)-t\log\delta\mathbb{E}^2(\sigma_0^2).
\end{aligned}$$

We can conclude that

$$\begin{aligned}
\frac{\lambda}{\delta}\mathbb{E}(G_t^2,\sigma_t^2)&=\mathbb{V}\mathrm{ar}(\sigma_0^2)\left(1-\mathrm{e}^{t\Psi(1)}-\frac{\log\delta(1-\mathrm{e}^{t\Psi(1)})}{-\Psi(1)}\right)\\
&=\mathbb{V}\mathrm{ar}(\sigma_0^2)\left(\frac{(1-\mathrm{e}^{t\Psi(1)})(-\Psi(1)-\log\delta)}{-\Psi(1)}\right)>0.
\end{aligned}$$

□

Remark 34. Proposition 2.3.1 tells us that log-returns are uncorrelated, while the squared log-returns are correlated. This agrees with empirical results. In this model the autocorrelation function of the squared log-return decreases exponentially. Furthermore $\mathbb{V}\mathrm{ar}(G_t^{(r)})$ is linear in r.

Proposition 2.3.2. *Suppose that the Lévy process L has finite variance and zero mean, and that $\Psi(1)<0$. Let $(\sigma_t)_{t\geq 0}$ be the stationary volatility process. Then the process $((G_{ri}^{(r)})^2)_{i\in\mathbb{N}}$ has for every fixed $r>0$ the autocorrelation structure of an ARMA(1,1) process.*

Proof. Let $\gamma(h) := \mathbb{C}\mathrm{ov}((G_{ri}^{(r)})^2, (G_{r(i+h)}^{(r)})^2)$ be the autocovariance function, for $h \in \mathbb{N}_0$, and denote by $\rho(h) := \mathbb{C}orr((G_{ri}^{(r)})^2, (G_{r(i+h)}^{(r)})^2)$ the autocorrelation function of the discrete time process $((G_{ri}^{(r)})^2)_{i\in\mathbb{N}}$. Then

$$\frac{\rho(h)}{\rho(1)} = \frac{\gamma(h)}{\gamma(1)} = \mathrm{e}^{-(h-1)r|\Psi(1)|}, \quad h \in \mathbb{N}.$$

For $h = 1$ we have $\rho(1) = \gamma(1)/\mathbb{V}\mathrm{ar}(G_r^2)$. □

We now state a theorem telling us the stationary distribution σ_∞^2 is *self-decomposable.*

Theorem 2.3.3. *The stationary distribution σ_∞^2 is self-decomposable, i.e. such that*

$$\sigma_\infty^2 \stackrel{\mathrm{d}}{=} k\sigma_\infty^2 + Y,$$

where $k \in (0, 1)$ and Y is independent of σ_∞^2.

Proof. The auxiliary process X is spectrally negative and $X_t \to +\infty$ a.s. as $t \to \infty$. From this follows that the stopping time T_h, defined for $h > 0$ by

$$T_h := \inf\{t \geq 0 : X_t = h\},$$

is almost surely finite. Consider the σ-algebra generated by $(X_s)_{0\leq s\leq t}$, $\mathcal{F}_t$, and the stopping time σ-algebra $\mathcal{F}_{T_h}$. By strong Markov property $(X_{T_h+t} - X_{T_h})_{t\geq 0}$ is independent of $\mathcal{F}_{T_h}$ and has the same law as $(X_t)_{t\geq 0}$.

$$\sigma_\infty^2 \stackrel{\mathrm{d}}{=} \beta \int_0^\infty \mathrm{e}^{X_t}\mathrm{d}t = \beta \int_0^{T_h} \mathrm{e}^{-X_t}\mathrm{d}t + \beta \int_{T_h}^\infty \mathrm{e}^{-X_t}\mathrm{d}t =: A_h + B_h,$$

tells us that A_h is $\mathcal{F}_{T_h}$-measurable that

$$B_h = \beta \int_{T_h}^\infty \mathrm{e}^{-(X_t - X_{T_h})}\mathrm{e}^{-X_{T_h}}\mathrm{d}t = \mathrm{e}^{-h}\beta \int_{T_h}^\infty \mathrm{e}^{-(X_t - X_{T_h})}\mathrm{d}t$$

is independent of A_h and has the same law as $\mathrm{e}^{-h}\sigma_\infty^2$. Hence for $h > 0$

$$\sigma_\infty^2 \stackrel{\mathrm{d}}{=} A_h + \mathrm{e}^{-h}\sigma_\infty^2$$

with A_h and σ_∞^2 are independent. □

Under suitable conditions once can show that σ_∞^2 has Pareto like tails (see [28]). See also [9] for higher and joint moments.

Remark 35. We will often use a different parametrisation for the COGARCH model. Take $\delta = \mathrm{e}^{-\eta}$ and $\lambda = \varphi\mathrm{e}^{-\eta}$ so that the auxiliary process becomes

$$X_t = \eta t - \sum_{0<s\leq t} \log(1 + \varphi(\Delta L_t)^2)$$

and the volatility process is solution of the following differential equation

$$\mathrm{d}\sigma_{t+}^2 = (\beta - \eta\sigma_t^2)\mathrm{d}t + \varphi\sigma_t^2\mathrm{d}[L]_t^d.$$

The following plots show trajectories of G_t, $G_t^{(1)}$ and σ_t^2. The simulation was done by choosing parametrisation introduced in Remark 35 and a variance gamma process as driving Lévy process with $\mu = 0$ and $\tau = \sigma = 1$[1], so that $\mathbb{E}(L_1) = 0$ and $\mathbb{E}(L_1^2) = 1$. In this case for $\tau > 0$

$$\mathbb{E}(\mathrm{e}^{iuL_t}) = \left(1 + \frac{\tau u^2}{2}\right)^{-t/\tau}$$

and

$$\nu_L(\mathrm{d}x) = \frac{1}{\tau|x|} \exp\left(-\sqrt{2/\tau}|x|\right) \mathrm{d}x, \quad x \neq 0.$$

By Equation (2.11)

$$\Psi(1) = -\eta + \varphi \int_{\mathbb{R}} x^2 \nu_L(\mathrm{d}x) = \varphi - \eta$$

and

$$\begin{aligned}
\Psi(2) &= -2\eta + \int_{\mathbb{R}} ((1 + \varphi x^2)^2 - 1) \frac{1}{\tau|x|} \mathrm{e}^{-\sqrt{2/\tau}|x|} \mathrm{d}x \\
&= -2\eta + \int_{\mathbb{R}} (2\varphi x^2 + \varphi^2 x^4) \frac{1}{\tau|x|} \mathrm{e}^{-\sqrt{2/\tau}|x|} \mathrm{d}x \\
&= -2\eta + 2\varphi + \frac{2\varphi^2}{\tau} \int_0^\infty x^3 \mathrm{e}^{-\sqrt{2/\tau}|x|} \mathrm{d}x = 2\varphi - 2\eta + 3\tau\varphi^2.
\end{aligned}$$

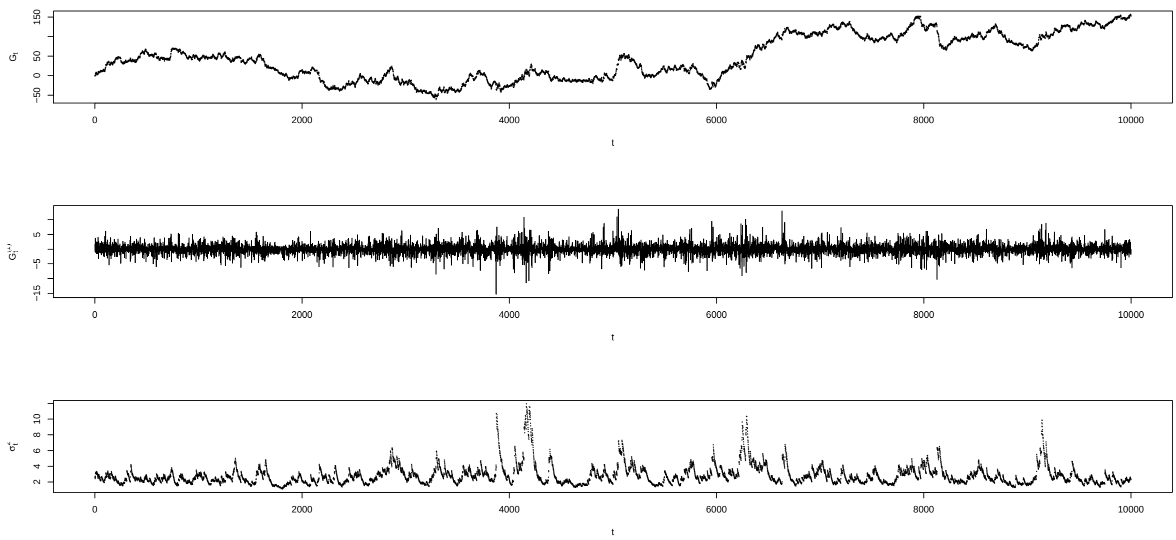

Figure 2.1: Simulation of a variance gamma driven COGARCH(1,1) process $(G_t)_{0 \leq t \leq 10000}$ with parameters $\beta = 0.04$, $\eta = 0.053$, $\varphi = 0.038$, $\tau = \sigma = 1$ and $\mu = 0$, log-return process $(G_t^{(1)})_{0 \leq t \leq 10000}$ and squared volatility process $(\sigma_t^2)_{0 \leq t \leq 10000}$.

[1] There shouldn't be confusion between the volatility process and the parameter of the variance gamma we called σ in the first chapter.

2.4 Asymmetric COGARCH(1,1) process

As is well known in literature, there is an asymmetric response of the volatility to positive and negative past returns. The so-called leverage effect refers to the observed tendency of the volatility to be negatively correlated with stock returns: volatility tends to increase in response to bad news and to fall in response to good news. It also been documented that the effect is asymmetric: declines in stock price are accompanied by larger increases in volatility than the decline in volatility which accompanies rising prices. Therefore, it is important to include the asymmetric term in financial time series models (see [39], [45] and [47]). Then, new discrete time models were introduced. Ding, Granger and Engle proposed in [13] an Asymmetric Power GARCH (APGARCH) model, which contains classical ARCH and GARCH processes, as well as the GJR-GARCH model (see [20]) and the Threshold GARCH (TGARCH) model developed by Zakoian [53]. We will discuss continuous time APGARCH presented in [31] and continuous time GJR-GARCH studied by Behme, Klüppelberg and Mayr recently in [5].

2.4.1 Discrete time APGARCH process

Definition 2.1. Let $(\epsilon_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables with $\mathbb{E}(\epsilon_n) = 0$ and $\mathbb{V}\mathrm{ar}(\epsilon_n) = 1$. The process $(Y_n)_{n\in\mathbb{N}}$ is called Asymmetric Power GARCH(p, q) if it satisfies the following form (cf. [39])

$$Y_n = \epsilon_n \sigma_n \tag{2.15}$$

$$\sigma_n^\tau = \beta + \sum_{i=1}^{q} \lambda_i h(Y_{n-i}) + \sum_{j=1}^{p} \delta_j \sigma_{n-j}^\tau, \tag{2.16}$$

with $h(x) := (|x| - \gamma x)^\tau$, $\beta > 0$, $\tau > 0$, $\lambda_i > 0$, $\delta_j > 0$ and $|\gamma_i| < 1$.

Remark 36. The function $h(x)$ is strictly positive for all $x \in \mathbb{R} \setminus \{0\}$ and $\tau > 0$ because $|x| > \gamma x$ if and only if $\gamma \in (-1, 1)$.

Remark 37. For $\tau = 2$ and $\gamma_i = 0$ for each $i = 1, \cdots, q$ $h(x) = x^2$ and we obtain the discrete time GARCH(p, q) model. In general for $\tau = 2$ we have the GJR-GARCH(p, q) (see [20]). In this case if $0 \le \gamma_i < 1$

$$\begin{aligned}
\sigma_n^2 &= \beta + \sum_{i=1}^{q} \lambda_i (|Y_{n-i}| - \gamma_i Y_{n-i})^2 + \sum_{j=1}^{p} \delta_j \sigma_{n-j}^2 \\
&= \beta + \sum_{i=1}^{q} \lambda_i (1-\gamma_i)^2 Y_{n-i}^2 + 4 \sum_{i=1}^{q} \lambda_i Y_{n-i}^2 \mathbf{1}_{\{Y_{n-i}<0\}} + \sum_{j=1}^{p} \delta_j \sigma_{n-j}^2 \\
&= \beta + \sum_{i=1}^{q} \lambda_i^* Y_{n-i}^2 + \sum_{i=1}^{q} \gamma_i^* Y_{n-i}^2 \mathbf{1}_{\{Y_{n-i}<0\}} + \sum_{j=1}^{p} \delta_j \sigma_{n-j}^2
\end{aligned}$$

with $\lambda_i^* = \lambda_i(1-\gamma_i)^2$ and $\gamma_i^* = 4\lambda_i\gamma_i$. If $-1 < \gamma_i < 0$, then

$$\begin{aligned}\sigma_n^2 &= \beta + \sum_{i=1}^{q} \lambda_i(1+\gamma_i)^2 Y_{n-i}^2 - 4\sum_{i=1}^{q} \lambda_i Y_{n-i}^2 \mathbf{1}_{\{Y_{n-i}>0\}} + \sum_{j=1}^{p} \delta_j \sigma_{n-j}^2 \\ &= \beta + \sum_{i=1}^{q} \lambda_i^* Y_{n-i}^2 + \sum_{i=1}^{q} \gamma_i^* Y_{n-i}^2 \mathbf{1}_{\{Y_{n-i}>0\}} + \sum_{j=1}^{p} \delta_j \sigma_{n-j}^2\end{aligned}$$

with $\lambda_i^* = \lambda_i(1+\gamma_i)^2$ and $\gamma_i^* = -4\lambda_i\gamma_i$.

2.4.2 Continuous time APGARCH process

Assume that $p = q = 1$, so that we have

$$Y_n = \epsilon_n \sigma_n, \quad \sigma_n^\tau = \beta + \lambda(|Y_{n-1}| - \gamma Y_{n-1})^\tau + \delta\sigma_{n-1}^\tau,$$

where $\beta > 0$, $\tau > 0$, $\lambda \geq 0$, $\delta \geq 0$ and $-1 < \gamma < 1$.
We will replace the innovations of the discrete time APGARCH model through the increments of a Lévy process. We observe that

$$\begin{aligned}\sigma_n^\tau &= \beta + \lambda(|Y_{n-1}| - \gamma Y_{n-1})^\tau + \delta\sigma_{n-1}^\tau \\ &= \beta + \lambda(|\epsilon_{n-1}\sigma_{n-1}| - \gamma \epsilon_{n-1}\sigma_{n-1})^\tau + \delta\sigma_{n-1}^\tau \\ &= \beta + \lambda(|\epsilon_{n-1}| - \gamma\epsilon_{n-1})^\tau \sigma_{n-1}^\tau + \delta\sigma_{n-1}^\tau \\ &= \beta + (\lambda h(\epsilon_{n-1}) + \delta)\sigma_{n-1}^\tau.\end{aligned} \tag{2.17}$$

Iteration of (2.17) gives us

$$\begin{aligned}\sigma_n^\tau &= \beta + (\lambda h(\epsilon_{n-1}) + \delta)\sigma_{n-1}^\tau \\ &= \beta + (\lambda h(\epsilon_{n-1}) + \delta)(\beta + (\lambda h(\epsilon_{n-2}) + \delta)\sigma_{n-2}^\tau) \\ &= \cdots \\ &= \beta \sum_{i=0}^{n-1} \prod_{j=i+1}^{n-1} (\delta + \lambda h(\epsilon_j)) + \sigma_0^\tau \prod_{j=0}^{n-1} (\delta + \lambda h(\epsilon_j)) \\ &\cong \beta \int_0^n \exp\left(\sum_{j=\lfloor s \rfloor +1}^{n-1} \log(\delta + \lambda h(\epsilon_j)) \right) \mathrm{d}s + \sigma_0^\tau \exp\left(\sum_{j=0}^{n-1} \log(\delta + \lambda h(\epsilon_j)) \right) \\ &= \beta \int_0^n \exp\left(\sum_{j=\lfloor s \rfloor +1}^{n-1} \left(\log\delta + \log\left(1 + \frac{\lambda}{\delta} h(\epsilon_j)\right)\right)\right) \mathrm{d}s + \\ &\quad + \sigma_0^\tau \exp\left(\sum_{j=0}^{n-1} \left(\log\delta + \log\left(1 + \frac{\lambda}{\delta} h(\epsilon_j)\right)\right)\right).\end{aligned}$$

We replaced the innovations ϵ_j by the increments of the Lévy process $L = (L_t)_{t\geq 0}$ with Lévy measure $\nu_L \neq 0$.
Define the auxiliary process $X = (X_t)_{t\geq 0}$ as follows

$$X_t := -t\log\delta - \sum_{0<s\leq t} \log\left(1 + \frac{\lambda}{\delta} h(\Delta L_s)\right), \quad t \geq 0,$$

where $\lambda > 0$, $0 < \delta < 1$, $|\gamma| < 1$, $\tau > 0$ and $h(x) = (|x| - \gamma x)^\tau$. With σ_0^τ a finite positive random variable independent of L define the càglàd volatility process analogously to (2.6) by

$$\sigma_t^\tau = \left(\beta \int_0^t \mathrm{e}^{X_s} \mathrm{d}s + \sigma_0^\tau\right) \mathrm{e}^{-X_{t-}}, \quad t \geq 0.$$

We now can define the integrated continuous time APGARCH(1,1) process $G = (G_t)_{t\geq 0}$ as the càdlàg process satisfying the following stochastic differential equation

$$\mathrm{d}G_t = \sigma_t \mathrm{d}L_t, \quad t \geq 0,$$

with $G_0 = 0$.
Analogously to the COGARCH(1,1) model G jumps at the same time as L does, with $\Delta G_t = \sigma_t \Delta L_t$.

2.4.3 Continuous time GJR-GARCH process

Continuous time APGARCH(1,1) model includes for $\tau = 2$ the continuous time GJR-GARCH(1,1) process analyzed in more details in [5]. For $p = q = 1$, $\beta > 0$, $\lambda > 0$, $\delta \in (0,1)$ and $\gamma \in (-1,1)$ equation (2.16) implies

$$\begin{aligned}
\sigma_n^2 &= \beta + \lambda(|Y_{n-1}| - \gamma Y_{n-1})^2 + \delta\sigma_{n-1}^2 \\
&= \beta + \lambda(|\epsilon_{n-1}\sigma_{n-1}| - \gamma\epsilon_{n-1}\sigma_{n-1})^2 + \delta\sigma_{n-1}^2 \\
&= \beta + \lambda\epsilon_{n-1}^2\sigma_{n-1}^2 - 2\lambda\gamma|\epsilon_{n-1}\sigma_{n-1}|\epsilon_{n-1}\sigma_{n-1} + \lambda\gamma^2\epsilon_{n-1}2\sigma_{n-1}^2 + \delta\sigma_{n-1}^2 \\
&= \beta + \lambda(1-\gamma)^2\epsilon_{n-1}^2\sigma_{n-1}^2\mathbf{1}_{\{\epsilon_{n-1}\geq 0\}} + \lambda(1+\gamma)^2\epsilon_{n-1}^2\sigma_{n-1}^2\mathbf{1}_{\{\epsilon_{n-1}<0\}} + \delta\sigma_{n-1}^2 \\
&= \beta + \lambda^*\epsilon_{n-1}^2\sigma_{n-1}^2\mathbf{1}_{\{\epsilon_{n-1}\geq 0\}} + \gamma^*\epsilon_{n-1}^2\sigma_{n-1}^2\mathbf{1}_{\{\epsilon_{n-1}<0\}} + \delta\sigma_{n-1}^2
\end{aligned}$$

where $\lambda^* := \lambda(1-\gamma)^2$ and $\gamma^* := \lambda(1+\gamma)^2$. Hence

$$\begin{aligned}
\sigma_n^2 &= \beta + \left(\delta + (\lambda^*\mathbf{1}_{\{\epsilon_{n-1}\geq 0\}} + \gamma^*\mathbf{1}_{\{\epsilon_{n-1}<0\}})\epsilon_{n-1}^2\right)\sigma_{n-1}^2 \\
&= \beta + \left(\delta + (\lambda^*\mathbf{1}_{\{\epsilon_{n-1}\geq 0\}} + \gamma^*\mathbf{1}_{\{\epsilon_{n-1}<0\}})\epsilon_{n-1}^2\right) \cdot \\
&\quad \cdot \left[\beta + \left(\delta + (\lambda^*\mathbf{1}_{\{\epsilon_{n-2}\geq 0\}} + \gamma^*\mathbf{1}_{\{\epsilon_{n-2}<0\}})\epsilon_{n-2}^2\right)\right]\sigma_{n-2}^2 \\
&= \cdots \\
&= \beta\sum_{j=0}^{n-1}\prod_{k=j+1}^{n-1}\left(\beta + (\lambda^*\mathbf{1}_{\{\epsilon_k\geq 0\}} + \gamma^*\mathbf{1}_{\{\epsilon_k<0\}})\right) + \sigma_0^2\prod_{k=0}^{n-1}\left(\beta + (\lambda^*\mathbf{1}_{\{\epsilon_k\geq 0\}} + \gamma^*\mathbf{1}_{\{\epsilon_k<0\}})\right) \\
&= \beta\int_0^n \exp\left(\sum_{k=\lfloor s\rfloor+1}^{n-1}\log(\delta + (\lambda^*\mathbf{1}_{\{\epsilon_k\geq 0\}} + \gamma^*\mathbf{1}_{\{\epsilon_k<0\}})\epsilon_k^2)\right)\mathrm{d}s + \\
&\quad + \sigma_0^2\exp\left(\sum_{k=0}^{n-1}\log(\delta + (\lambda^*\mathbf{1}_{\{\epsilon_k\geq 0\}} + \gamma^*\mathbf{1}_{\{\epsilon_k<0\}})\epsilon_k^2)\right) \\
&= \beta\int_0^n \exp\left(\sum_{k=\lfloor s\rfloor+1}^{n-1}\left(\log\delta + \log\left(1 + \frac{\lambda^*\mathbf{1}_{\{\epsilon_k\geq 0\}} + \gamma^*\mathbf{1}_{\{\epsilon_k<0\}}}{\delta}\epsilon_k^2\right)\right)\right)\mathrm{d}s + \\
&\quad + \sigma_0^2\exp\left(\sum_{k=0}^{n-1}\left(\log\delta + \log\left(1 + \frac{\lambda^*\mathbf{1}_{\{\epsilon_k\geq 0\}} + \gamma^*\mathbf{1}_{\{\epsilon_k<0\}}}{\delta}\epsilon_k^2\right)\right)\right).
\end{aligned}$$

Hence, if we define the process $X = (X_t)_{t\geq 0}$ as

$$X_t := -t\log\delta - \sum_{0<s\leq t}\log\left(1 + \frac{\lambda^*\mathbf{1}_{\{\Delta L_s\geq 0\}} + \gamma^*\mathbf{1}_{\{\Delta L_s<0\}}}{\delta}(\Delta L_s)^2\right)$$

one obtains the integrated continuous time GJR-GARCH(1,1) process $G = (G_t)_{t\geq 0}$ defined by

$$\mathrm{d}G_t = \sigma_t \mathrm{d}L_t,$$

with

$$\sigma_t^2 := \left(\beta\int_0^t \mathrm{e}^{X_s}\mathrm{d}s + \sigma_0^2\right)\mathrm{e}^{-X_{t-}}, \quad t\geq 0.$$

Remark 38. Choosing the parameters $\delta = \mathrm{e}^{-\eta}$, $\lambda = \varphi\mathrm{e}^{-\eta}$ it follows that $\lambda^* = \varphi\mathrm{e}^{-\eta}(1-\gamma)^2$ and $\gamma^* = \varphi\mathrm{e}^{-\eta}(1+\gamma)^2$ we can write

$$\begin{aligned} X_t &= \eta t - \sum_{0<s\leq t}\log\left(1 + \left[(1-\gamma)^2\mathbf{1}_{\{\Delta L_s\geq 0\}} + (1+\gamma)^2\mathbf{1}_{\{\Delta L_s<0\}}\right]\varphi(\Delta L_s)^2\right) \\ &= \eta t - \sum_{0<s\leq t}\log(1+\varphi h(\Delta L_s)). \end{aligned}$$

This kind of model has the same properties as the COGARCH(1,1) process (see [5]). We list them without proofs since one can prove these results analogously we did for the symmetric model.

Proposition 2.4.1. *Suppose that* $\mathbb{E}(|L_1|^\tau) < \infty$*. Then* X *is a spectrally negative Lévy process with bounded variation. Moreover the characteristic triplet is such that* $C_X = 0$,

$$\nu_X([0,\infty)) = 0,$$

$$\nu_X((-\infty,-x]) = \nu_L\left(\left\{y\in\mathbb{R} : h(y)\geq\sqrt{(\mathrm{e}^x-1)\delta/\lambda}\right\}\right), \quad x>0$$

and

$$\gamma_{X,0} := \gamma_X - \int_{[-1,1]} x\nu_X(\mathrm{d}x) = -\log\delta.$$

Proposition 2.4.2. *The process* $(\sigma_t^2)_{t\geq 0}$ *satisfies the following stochastic differential equation*

$$\mathrm{d}\sigma_{t+}^\tau = \beta\mathrm{d}t + \sigma_t^\tau\mathrm{e}^{X_{t-}}\mathrm{d}(\mathrm{e}^{-X_t}), \quad t>0$$

and we have

$$\sigma_t^\tau = \sigma_0^\tau + \beta t + \log\delta\int_0^t\sigma_s^\tau\mathrm{d}s + \frac{\lambda}{\delta}\sum_{0<s<t}\sigma_s^\tau h(\Delta L_s). \tag{2.18}$$

Remark 39. If we use the parametrisation with β, η, φ and γ we obtain

$$\mathrm{d}\sigma_{t+}^\tau = (\beta - \eta\sigma_t^\tau)\mathrm{d}t + \varphi\sigma_t^\tau\mathrm{d}\left(\sum_{0<s\leq t}h(\Delta L_s)\right).$$

Theorem 2.4.3. *Suppose that*

$$\int_{\mathbb{R}} \log\left(1+\frac{\lambda}{\delta}h(y)\right)\nu_L(\mathrm{d}y) < -\log\delta, \tag{2.19}$$

then the process $(\sigma_t^\tau)_{t\geq 0}$ converges in distribution as $t\to\infty$ to a finite random variable σ_∞^τ such that

$$\sigma_\infty^\tau \overset{\mathrm{d}}{=} \beta\int_0^\infty \mathrm{e}^{-X_t}\mathrm{d}t.$$

Conversely, if (2.19) does not hold, then $\sigma_t^\tau \xrightarrow{\mathrm{p}} \infty$ as $t\to\infty$.

Theorem 2.4.4. *$(\sigma_t^\tau)_{t\geq 0}$ an $(\sigma_t^\tau, G_t)_{t\geq 0}$ are time homogeneous Markov processes, moreover if σ_∞^τ exists and $\sigma_0^\tau \overset{\mathrm{d}}{=} \sigma_\infty^\tau$ (independent of $(L_t)_{t\geq 0}$), then $(\sigma_t^\tau)_{t\geq 0}$ is strictly stationary and $(G_t)_{t\geq 0}$ ia a process with stationary increments.*

Lemma 2.4.5. *Keep $c>0$*

(a) Let $\lambda>0$, then the Laplace transform $\mathbb{E}(\mathrm{e}^{-cX_t})$ of X_t at c is finite for some $t>0$, or equivalently for every $t>0$, if and only if $\mathbb{E}(L_1^{\tau c})<\infty$.

(b) If $\mathbb{E}(\mathrm{e}^{-cX_1})<\infty$ $|\Psi(c)|<\infty$, where $\Psi(c)=\log\mathbb{E}(\mathrm{e}^{-cX_1})$ and $\mathbb{E}(\mathrm{e}^{-cX_t})=\mathrm{e}^{t\Psi(c)}$, with

$$\Psi(c) = c\log\delta + \int_{\mathbb{R}}\left[\left(1+\frac{\lambda}{\delta}h(y)\right)^c - 1\right]\nu_L(\mathrm{d}y).$$

(c) If $\mathbb{E}(|L_1|^\tau)<\infty$ and $\Psi(1)<0$, then $\sigma_t^\tau \xrightarrow{\mathrm{d}} \sigma_\infty^\tau$, where σ_∞^τ is a finite random variable.

(d) If $\Psi(c)<0$ for some $c>0$, then $\Psi(d)<0$ for $d\in(0,c)$.

Proposition 2.4.6. *Let $\lambda>0$, $t>0$ and $\tilde{h}\geq 0$.*

(a) $\mathbb{E}(\sigma_t^\tau)<\infty$ if and only if $\mathbb{E}(|L_1|^\tau)<\infty$ and $\mathbb{E}(\sigma_0^\tau)<\infty$. If this is so

$$\mathbb{E}(\sigma_t^\tau) = -\frac{\beta}{\Psi(1)} + \left(\mathbb{E}(\sigma_0^\tau)+\frac{\beta}{\Psi(1)}\right)\mathrm{e}^{t\Psi(1)}.$$

(If $\Psi(1)=0$ the right hand side has to be interpreted as its limit as $\Psi(1)\to 0$, and in this case $\mathbb{E}(\sigma_t^\tau)=\beta t+\mathbb{E}(\sigma_0^\tau)$.)

(b) $\mathbb{E}(\sigma_t^{2\tau})<\infty$ if and only if $\mathbb{E}(|L_1|^{2\tau})<\infty$ and $\mathbb{E}(\sigma_0^{2\tau})<\infty$. In this case

$$\begin{aligned}\mathbb{E}(\sigma_t^{2\tau}) = {} & \frac{2\beta^2}{\Psi(1)\Psi(2)} + \frac{2\beta^2}{\Psi(2)-\Psi(1)}\left(\frac{\mathrm{e}^{t\Psi(2)}}{\Psi(2)} - \frac{\mathrm{e}^{t\Psi(1)}}{\Psi(1)}\right) + \\ & + 2\beta\mathbb{E}(\sigma_0^\tau)\left(\frac{\mathrm{e}^{t\Psi(2)}-\mathrm{e}^{t\Psi(1)}}{\Psi(2)-\Psi(1)}\right) + \mathbb{E}(\sigma_0^{2\tau})\mathrm{e}^{t\Psi(2)}.\end{aligned}$$

(Also here the right hand side has to be interpreted as its limit if some of the denominators are zero.) Moreover

$$\mathbb{C}\mathrm{ov}\left(\sigma_t^\tau, \sigma_{t+\tilde{h}}^\tau\right) = \mathbb{V}\mathrm{ar}(\sigma_t^\tau)\mathrm{e}^{\tilde{h}\Psi(1)}.$$

Proposition 2.4.7. *Let $\lambda > 0$. Then the k-th moment of σ_∞^τ (and of σ_t^τ if the volatility process is strictly stationary) is finite if and only if $\mathbb{E}(|L_1|^{\tau k}) < \infty$ and $\Psi(k) < 0$ for $k \in \mathbb{N}$ and we have*

$$\mathbb{E}(\sigma_\infty^{\tau k}) = k!\beta^k \prod_{l=1}^{k} \frac{1}{-\Psi(l)}.$$

Corollary 2.4.8. *If $(\sigma_t^\tau)_{t\geq 0}$ is strictly stationary with $\sigma_0^\tau \overset{\mathrm{d}}{=} \sigma_\infty^\tau$, then*

$$\mathbb{E}(\sigma_\infty^\tau) = -\frac{\beta}{\Psi(1)}$$

$$\mathbb{E}(\sigma_\infty^{2\tau}) = \frac{2\beta^2}{\Psi(1)\Psi(2)}$$

$$\mathbb{C}\mathrm{ov}\left(\sigma_t^\tau, \sigma_{t+\tilde{h}}^\tau\right) = \beta^2 \left(\frac{2}{\Psi(1)\Psi(2)} - \frac{1}{\Psi^2(1)}\right) \mathrm{e}^{\tilde{h}\Psi(1)}$$

provided that $\mathbb{E}(|L_1|^{\tau k}) < \infty$ and $\Psi(k) < 0$ with $k = 1$ for $\mathbb{E}(\sigma^2\tau_\infty)$ and $k = 2$ for $\mathbb{E}(\sigma_\infty^{2\tau})$ and $\mathbb{C}\mathrm{ov}\left(\sigma_t^\tau, \sigma_{t+\tilde{h}}^{2\tau}\right)$.

Theorem 2.4.9. *Let $k \in \mathbb{N}$, $\delta \in (0,1)$ and $\lambda \geq 0$. Then σ_∞^τ exists and has finite k-th moment if and only if*

$$\frac{1}{k}\int_{\mathbb{R}} \left[\left(1 + \frac{\lambda}{\delta}h(y)\right)^k - 1\right] \nu_L(\mathrm{d}y) < -\log\delta.$$

Proposition 2.4.10. *(a) For any Lévy process L with nonzero Lévy measure such that*

$$\int_{\mathbb{R}} \log(1 + h(y))\nu_L(\mathrm{d}y) < \infty$$

there exist $\delta, \lambda \in (0,1)$ such that σ_∞^τ exists, but $\mathbb{E}(\sigma_\infty^\tau) = \infty$.

(b) $k \in \mathbb{N}$, for any Lévy process L such that $\mathbb{E}(|L_1|^{\tau k}) < \infty$ and $\forall\delta \in (0,1)$ there exists $\lambda_\delta > 0$ such that σ_∞^τ exists with $\mathbb{E}(\sigma_\infty^{\tau k}) < \infty$ for every (λ, δ) such that $0 \leq \lambda \leq \lambda_\delta$.

(c) $\lambda > 0$. For no Lévy process L with nonzero Lévy measure do the moments of all orders of σ_∞^τ exist.

It is not known how to calculate the moments for every $\tau > 0$. Therefore only models with $\tau = 2$ will be considered in the following.

Proposition 2.4.11. *Suppose that L is a quadratic pure jump process (i.e. $C_L = 0$) with $\mathbb{E}(L_1^2) < \infty$ and $\mathbb{E}(L_1) = 0$, and that $\Psi(1) < 0$. Let $(\sigma_t^2)_{t\geq 0}$ be the stationary volatility process with $\sigma_0^2 \overset{\mathrm{d}}{=} \sigma_\infty^2$, then for $t \geq 0$, $\tilde{h} \geq r > 0$*

$$\mathbb{E}(G_t^{(r)}) = 0$$

$$\mathbb{E}((G_t^{(r)})^2) = -\frac{\beta r}{\Psi(1)}\mathbb{E}(L_1^2)$$

$$\mathbb{C}\mathrm{ov}\left(G_t^{(r)}, G_{t+\tilde{h}}^{(r)}\right) = 0.$$

Assume furthermore that $\mathbb{E}(L_1^4) < \infty$ *and* $\Psi(2) < 0$*, then*

$$\mathbb{C}\mathrm{ov}\left((G_t^{(r)})^2, (G_{t+\tilde{h}}^{(r)})^2\right) = \left(\frac{\mathrm{e}^{-r\Psi(1)} - 1}{-\Psi(1)}\right)\mathbb{E}(L_1^2)\mathrm{e}^{\tilde{h}\Psi(1)}\mathbb{C}\mathrm{ov}(G_r^2, \sigma_r^2).$$

If further $\lambda > 0$, $\mathbb{E}(L_1^8) < 0$, $\Psi(4) < 0$, $\int_{|x|\leq 1} |x|\nu_L(\mathrm{d}x) < \infty$ *and* $\int_{\mathbb{R}} x^3\nu_L(\mathrm{d}x) = 0$*, then* $\mathbb{C}\mathrm{ov}\left((G_t^{(r)})^2, (G_{t+\tilde{h}}^{(r)})^2\right) > 0$.

Proposition 2.4.12. *Suppose that* $\mathbb{E}(L_1) = 0$, $\mathbb{E}(L_1^2) < \infty$*, and that* $\Psi(1) < 0$*. Let* $(\sigma_t^2)_{t\geq 0}$ *be the stationary version. Then the process* $((G_{ri}^{(r)})^2)_{i\in\mathbb{N}}$ *has for every fixed* $r > 0$ *the autocorrelation structure of an ARMA(1,1) process.*

Theorem 2.4.13. *The stationary distribution* σ_∞^τ *is self-decomposable, i.e. such that*

$$\sigma_\infty^\tau \stackrel{\mathrm{d}}{=} k\sigma_\infty^\tau + Y,$$

where $k \in (0,1)$ *and* Y *is independent of* σ_∞^τ*.*

We choose L symmetric so that the asymmetry of the model originates in γ only. In particular we will use $\mathbb{E}(L_1) = 0$ and $\mathbb{E}(L_1^2) = 1$.

Remark 40. For a symmetric Lévy process the sign of the parameter γ is irrelevant because positive and negative jumps of the same size appear with the same probability. Hence we assume from now on that $\gamma \in [0,1)$.

Remark 41. Asymmetry of a COGARCH process can also be achieved by choosing an asymmetric Lévy process as driving process. Replacing in (2.18) the term $h(\Delta L_s)$ for L with symmetric Lévy measure ν_L by $(\Delta L_s)^2$ with the following asymmetric measure

$$\nu(\mathrm{d}x) = \nu_L(\mathrm{d}x)((1-\gamma)\mathbf{1}_{\{x\geq 0\}} + (1+\gamma)\mathbf{1}_{\{x<0\}})$$

yields the same model. However one prefers using the parameter γ in order to model the asymmetry because we can estimate this parameter by means of statistical procedures.

We now show a few plots regarding the behavior of the trajectories of the GJR-COGARCH(1,1) model and related processes for different values of γ in order to understand its influence on the asymmetry.
We simulated sample paths of a continuous time GJR-GARCH(1,1) process, $(G_t)_{t\geq 0}$, driven by a variance gamma with $\sigma = \tau = 1$ and $\mu = 0$. The parameters of the model were chosen as $\beta = 0.04$, $\eta = 0.053$, $\varphi = 0.038$ and $\gamma \in \{0, 0.2, 0.4, 0.6\}$. We can see if a large value is chosen for γ the negative jumps of the driving Lévy process are weighted more than positive jumps of the same size. In pictures showing log-returns sample paths we can observe bigger volatility clustering for high values of γ too.

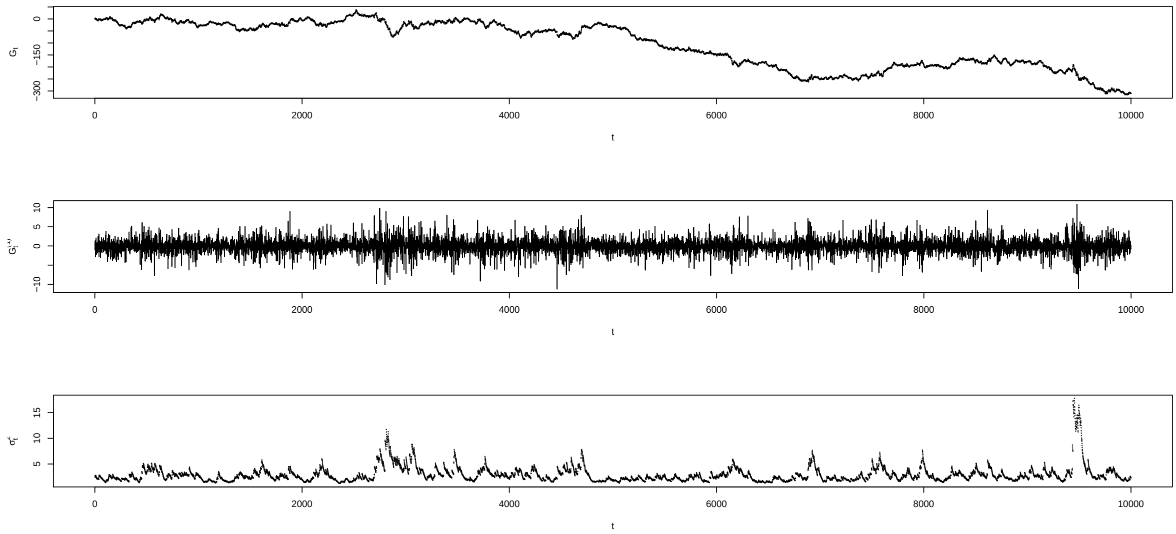

Figure 2.2: Simulation of a variance gamma driven GJR-COGARCH(1,1) process $(G_t)_{0\leq t\leq 10000}$ with parameters $\beta = 0.04$, $\eta = 0.053$, $\varphi = 0.038$, $\gamma = 0$, $\tau = \sigma = 1$ and $\mu = 0$, log-return process $(G_t^{(1)})_{0\leq t\leq 10000}$ and squared volatility process $(\sigma_t^2)_{0\leq t\leq 10000}$.

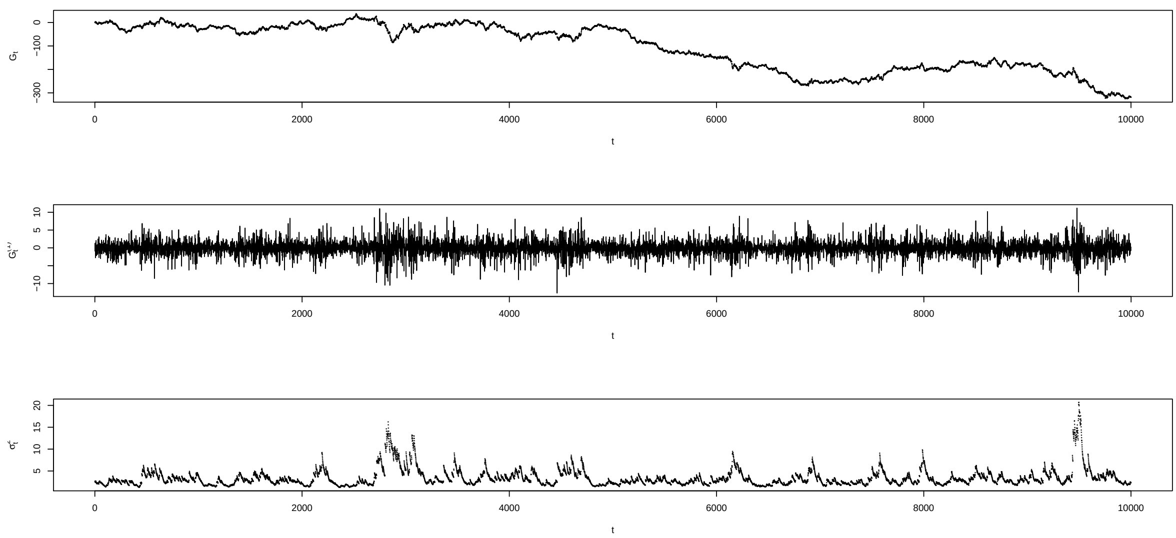

Figure 2.3: Simulation of a variance gamma driven GJR-COGARCH(1,1) process $(G_t)_{0\leq t\leq 10000}$ with parameters $\beta = 0.04$, $\eta = 0.053$, $\varphi = 0.038$, $\gamma = 0.2$, $\tau = \sigma = 1$ and $\mu = 0$, log-return process $(G_t^{(1)})_{0\leq t\leq 10000}$ and squared volatility process $(\sigma_t^2)_{0\leq t\leq 10000}$.

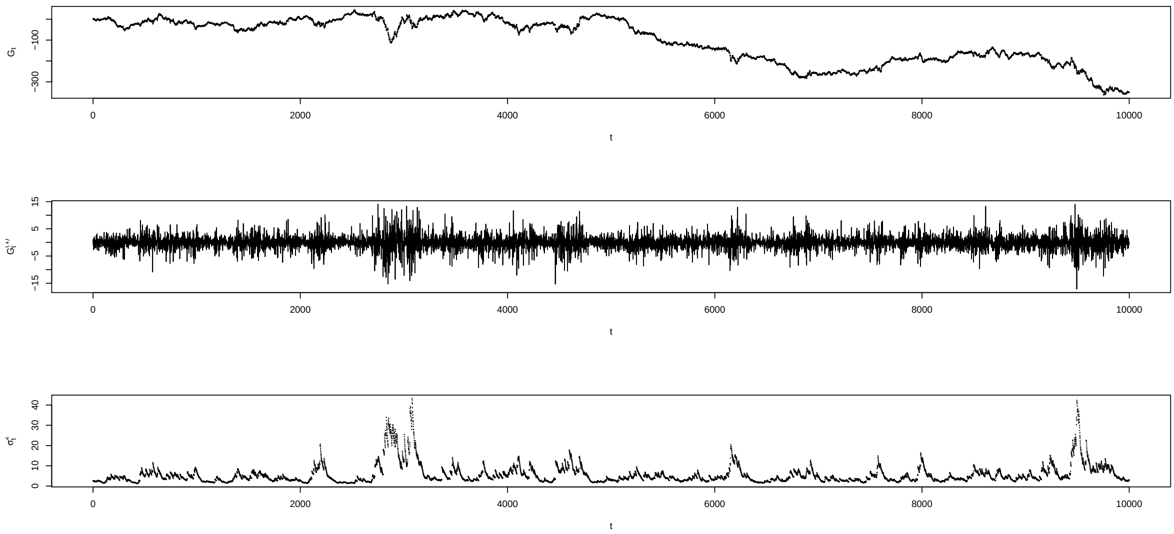

Figure 2.4: Simulation of a variance gamma driven GJR-COGARCH(1,1) process $(G_t)_{0\le t\le 10000}$ with parameters $\beta = 0.04$, $\eta = 0.053$, $\varphi = 0.038$, $\gamma = 0.4$, $\tau = \sigma = 1$ and $\mu = 0$, log-return process $(G_t^{(1)})_{0\le t\le 10000}$ and squared volatility process $(\sigma_t^2)_{0\le t\le 10000}$.

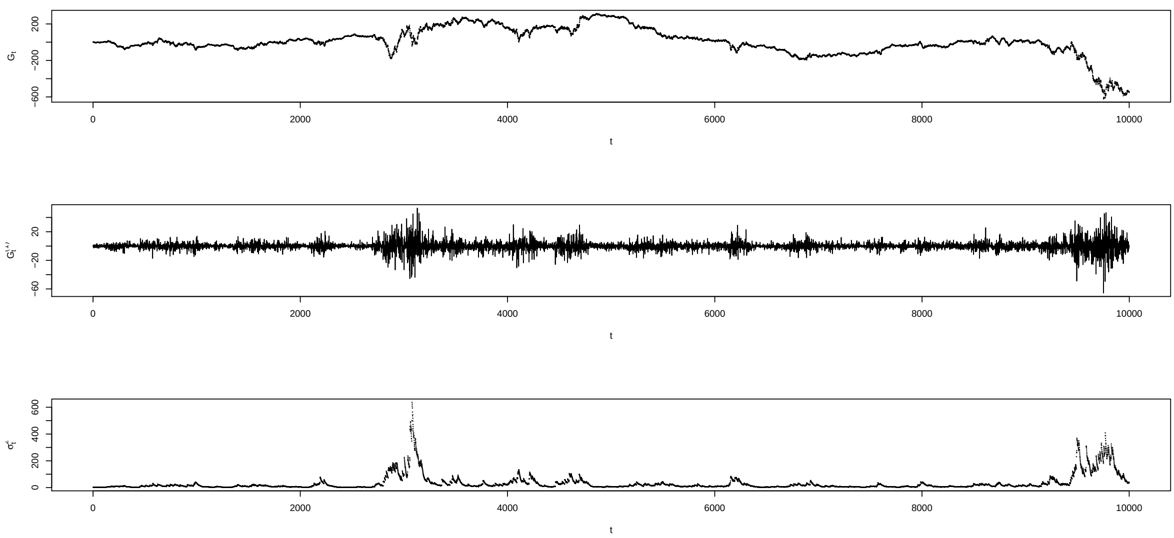

Figure 2.5: Simulation of a variance gamma driven GJR-COGARCH(1,1) process $(G_t)_{0\le t\le 10000}$ with parameters $\beta = 0.04$, $\eta = 0.053$, $\varphi = 0.038$, $\gamma = 0.6$, $\tau = \sigma = 1$ and $\mu = 0$, log-return process $(G_t^{(1)})_{0\le t\le 10000}$ and squared volatility process $(\sigma_t^2)_{0\le t\le 10000}$.

Chapter 3

Pseudo-maximum likelihood estimation for asymmetric COGARCH processes

For the COGARCH model different methods have been suggested: method of moments, pseudo-maximum likelihood (PML), Markov chain Monte Carlo (MCMC) and optimal prediction-based estimating functions (OPBEFs). See respectively [23], [36], [41] and [9]. The method of moments estimator, which is consistent and asymptotically normal under regularity conditions, is suitable only for equally spaced time series. On the other hand MCMC method is suitable for irregularly spaced time data, but it was just proposed for COGARCH models driven by a compound Poisson process. The pseudo-maximum likelihood estimator can be applied both for every driven Lévy process and for irregularly spaced data. It has also been proved by Kim and Lee in [26], once modified the likelihood function, that PML estimator is consistent and asymptotically normal. OPBEFs method needs higher moments and uses observations separeted by a constant time lag as well. Asymptotic properties are well known (see [50]). We refer to [21], [6] and [19] for estimation in discrete time models.
For the asymmetric model, by following [23] and [36], Behme, Klüppelberg and Mayr suggested in [5] the method of moments and the pseudo-maximum likelihood.
The aim of this chapter is the estimation of the GJR-COGARCH(1,1) model parameters $(\beta, \eta, \varphi, \gamma)$. Following [36] and [26] we will introduce a new pseudo-maximum likelihood estimator and prove the weak consistency. For doing that firstly we show how to fit the continuous time model to irregularly spaced time series data using discrete time GJR-GARCH methodology, by approximating the GJR-COGARCH with an embedded sequence of discrete time GJR-GARCH series. We need to discretize the continuous time volatility process and for this reason a "first jump" approximation is used (see [35] for details about this methodology). At the end we summarize the method of moments in order to compare this kind of approach with the pseudo-maximum likelihood estimator by means of Monte Carlo simulations.

3.1 Discrete approximation of the GJR-COGARCH

Take a deterministic sequence N_n such that $\lim_{n\to\infty} N_n = \infty$ and divide the finite interval $[0,T]$, with $T>0$, in N_n subintervals of length $\Delta_{t_k}(n) := t_k(n) - t_{k-1}(n)$, for $k = 1, \cdots, N_n$, in the following way

$$0 = t_0(n) < t_1(n) < \cdots < t_{N_n}(n) = T.$$

Assume that $\max_{k=1,\cdots,N_n} \Delta_{t_k}(n) \to 0$ as $n \to \infty$, and define for $n \in \mathbb{N}$ a discrete time process $(G_{n,k})_{k=1,\cdots,N_n}$ such that for $k = 1, \cdots, N_n$

$$G_{n,k} = G_{n,k-1} + \tilde{\sigma}_{n,k-1}\sqrt{\Delta_{t_k}(n)}\epsilon_{n,k}, \tag{3.1}$$

where $G_{n,0} = 0$ and the variance $\tilde{\sigma}^2_{n,k}$ follows the recursion

$$\begin{aligned}\tilde{\sigma}^2_{n,k} &= \beta\Delta_{t_k}(n) + \\ &+ \left(1 + \left[(1-\gamma)^2\mathbf{1}_{\left\{\epsilon_{n,k-1}\geq 0\right\}} + (1+\gamma)^2\mathbf{1}_{\left\{\epsilon_{n,k-1}<0\right\}}\right]\varphi\Delta_{t_k}(n)\epsilon^2_{n,k-1}\right)\mathrm{e}^{-\eta\Delta_{t_k}(n)}\tilde{\sigma}_{n,k-1} \\ &= \beta\Delta_{t_k}(n) + \mathrm{e}^{-\eta\Delta_{t_k}(n)}\tilde{\sigma^2}_{n,k-1} + \varphi\mathrm{e}^{-\eta\Delta_{t_k}(n)}(|Y_{n,k}| - \gamma Y_{n,k})^2\end{aligned} \tag{3.2}$$

with $Y_{n,k} := G_{n,k} - G_{n,k-1}$.
The innovations $(\epsilon_{n,k})_{k=1,\cdots,N_n}$ for $n \in \mathbb{N}$ are constructed using a *first jump approximation* of the Lévy process as follows. Take a strictly positive sequence $1 \geq a_n \downarrow 0$ satisfying $\lim_{n\to\infty} \Delta_{t_k}(n)\bar{\nu}^2_L(a_n) = 0$, where $\bar{\nu}_L(x) = \int_{|y|>x} \nu_L(\mathrm{d}y)$ is the tail of the Lévy measure. Such a sequence always exists as $\lim_{x\downarrow 0} x^2\bar{\nu}_L(x) = 0$ for any Lévy measure. For $n \in \mathbb{N}$ define the following stopping times

$$\tau_{n,k} := \inf\left\{t \in [t_{k-1}(n), t_k(n)) : |\Delta L_t| \geq a_n\right\}, \quad k = 1, \cdots, N_n.$$

$\tau_{n,k}$ is the time of the first jump of the driving Lévy process in the k-th interval whose magnitude exceeds a_n, if such a jump occurs. Thanks to the strong Markov property

$$\left(\mathbf{1}_{\left\{\tau_{n,k}<\infty\right\}}\Delta L_{\tau_{n,k}}\right)_{k=1,\cdots,N_n}$$

is for each $n \in \mathbb{N}$ a sequence of independent and identically distributed random variables with law

$$\frac{\nu_L(\mathrm{d}x)\mathbf{1}_{\{|x|>a_n\}}}{\bar{\nu}_L(a_n)}(1 - \mathrm{e}^{-\Delta_{t_k}(n)\bar{\nu}_L(a_n)}), \quad x \in \mathbb{R}\setminus\{0\}, k = 1, \cdots, N_n,$$

and with mass $\mathrm{e}^{-\Delta_{t_k}(n)\bar{\nu}_L(a_n)}$ at 0. These random variables have finite mean, $\varsigma_k(n)$, and variance, $\xi^2_k(n)$, since $\mathbb{E}(L_1^2) = 1$. So, the innovations required are

$$\epsilon_{n,k} = \frac{\mathbf{1}_{\left\{\tau_{n,k}<\infty\right\}}\Delta L_{\tau_{n,k}} - \varsigma_k(n)}{\xi_k(n)},$$

where for $n \in \mathbb{N}$ $\mathbb{E}(\epsilon_{n,1}) = 0$ and $\mathbb{E}^2(\epsilon_{n,1}) = 1$. Finally we take $\tilde{\sigma}^2_{n,0}$ independent of the $\epsilon_{n,k}$.

Remark 42. Equations (3.1) and (3.2) specify a GJR-GARCH(1,1)-type recursion. In the classical discrete time GJR-GARCH(1,1) process

$$\sigma_k^2 = a + bh(\sigma_{k-1}\epsilon_{k-1}) + c\sigma_{k-1}^2 \tag{3.3}$$

with $a, b, c, \gamma > 0$. If $\Delta_{t_k}(n)$ does not depend on k, then (3.2) is equivalent to (3.3) after rescaling by $\Delta_t(n)$ and a new parametrisation, and (3.1) becomes a rescaled GJR-GARCH(1,1) equation for the differenced sequence $G_{n,k} - G_{n,k-1}$. In general, even if the time grid is non equally spaced, we have convergence to the GJR-COGARCH(1,1).

Embed the discrete time processes $G_{n,\cdot}$ and $\sigma_{n,\cdot}^2$ into continuous time versions G_n and σ_n^2 with

$$G_n(t) := G_{n,k} \quad \text{and} \quad \sigma_n^2(t) = \sigma_{n,k}^2 \quad \text{for} \quad t \in [t_{k-1}(n), t_k(n)),$$

$0 \le t \le T$ and $G_n(0) = 0$. The processes G_n and σ_n^2 are in $\mathbb{D}[0,T]$, the space of càdlàg real-valued stochastic processes on $[0,T]$. The following theorem (see [5]) shows how a COGARCH process can be obtained as the limit of an embedded sequence of discrete time GARCH series.

Theorem 3.1.1. *The Skorokhod distance*[1] *between the processes (G, σ^2) and the discretized, piecewise constant processes $(G_n, \sigma_n^2)_{n\in\mathbb{N}}$ converges to 0 in probability as $n \to \infty$, i.e.*

$$\rho\left((G_n, \sigma_n^2), (G, \sigma^2)\right) \xrightarrow{\text{p}} 0, \quad as \quad n \to \infty.$$

Consequently, we also have convergence in distribution in $\mathbb{D}[0,T] \times \mathbb{D}[0,T]$, that is

$$(G_n, \sigma_n^2) \xrightarrow{\text{d}} (G, \sigma^2), \quad as \quad n \to \infty.$$

3.2 Estimation via pseudo-maximum likelihood

$G = (G_t)_{t\ge 0}$ is observed discretely with irregular time spaces. For each $n \in \mathbb{N}$ we set $N = N_n$,

$$0 = t_0 < t_1 < \cdots < t_N < \infty, \quad \Delta_{t_k} := t_k - t_{k-1}$$

and

$$Y_{n,k} := G_{t_k} - G_{t_{k-1}} = \int_{(t_{k-1}, t_k)} \sigma_s \mathrm{d}L_s.$$

$\Delta := \Delta_n := \max(\Delta_{t_1}, \cdots, \Delta_{t_N})$, where Δ_{t_k} are allowed to be nonidentical. We assume that $\Delta \to 0$ and $t_N \to \infty$ as $n \to \infty$.

[1] The Skorokhod distance J_1 between two $\mathbb{R}^d$-valued processes U and V, each in $\mathbb{D}^d[0,T]$ (the space of càdlàg $\mathbb{R}^d$-valued stochastic processes on $[0,T]$) is defined by

$$\rho(U,V) = \inf_{\lambda\in\Lambda} \left\{ \sup_{0\le t\le T} \left\| U_t - V_{\lambda(t)} \right\| + \sup_{0\le t\le T} |\lambda(t) - t| \right\}$$

where Λ is the set of strictly increasing continuous function with $\lambda(0) = 0$ and $\lambda(T) = T$.

Let $\vartheta^\circ = (\beta^\circ, \varphi^\circ, \eta^\circ, \gamma^\circ) \in \Theta$ be the vector of the (unknown) true parameters, $\vartheta = (\beta, \varphi, \eta, \gamma)$ and

$$\begin{aligned}\Theta := \{&\vartheta = (\beta, \varphi, \eta, \gamma) : \beta_* \le \beta \le \beta^*, \eta_* \le \eta \le \eta^*, \varphi* \le \varphi \le \varphi^*,\\ &\gamma_* \le \gamma \le \gamma^*, \eta - \varphi(1+\gamma^2) \ge c_*\}\end{aligned}$$

the parametric space with $0 < \beta_* < \beta^* < \infty$, $0 < \eta_* < \eta^* < \infty$, $0 < \varphi_* < \varphi^* < \infty$, $0 \le \gamma_* < \gamma^* < 1$, $0 < c_* < \infty$.
Let

$$\tilde{\sigma}^2_{n,0}(\vartheta) := \frac{\beta}{\eta - \varphi(1+\gamma^2)}$$

and

$$\tilde{\sigma}^2_{n,k}(\vartheta) := \beta\Delta_{t_k} + \mathrm{e}^{-\eta\Delta_{t_k}}\tilde{\sigma}^2_{n,k-1}(\vartheta) + \varphi \mathrm{e}^{-\eta\Delta_{t_k}} h(Y_{n,k})$$

for $k = 1, \cdots, N$ so that

$$\tilde{\sigma}^2_{n,k}(\vartheta) = \beta \sum_{i=0}^{k-1} \Delta_{t_{k-i}} \mathrm{e}^{-\eta(t_k - t_{k-i})} + \mathrm{e}^{-\eta t_k}\tilde{\sigma}^2_{n,0}(\vartheta) + \varphi \mathrm{e}^{-\eta t_k} \sum_{i=1}^{k} \mathrm{e}^{\eta t_{k-i}} h(Y_{n,k-i+1})$$

with $h(\Delta L_u) = (|\Delta L_u| - \gamma \Delta L_u)^2$, $0 \le \gamma < 1$. We can see $\tilde{\sigma}^2_{n,k}(\vartheta)$ as an estimate of $\sigma^2_{t_k}$ when $\vartheta = \vartheta^\circ$.

Our aim is to use a pseudo-maximum likelihood method to estimate the parameters $(\beta, \eta, \varphi, \gamma)$ from $Y_{n,1}, \cdots, Y_{n,N}$. We derive the pseudo-likelihood as follows.
Since $(\sigma_t^2)_{t\ge 0}$ is Markovian, then $Y_{n,k}$ is conditionally independent of $Y_{n,k-1}, Y_{n,k-2}, \cdots$, given $\mathcal{F}_{t_{k-1}}$, where $(\mathcal{F}_t)_{t\ge0}$ is the natural filtration of L satisfying the usual conditions. Consequently $\mathbb{E}(Y_{n,k}|\mathcal{F}_{t_{k-1}}) = 0$ and

$$\begin{aligned}\rho^2_{n,k} = \mathbb{E}(Y_{n,k}|\mathcal{F}_{t_{k-1}}) &= \mathbb{E}(L_1^2) \int_{t_{k-1}}^{t_k} \Big((\sigma^2_{t_{k-1}} - \mathbb{E}(\sigma_0^2))\mathrm{e}^{s-t_{k-1}\Psi(1)} + \mathbb{E}(\sigma^2_{s-t_{k-1}})\Big)\,\mathrm{d}s\\
&= (\sigma^2_{t_{k-1}} - \mathbb{E}(\sigma_0^2)) \int_{t_{k-1}}^{t_k} \mathrm{e}^{s-t_{k-1}\Psi(1)}\mathrm{d}s + \int_{t_{k-1}}^{t_k} \mathbb{E}(\sigma^2_{s-t_{k-1}})\mathrm{d}s\\
&= (\sigma^2_{t_{k-1}} - \mathbb{E}(\sigma_0^2)) \frac{\mathrm{e}^{\Delta_{t_k}\Psi(1)} - 1}{\Psi(1)} + \Delta_{t_k}\mathbb{E}(\sigma_0^2).\end{aligned}$$

Since

$$\begin{aligned}\Psi(1) &= -\eta^\circ + \int_{\mathbb{R}} ((1 + \varphi^\circ(|x| - \gamma^\circ x)^2) - 1)\nu_L(\mathrm{d}x)\\
&= -\eta^\circ + \varphi^\circ \int_{\mathbb{R}} (|x|^2 - 2\gamma^\circ x|x| + \gamma^{\circ 2}x^2)\nu_L(\mathrm{d}x)\\
&= -\eta^\circ + \varphi^\circ \left[(1+\gamma^{\circ 2}) \int_{\mathbb{R}} x^2 \nu_L(\mathrm{d}x) - 2\gamma^\circ \int_{\mathbb{R}} |x| x \nu_L(\mathrm{d}x)\right] = -\eta^\circ + \varphi^\circ(1+\gamma^{\circ 2}),\end{aligned}$$

then we have

$$\begin{aligned}\rho^2_{n,k} =& (\sigma^2_{t_{k-1}} - \mathbb{E}(\sigma_0^2)) \frac{\mathrm{e}^{(\varphi^\circ(1+\gamma^{\circ 2}) - \eta^\circ)\Delta_{t_k}} - 1}{-\eta^\circ + \varphi^\circ(1+\gamma^{\circ 2})} + \Delta_{t_k}\mathbb{E}(\sigma_0^2)\\
=& \left(\sigma^2_{t_{k-1}} - \frac{\beta^\circ}{\eta^\circ - \varphi^\circ(1+\gamma^{\circ 2})}\right) \frac{\mathrm{e}^{(-\eta^\circ + \varphi^\circ(1+\gamma^{\circ 2}))\Delta_{t_k}} - 1}{-\eta^\circ + \varphi^\circ(1+\gamma^{\circ 2})} + \frac{\beta^\circ \Delta_{t_k}}{\eta^\circ - \varphi^\circ(1+\gamma^{\circ 2})}\end{aligned}$$

and we can use

$$\tilde{\rho}^2_{n,k}(\vartheta) = \left(\tilde{\sigma}^2_{n,k-1}(\vartheta) - \frac{\beta}{\eta - \varphi(1+\gamma^2)}\right)\left(\frac{\mathrm{e}^{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}} - 1}{\varphi(1+\gamma^2) - \eta}\right) + \frac{\beta \Delta_{t_k}}{\eta - \varphi(1+\gamma^2)}$$

as estimates of the conditional variances of $Y_{n,k}$ when $\vartheta = \vartheta^\circ$.
We assume that $Y_{n,k}$ have conditionally gaussian distribution with zero mean and variance $\rho^2_{n,k}$ and use recursive conditioning to write for $m = m_n \in \mathbb{N}$ a pseudo-log-likelihood function for $Y_{n,1}, \cdots, Y_{n,N}$ as

$$\tilde{\mathcal{L}}_N(\vartheta) = -\frac{1}{2}\sum_{k=m}^{N} \frac{Y^2_{n,k}}{\rho^2_{n,k}} - \frac{1}{2}\sum_{k=m}^{N} \log \rho^2_{n,k} - \frac{N-m+1}{2}\log 2\pi. \tag{3.4}$$

We must substitute in (3.4) a calculable quantity for $\rho^2_{n,k}$, hence we need such for $\sigma^2_{t_{k-1}}$. We discretize the continuous time volatility process as was done in the previous section and in Theorem 3.1.1. Consequently we define the pseudo-gaussian log-likelihood function of ϑ as

$$\mathcal{L}_N(\vartheta) := \sum_{k=m}^{N} l_{n,k}(\vartheta)\Delta t_k, \quad l_{n,k}(\vartheta) = -\left(\frac{Y^2_{n,k}}{\tilde{\rho}^2_{n,k}(\vartheta)} + \log\frac{\tilde{\rho}^2_{n,k}(\vartheta)}{\Delta t_k}\right)$$

and we use as estimator $\hat{\vartheta}_n$ the measurable maximum point of $\mathcal{L}_N$, i.e.

$$\mathcal{L}_N(\hat{\vartheta}) = \max_{\vartheta\in\Theta} \mathcal{L}_N(\vartheta).$$

Remark 43. This pseudo-log-likelihood function is slightly different from (3.4) in which Δ_{t_k} does not appear. The reason will be more clear later when we prove the consistency of the estimator. Without this term the pseudo-maximum likelihood estimator is not consistent, for this reason we propose this kind of estimator. Behme, Klüppelberg and Mayr [5] (and Maller, Müller and Szimayer [36] for the symmetric model) adopted the version without Δ_{t_k}.

In order to prove the consistency theorem we need *regularity conditions* on the driving Lévy process and the sampling scheme.

Theorem 3.2.1. *Under*
C1. $\vartheta \in \Theta$, $\Delta \to 0$, $t_N \to \infty$, $t_m = o(t_N)$, $\mathrm{e}^{-\eta_* t_m} = O(\Delta^{1/2})$ *as* $n \to \infty$.
C2. $C_L = 0$, *i.e. there is no Brownian component* ($(L_t)_{t\geq 0}$ *is a quadratic pure jump Lévy process).*
C3. $\mathbb{E}(L_1) = 0$, $\mathbb{E}(L_1^2) = 1$, $\mathbb{E}(L_1^4) < \infty$ *and* $\Psi(2) < 0$, *where* $\Psi(z) = \log \mathbb{E}(\mathrm{e}^{-zX_1})$.

$$\hat{\vartheta}_n \xrightarrow{\mathrm{P}} \vartheta^\circ.$$

To prove the theorem we need some preliminary results we are going to show.

3.2.1 Proof of the consistency

To prove the consistency of the estimator we look for a function $\Upsilon(\vartheta)$ such that

$$\frac{1}{t_N}\mathcal{L}_N(\vartheta) \xrightarrow{\mathrm{P}} \Upsilon(\vartheta) \quad \text{and} \quad \Upsilon(\vartheta^\circ) = \max_{\vartheta\in\Theta} \Upsilon(\vartheta),$$

so that we can say that if the functions $\mathcal{L}_N$ and Υ are close each other, then their maximum points should be close each other too. Actually we need the uniform convergence in probability in order to guarantee this result and the following propositions and lemmas, though quite technical, aim to show this kind of convergence.

From now on C denotes a generic constant. We start extending the time domain of the processes L and X to $\mathbb{R}$ by letting

$$L_t := -L^*_{(-t)-}, \quad -\infty < t < 0$$

$$X_t := \eta^\circ t + \sum_{t<s\leq 0} \log\left(1 + \varphi^\circ h^\circ(\Delta L_s)\right), \quad -\infty < t < 0$$

with L^* independent copy of L.

Remark 44. It is clear that L and X are both càdlàg Lévy processes.

We now define

$$\sigma_u^2 := \beta^\circ \int_{-\infty}^{u} \mathrm{e}^{X_v - X_{u-}} \mathrm{d}v, \quad u \leq 0.$$

Lemma 3.2.2. *σ_u^2 is square integrable.*

Proof.

$$\sigma_u^2 = \beta^\circ \int_{-\infty}^{u} \mathrm{e}^{X_v - X_{u-}} \mathrm{d}v \stackrel{\mathrm{d}}{=} \beta^\circ \int_{-\infty}^{u} \mathrm{e}^{X_{v-u-}} \mathrm{d}v = \beta^\circ \int_{-\infty}^{0} \mathrm{e}^{X_s} \mathrm{d}s$$

and

$$\mathbb{E}\left(\int_{-\infty}^{0} \mathrm{e}^{X_s}\mathrm{d}s\right)^2 = \mathbb{E}\int_{-\infty}^{0}\int_{-\infty}^{0} \mathrm{e}^{X_r}\mathrm{e}^{X_s}\mathrm{d}r\mathrm{d}s = \mathbb{E}\int_{-\infty}^{0}\int_{-\infty}^{0} \mathrm{e}^{X_s - X_r}\mathrm{e}^{2X_r}\mathrm{d}r\mathrm{d}s =$$
$$= \int_{-\infty}^{0}\int_{-\infty}^{s} \mathbb{E}(\mathrm{e}^{X_s - X_r})\mathbb{E}(\mathrm{e}^{2X_r})\mathrm{d}r\mathrm{d}s + \int_{-\infty}^{0}\int_{s}^{0} \mathbb{E}(\mathrm{e}^{X_r - X_s})\mathbb{E}(\mathrm{e}^{2X_s})\mathrm{d}r\mathrm{d}s < \infty.$$

□

Remark 45. Note that $(\sigma_u^2)_{u\leq 0}$ is strictly stationary as

$$\beta^\circ \int_{-\infty}^{u} \mathrm{e}^{X_v - X_{u-}} \mathrm{d}v \stackrel{\mathrm{d}}{=} \beta^\circ \int_{-\infty}^{0} \mathrm{e}^{X_s} \mathrm{d}s \stackrel{a.s.}{=} \beta^\circ \int_{-\infty}^{0} \mathrm{e}^{X_s - X_{0-}} \mathrm{d}s = \sigma_0^2.$$

Define

$$\sigma_0^2(\vartheta) := \beta/\eta + \varphi \int_{(-\infty,0)} \mathrm{e}^{\eta u} \sigma_u^2 \mathrm{d}\left(\sum_{0<s\leq u} h(\Delta L_s)\right)$$

and for $t > 0$

$$\sigma_t^2(\vartheta) := \beta/\eta + (\sigma_0^2(\vartheta) - \beta/\eta)\mathrm{e}^{-\eta t} + \varphi \mathrm{e}^{-\eta t} \int_{(0,t)} \mathrm{e}^{\eta u} \sigma_u^2 \mathrm{d}\left(\sum_{0<s\leq u} h(\Delta L_s)\right)$$
$$= \beta/\eta + \varphi \int_{(-\infty,t)} \mathrm{e}^{-\eta(t-u)} \sigma_u^2 \mathrm{d}\left(\sum_{0<s\leq u} h(\Delta L_s)\right).$$

Remark 46. Clearly $(\sigma_t^2(\vartheta))_{t\geq 0}$ is a càglàd process.

Lemma 3.2.3. $\mathbb{E}(\sigma_0^2(\vartheta)) < \infty$ *for all* $\vartheta \in \Theta$.

Proof.

$$\mathbb{E}\int_{(-\infty,0)} \mathrm{e}^{\eta u}\sigma_u^2 \mathrm{d}\left(\sum_{0<s\leq u} h(\Delta L_s)\right) =$$
$$= \mathbb{E}\int_{(-\infty,0)} \mathrm{e}^{\eta u}\sigma_u^2 \mathrm{d}\left((1+\gamma^2)\sum_{0<s\leq u}(\Delta L_s)^2 - 2\gamma \sum_{0<s\leq u}(\Delta L_s)|\Delta L_s|\right)$$
$$= \begin{cases} (1-\gamma)^2 \int_{(-\infty,0)} \mathrm{e}^{\eta u}\sigma_u^2 \mathrm{d}[L]_u & \text{if} \quad \Delta L_u > 0 \\ (1+\gamma)^2 \int_{(-\infty,0)} \mathrm{e}^{\eta u}\sigma_u^2 \mathrm{d}[L]_u & \text{if} \quad \Delta L_u \leq 0. \end{cases}$$

Since

$$\mathbb{E}\int_{(-\infty,0)} \mathrm{e}^{\eta u}\sigma_u^2 \mathrm{d}[L]_u = \mathbb{E}\sum_{u\leq 0} \mathrm{e}^{\eta u}\sigma_u^2(\Delta L_u)^2 = \mathbb{E}\int_{(-\infty,0]}\int_{\mathbb{R}} x^2 \mathrm{e}^{\eta u}\sigma_u^2 N(\mathrm{d}u,\mathrm{d}x) =$$
$$= \mathbb{E}\int_{-\infty}^{0}\int_{\mathbb{R}} x^2 \mathrm{e}^{\eta u}\sigma_u^2 \nu_L(\mathrm{d}x)\mathrm{d}u = \int_{\mathbb{R}} x^2\nu_L(\mathrm{d}x)\int_{-\infty}^{0}\mathbb{E}(\sigma_u^2)\mathrm{e}^{\eta u}\mathrm{d}u =$$
$$= \mathbb{E}(\sigma_0^2)\int_{-\infty}^{0} \mathrm{e}^{\eta u}\mathrm{d}u < \infty.$$

□

Lemma 3.2.4. $(\sigma_t^2(\vartheta))_{t\geq 0}$ *is strictly stationary and satisfies the following stochastic differential equation*

$$\mathrm{d}\sigma_{t+}^2(\vartheta) = (\beta - \eta\sigma_t^2(\vartheta))\mathrm{d}t + \varphi\sigma_t^2\mathrm{d}\left(\sum_{0<s\leq t} h(\Delta L_s)\right).$$

Especially $\sigma_t^2(\vartheta) \geq \beta/\eta$ *a.s. and* $\mathbb{E}(\sigma_t^4(\vartheta)) < \infty$.

Proof. Since

$$\sigma_t^2(\vartheta) = \beta/\eta + (\sigma_0^2(\vartheta) - \beta/\eta)\mathrm{e}^{-\eta t} + \varphi \mathrm{e}^{-\eta t}\int_{(0,t)} \mathrm{e}^{\eta u}\sigma_u^2\mathrm{d}\left(\sum_{0<s\leq u} h(\Delta L_s)\right)$$

$$\sigma_{t+}^2(\vartheta) - \sigma_{0+}^2(\vartheta) = (\sigma_0^2(\vartheta) - \beta/\eta)\mathrm{e}^{-\eta t} + \varphi\mathrm{e}^{-\eta t}\int_{(0,t]} \mathrm{e}^{\eta u}\sigma_u^2\mathrm{d}\left(\sum_{0<s\leq u} h(\Delta L_s)\right) +$$
$$- \varphi\int_{(-\infty,0]} \mathrm{e}^{\eta u}\sigma_u^2\mathrm{d}\left(\sum_{0<s\leq u} h(\Delta L_u)\right)$$
$$= (\sigma_0^2(\vartheta) - \beta/\eta)(\mathrm{e}^{-\eta t} - 1) + \varphi\mathrm{e}^{-\eta t}\int_{(0,t]} \mathrm{e}^{\eta u}\sigma_u^2\mathrm{d}\left(\sum_{0<s\leq u} h(\Delta L_s)\right).$$

$$
\begin{aligned}
\int_0^t (\beta - \eta\sigma_s^2(\vartheta))\mathrm{d}s &= \int_0^t (\beta - \eta\sigma_{s+}^2(\vartheta))\mathrm{d}s \\
&= \int_0^t \left[-\eta(\sigma_0^2(\vartheta) - \beta/\eta)\mathrm{e}^{-\eta s} - \eta\varphi\mathrm{e}^{-\eta s} \int_{(0,s]} \mathrm{e}^{\eta u}\sigma_u^2 \mathrm{d}\left(\sum_{0<k\le u} h(\Delta L_k) \right) \right] \mathrm{d}s \\
&= (\sigma_0^2(\vartheta) - \beta/\eta)(\mathrm{e}^{-\eta t} - 1) - \varphi \int_{(0,t]} \int_u^t \eta \mathrm{e}^{-\eta s}\mathrm{d}s\mathrm{e}^{\eta u}\sigma_u^2 \mathrm{d}\left(\sum_{0<k\le u} h(\Delta L_k) \right) \\
&= (\sigma_0^2(\vartheta) - \beta/\eta)(\mathrm{e}^{-\eta t} - 1) + \varphi \int_{(0,t]} \mathrm{e}^{-\eta(t-u)}\sigma_u^2 \mathrm{d}\left(\sum_{0<k\le u} h(\Delta L_k) \right) + \\
&\quad - \varphi \int_{(0,t]} \sigma_u^2 \mathrm{d}\left(\sum_{0<k\le u} h(\Delta L_k) \right).
\end{aligned}
$$

Hence, as

$$
\sigma_{t+}^2(\vartheta) - \sigma_{0+}^2(\vartheta) = \int_0^t (\beta - \eta\sigma_s^2(\vartheta))\mathrm{d}s + \varphi \int_{(0,t]} \sigma_u^2 \mathrm{d}\left(\sum_{0<k\le u} h(\Delta L_k) \right)
$$

$$
\mathrm{d}\sigma_{t+}^2(\vartheta) = (\beta - \eta\sigma_t^2(\vartheta))\mathrm{d}t + \varphi\sigma_t^2 \mathrm{d}\left(\sum_{0<s\le t} h(\Delta L_s) \right).
$$

We also have that for every $s, t \ge 0$

$$
\begin{aligned}
\sigma_t^2(\vartheta) &= \beta/\eta + \varphi \int_{(-\infty,t)} \mathrm{e}^{-\eta(t-u)}\sigma_u^2 \mathrm{d}\left(\sum_{0<k\le u} h(\Delta L_k) \right) \\
&= \beta/\eta + \varphi \int_{(-\infty,s)} \mathrm{e}^{-\eta(t-y)}\mathrm{e}^{\eta(t-s)}\sigma_{y+t-s}^2 \mathrm{d}\left(\sum_{0<k\le y+t-s} h(\Delta L_k) \right) \\
&\overset{\mathrm{d}}{=} \beta/\eta + \varphi \int_{(-\infty,s)} \mathrm{e}^{-\eta(s-y)}\sigma_y^2 \mathrm{d}\left(\sum_{0<k\le y} h(\Delta L_k) \right) = \sigma_s^2(\vartheta).
\end{aligned}
$$

Since $\sigma_0^2(\vartheta) \ge \beta/\eta$ a.s., then $\sigma_t^2(\vartheta) \ge \beta/\eta$ a.s. too.
Moreover

$$
\begin{aligned}
\sigma_0^2(\vartheta) &= \beta/\eta + \varphi \int_{(-\infty,0)} \mathrm{e}^{\eta u}\sigma_u^2 \mathrm{d}\left(\sum_{0<s\le u} h(\Delta L_s) \right) \\
&= \beta/\eta + \varphi \sum_{j=0}^{\infty} \int_{(-j-1,-j]} \mathrm{e}^{\eta u}\sigma_u^2 \mathrm{d}\left(\sum_{0<s\le u} h(\Delta L_s) \right),
\end{aligned}
$$

which implies that

$$
\sigma_0^2(\vartheta) \le \beta/\eta + \varphi \sum_{j=0}^{\infty} \mathrm{e}^{-\eta j} \int_{(-j-1,-j]} \sigma_u^2 \mathrm{d}\left(\sum_{0<s\le u} h(\Delta L_s) \right).
$$

It follows that

$$\sigma_0^4(\vartheta) \leq \beta^2/\eta^2 + \left[\sum_{j=0}^{\infty} \mathrm{e}^{-\eta j} \int_{(-j-1,-j]} \sigma_u^2 \mathrm{d}\left(\sum_{0<s\leq u} h(\Delta L_s)\right)\right]^2 + $$
$$+ 2\beta/\eta \sum_{j=0}^{\infty} \mathrm{e}^{-\eta j} \int_{(-j-1,-j]} \sigma_u^2 \mathrm{d}\left(\sum_{0<s\leq u} h(\Delta L_s)\right)$$

We just have to deal with

$$\left[\sum_{j=0}^{\infty} \mathrm{e}^{-\eta j} \int_{(-j-1,-j]} \sigma_u^2 \mathrm{d}\left(\sum_{0<k\leq u} h(\Delta L_k)\right)\right]^2 =$$
$$= \sum_{j=0}^{\infty} \mathrm{e}^{-2\eta j} \left[\int_{(-j-1,-j]} \sigma_u^2 \mathrm{d}\left(\sum_{0<k\leq u} h(\Delta L_k)\right)\right]^2 +$$
$$+ \sum_{j=0}^{\infty} \sum_{i=0 \wedge i\neq j}^{\infty} \mathrm{e}^{-\eta i}\mathrm{e}^{-\eta j} \int_{(-j-1,-j]} \sigma_u^2 \mathrm{d}\left(\sum_{0<k\leq u} h(\Delta L_k)\right) \int_{(-i-1,-i]} \sigma_u^2 \mathrm{d}\left(\sum_{0<k\leq u} h(\Delta L_k)\right).$$

Clearly

$$\mathbb{E} \sum_{j=0}^{\infty} \mathrm{e}^{-\eta j} \left[\int_{(-j-1,-j]} \sigma_u^2 \mathrm{d}\left(\sum_{0<k\leq u} h(\Delta L_k)\right)\right]^2 = \sum_{j=0}^{\infty} \mathrm{e}^{-\eta j} \mathbb{E} \left[\int_{(-j-1,-j]} \sigma_u^2 \mathrm{d}\left(\sum_{0<k\leq u} h(\Delta L_k)\right)\right]^2.$$

It is known that

$$\begin{aligned}
\mathbb{E} \sum_{-j-1<u\leq -j} \sigma_u^4 (\Delta L_u)^4 &= \mathbb{E} \int_{(-j-1,-j]} \int_{\mathbb{R}} \sigma_u^4 x^4 \nu_L(\mathrm{d}x) \mathrm{d}u \\
&= \int_{\mathbb{R}} x^4 \nu_L(\mathrm{d}x) \int_{(-j-1,-j]} \mathbb{E}(\sigma_u^4) \mathrm{d}u \\
&= \mathbb{E}(\sigma_0^4) \int_{\mathbb{R}} x^4 \nu_L(\mathrm{d}x) < \infty.
\end{aligned}$$

Remark 47. $\int_{\mathbb{R}} x^4 \nu_L(\mathrm{d}x)$ is finite since $\mathbb{E}(L_1^4) < \infty \Leftrightarrow \int_{|x|\geq 1} x^4 \nu_L(\mathrm{d}x) < \infty$ and because $\int_{|x|<1} x^4 \nu_L(\mathrm{d}x) \leq \int_{|x|<1} x^2 \nu_L(\mathrm{d}x) < \infty$ thanks to the properties of the Lévy measure.

For $u \neq s$

$$\begin{aligned}
\mathbb{E} \sum_{-j-1<u\leq -j} \sum_{-j-1<s\leq -j} \sigma_u^2 \sigma_s^2 (\Delta L_u)^2 (\Delta L_s)^2 &= \sum_{-j-1<u\leq -j} \mathbb{E} \left[\sigma_u^2 \sum_{-j-1<s\leq -j} \sigma_s^2 (\Delta L_s)^2 (\Delta L_u)^2\right] \\
&= \int_{-j-1}^{-j} \mathbb{E} \left(\sigma_u^2 \sum_{-j-1<s\leq -j} \sigma_s^2 (\Delta L_s)^2\right) \mathrm{d}u \\
&= \int_{-j-1}^{-j} \int_{-j-1}^{-j} \mathbb{E}(\sigma_u^2 \sigma_s^2) \mathrm{d}u \mathrm{d}s = \\
&= A(1 - \mathrm{e}^{\Psi(1)} - \mathrm{e}^{-\Psi(1)} + 1) + B
\end{aligned}$$

with A and B constants

$$A = \frac{\beta^{\circ 2}}{-\Psi^2(1)}\left(\frac{2}{\Psi(1)\Psi(2)} - \frac{1}{\Psi^2(1)}\right) \quad \text{and} \quad B = \frac{\beta^{\circ 2}}{\Psi^2(1)}.$$

Then

$$\sum_{j=0}^{\infty} \mathrm{e}^{-2\eta i}\left(\mathbb{E}(\sigma_0^4)\int_{\mathbb{R}} x^4 \nu_L(\mathrm{d}x) + \int_{-j-1}^{-j}\int_{-j-1}^{-j} \mathbb{E}(\sigma_u^2\sigma_s^2)\mathrm{d}u\mathrm{d}s\right) < \infty.$$

Furthermore

$$\begin{aligned}
&\sum_{j=0}^{\infty}\sum_{i=0\wedge i\neq j}^{\infty} \mathrm{e}^{-\eta i}\mathrm{e}^{-\eta j}\mathbb{E}\int_{(-j-1,-j]}\sigma_s^2\mathrm{d}[L]_s\int_{(-i-1,-i]}\sigma_u^2\mathrm{d}[L]_u = \\
&= \sum_{j=0}^{\infty}\sum_{i=0\wedge i\neq j}^{\infty} \mathrm{e}^{-\eta i}\mathrm{e}^{-\eta j}\mathbb{E}\left[\sum_{-j-1<s\leq -j}\sigma_s^2\sum_{-i-1<u\leq -i}\sigma_u^2(\Delta L_u)^2(\Delta L_s)^2\right] \\
&= \sum_{j=0}^{\infty}\sum_{i=0\wedge i\neq j}^{\infty} \mathrm{e}^{-\eta i}\mathrm{e}^{-\eta j}\int_{-j-1}^{-j}\mathbb{E}\left(\sigma_s^2\sum_{-i-1<u\leq -i}\sigma_u^2(\Delta L_u)^2\right)\mathrm{d}s \\
&= \sum_{j=0}^{\infty}\sum_{i=0\wedge i\neq j}^{\infty} \mathrm{e}^{-\eta i}\mathrm{e}^{-\eta j}\int_{-j-1}^{-j}\int_{-i-1}^{-i}\mathbb{E}(\sigma_s^2\sigma_u^2)\mathrm{d}u\mathrm{d}s < \infty
\end{aligned}$$

and the lemma is validated. □

Lemma 3.2.5. *$\sigma_0^2(\vartheta^\circ) = \sigma_0^2$ a.s., hence $\sigma_t^2(\vartheta^\circ) = \sigma_t^2$ a.s. for $t \geq 0$ and $\sigma_t^2 \geq \beta^\circ/\eta^\circ$.*

Proof.

$$\begin{aligned}
\sigma_0^2(\vartheta^\circ) &= \beta^\circ/\eta^\circ + \varphi^\circ\int_{(-\infty,0)}\mathrm{e}^{\eta^\circ u}\sigma_u^2\mathrm{d}\left(\sum_{0<k\leq u}h^\circ(\Delta L_k)\right) \\
&= \beta^\circ/\eta^\circ + \varphi^\circ\int_{(-\infty,0)}\mathrm{e}^{\eta^\circ u}\left(\beta^\circ\int_{-\infty}^{u}\mathrm{e}^{X_v - X_{-u-}}\mathrm{d}v\right)\mathrm{d}\left(\sum_{0<k\leq u}h^\circ(\Delta L_k)\right) \\
&= \beta^\circ/\eta^\circ + \beta^\circ\varphi^\circ\int_{-\infty}^{0}\left(\int_{(v,0)}\mathrm{e}^{\eta^\circ u - X_{u-}}\mathrm{d}\left(\sum_{0<k\leq u}h^\circ(\Delta L_k)\right)\right)\mathrm{e}^{X_v}\mathrm{d}v \\
&= \beta^\circ\int_{-\infty}^{0}\left[\mathrm{e}^{\eta^\circ v - X_v} + \varphi^\circ\int_{(v,0)}\mathrm{e}^{\eta^\circ u - X_{u-}}\mathrm{d}\left(\sum_{0<k\leq u}h^\circ(\Delta L_k)\right)\right]\mathrm{e}^{X_v}\mathrm{d}v.
\end{aligned}$$

Moreover if $S_s := \log(1+\varphi^\circ h^\circ(\Delta L_s))$

$$\begin{aligned}
e^{\eta^\circ v - X_v} &= e^{-\sum_{v<s\le 0}\log(1+\varphi^\circ h^\circ(\Delta L_s))} \\
&= \sum_{v<w\le 0}\left(e^{-\sum_{w\le s\le 0} S_s} - e^{-\sum_{w<s\le 0} S_s}\right) + 1 \\
&= \sum_{v<w\le 0} e^{-\sum_{w\le s\le 0} S_s}\left(1 - e^{-\sum_{w<s\le 0} S_s}/e^{-\sum_{w\le s\le 0} S_s}\right) + 1 \\
&= \sum_{v<w\le 0} e^{-\sum_{w\le s\le 0} S_s}\left(1 - e^{S_w}\right) + 1 \\
&= -\varphi^\circ \sum_{v<w\le 0} e^{-\sum_{w\le s\le 0} S_s} h^\circ(\Delta L_w) + 1 \\
&= -\varphi^\circ \int_{(v,0]} e^{\eta^\circ w - X_{w-}} \mathrm{d}\left(\sum_{0<k\le w} h^\circ(\Delta L_k)\right) + 1.
\end{aligned}$$

Then, since $h^\circ(\Delta L_0) = 0 = X_{0-}$ a.s.,

$$\sigma_0^2(\vartheta^\circ) = \beta^\circ \int_{-\infty}^0 e^{X_v - X_{0-}} \mathrm{d}v = \sigma_0^2$$

a.s.. We know that for every $t \ge 0$

$$\sigma_{0+}^2(\vartheta^\circ) = \sigma_{t+}^2(\vartheta^\circ) - \int_0^t (\beta^\circ - \eta^\circ \sigma_s^2(\vartheta^\circ))\mathrm{d}s - \varphi^\circ \int_{(0,t]} \sigma_s^2 \mathrm{d}\left(\sum_{0<k\le s} h(^\circ \Delta L_k)\right)$$

and

$$\sigma_{0+}^2 = \sigma_{t+}^2 - \int_0^t (\beta^\circ - \eta^\circ \sigma_s^2)\mathrm{d}s - \varphi^\circ \int_{(0,t]} \sigma_s^2 \mathrm{d}\left(\sum_{0<k\le s} h^\circ(\Delta L_k)\right),$$

then

$$\sigma_{t+}^2(\vartheta^\circ) + \eta^\circ \int_0^t \sigma_s^2(\vartheta^\circ)\mathrm{d}s = \sigma_t^2 + \eta^\circ \int_0^t \sigma_s^2 \mathrm{d}s \quad \text{a.s..}$$

It follows[2] that $\sigma_t^2(\vartheta^\circ) = \sigma_t^2$ a.s. and that $\sigma_t^2 \ge \beta^\circ/\eta^\circ$. □

Proposition 3.2.6.

$$\Upsilon(\vartheta) := -\mathbb{E}\left(\frac{\sigma_0^2}{\sigma_0^2(\vartheta)} + \log \sigma_0^2(\vartheta)\right)$$

has the unique maximum at $\vartheta = \vartheta^\circ$ and is uniformly continuous in $\vartheta \in \Theta$.

[2] Suppose that $f_t + \int_0^t f_s \mathrm{d}s = g_t + \int_0^t g_s \mathrm{d}s$. If we assume that $f_t \ne g_t$, for example $f_t > g_t$, then $\int_0^t g_s \mathrm{d}s - \int_0^t f_s \mathrm{d}s = \int_0^t (g_s - f_s)\mathrm{d}s < 0$, but at the same time $\int_0^t (g_s - f_s)\mathrm{d}s = f_t - g_t > 0$ and we have a contradiction.

Proof.

$$\begin{aligned}
\Upsilon(\vartheta) - \Upsilon(\vartheta^\circ) &= -\mathbb{E}\left(\frac{\sigma_0^2}{\sigma_0^2(\vartheta)} + \log\sigma_0^2(\vartheta)\right) + \mathbb{E}\left(\frac{\sigma_0^2}{\sigma_0^2(\vartheta^\circ)} + \log\sigma_0^2(\vartheta^\circ)\right) \\
&= -\mathbb{E}\left(\frac{\sigma_0^2}{\sigma_0^2(\vartheta)}\right) + 1 + \mathbb{E}\left(\log\frac{\sigma_0^2(\vartheta^\circ)}{\sigma_0^2(\vartheta)}\right) \\
&\leq -\mathbb{E}\left(\frac{\sigma_0^2}{\sigma_0^2(\vartheta)}\right) + 1 + \mathbb{E}\left(\frac{\sigma_0^2(\vartheta^\circ)}{\sigma_0^2(\vartheta)}\right) - 1 \\
&= \mathbb{E}\left(\frac{\sigma_0^2}{\sigma_0^2(\vartheta)}\right) - \mathbb{E}\left(\frac{\sigma_0^2}{\sigma_0^2(\vartheta)}\right) = 0.
\end{aligned}$$

For $\tilde{h} > 0$ and η_1, η_2 such that $\eta_* \leq \eta_2 < \eta_1 \leq \eta^*$

$$\begin{aligned}
&\left\|\int_{(-\infty,0)} |\mathrm{e}^{\eta_1 u} - \mathrm{e}^{\eta_2 u}|\sigma_u^2 \mathrm{d}[L]_u\right\|_2 = \left\|\int_{(-\infty,0)} \mathrm{e}^{\eta_2 u}|\mathrm{e}^{(\eta_1-\eta_2)u} - 1|\sigma_u^2 \mathrm{d}[L]_u\right\|_2 \leq \\
&\leq \left\|\sup_{-\tilde{h}<u<0} |\mathrm{e}^{(\eta_1-\eta_2)u} - 1| \int_{(-\infty,0)} \mathrm{e}^{\eta_2 u}\sigma_u^2 \mathrm{d}[L]_u + \mathrm{e}^{-\eta_2\tilde{h}} \int_{(-\infty,-\tilde{h}]} \mathrm{e}^{\eta_2(u+\tilde{h})}\sigma_u^2 \mathrm{d}[L]_u\right\|_2 \\
&= \left(\sup_{-\tilde{h}<u<0} |\mathrm{e}^{(\eta_1-\eta_2)} - 1| + \mathrm{e}^{-\eta_2\tilde{h}}\right) \left\|\int_{(-\infty,0]} \mathrm{e}^{\eta_2 u}\sigma_u^2 \mathrm{d}[L]_u\right\|_2
\end{aligned}$$

since

$$\int_{(-\infty,-\tilde{h}]} \mathrm{e}^{\eta_2(u+\tilde{h})}\sigma_u^2 \mathrm{d}[L]_u = \int_{(-\infty,0]} \mathrm{e}^{\eta_2 v}\sigma_{v-\tilde{h}}^2 \mathrm{d}[L]_{v-\tilde{h}} \stackrel{\mathrm{d}}{=} \int_{(-\infty,0]} \mathrm{e}^{\eta_2 u}\sigma_u^2 \mathrm{d}[L]_u.$$

Then

$$\begin{aligned}
&\lim_{\delta\downarrow 0} \sup_{|\eta_1-\eta_2|<\delta} \left\|\int_{(-\infty,0)} \mathrm{e}^{\eta_1 u}\sigma_u^2 \mathrm{d}[L]_u - \int_{(-\infty,0)} \mathrm{e}^{\eta_2 u}\sigma_u^2 \mathrm{d}[L]_u\right\|_2 \leq \\
&\leq \lim_{\delta\downarrow 0} \sup_{|\eta_1-\eta_2|<\delta} \left\|\int_{(-\infty,0)} |\mathrm{e}^{\eta_1 u} - \mathrm{e}^{\eta_2 u}|\sigma_u^2 \mathrm{d}[L]_u\right\|_2 = 0
\end{aligned}$$

as we can choose $\tilde{h}$ such that $\mathrm{e}^{-\eta_2\tilde{h}} < \epsilon \ \forall\delta \geq \tilde{\delta}$. □

Proposition 3.2.7. *Let* $\sigma_{n,k-1}^2(\vartheta) := \sigma_{t_{k-1}}^2(\vartheta)$, *then*

$$\frac{1}{t_N}\sum_{k=m}^{N}\left(\frac{\sigma_{n,k-1}^2}{\sigma_{n,k-1}^2(\vartheta)} + \log\sigma_{n,k-1}^2(\vartheta)\right)\Delta t_k \xrightarrow{\mathrm{p}} -\Upsilon(\vartheta).$$

Proof. For $t_{k-1} < u \leq t_k$

$$
\begin{aligned}
&|\sigma^2_{n,k-1}(\vartheta) - \sigma^2_u(\vartheta)| = \\
&\quad -\varphi \mathrm{e}^{-\eta u} \int_{(0,u)} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right)\Bigg| \\
&= \left|\varphi \mathrm{e}^{-\eta t_{k-1}} \int_{(-\infty,0)} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right) - \varphi \mathrm{e}^{-\eta u} \int_{(-\infty,0)} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right) + \right.\\
&\quad \left. +\varphi \mathrm{e}^{-\eta t_{k-1}} \int_{(0,t_{k-1})} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right) - \varphi \mathrm{e}^{-\eta u} \int_{(0,u)} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right)\right| \\
&= \left|\varphi \mathrm{e}^{-\eta t_{k-1}} \int_{(-\infty,t_{k-1})} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right) - \varphi \mathrm{e}^{-\eta u} \int_{(-\infty,u)} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right)\right| \\
&= \left|\varphi \int_{(-\infty,t_{k-1})} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right)(\mathrm{e}^{-\eta t_{k-1}} - \mathrm{e}^{-\eta u}) + \right.\\
&\quad \left. -\varphi \mathrm{e}^{-\eta u} \int_{(t_{k-1},t_k)} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right) + \varphi \mathrm{e}^{-\eta u} \int_{(u,t_k)} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right)\right| \\
&\leq \varphi(\mathrm{e}^{-\eta t_{k-1}} - \mathrm{e}^{-\eta t_k}) \int_{(-\infty,t_{k-1})} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right) + \\
&\quad + \left|\varphi \mathrm{e}^{-\eta u} \int_{(u,t_k)} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right) - \varphi \mathrm{e}^{-\eta u} \int_{(t_{k-1},t_k)} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right)\right| \\
&\leq \varphi \mathrm{e}^{-\eta u} \int_{(t_{k-1},t_k)} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right) + \\
&\quad + \varphi(\mathrm{e}^{-\eta t_{k-1}} - \mathrm{e}^{-\eta t_k}) \int_{(-\infty,t_{k-1})} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right).
\end{aligned}
$$

Hence, for $t_{k-1} < u \leq t_k$

$$
\begin{aligned}
&|\sigma^2_{n,k-1}(\vartheta) - \sigma^2_u(\vartheta)| \leq \\
&\quad \leq \varphi(\mathrm{e}^{-\eta t_{k-1}} - \mathrm{e}^{-\eta t_k}) \int_{(-\infty,t_{k-1})} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right) + \\
&\quad + \varphi \mathrm{e}^{\eta(t_k - t_{k-1})} \int_{(t_{k-1},t_k)} \sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right) \\
&= \varphi \mathrm{e}^{\eta \Delta t_k} \int_{(t_{k-1},t_k)} \sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right) + \varphi \mathrm{e}^{-\eta t_{k-1}} \int_{(-\infty,t_{k-1})} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right) + \\
&\quad - \varphi \mathrm{e}^{\eta(t_{k-1} - t_k)} \int_{(-\infty,t_{k-1})} \mathrm{e}^{\eta(s - t_{k-1})}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\leq s} h(\Delta L_i)\right).
\end{aligned}
$$

It follows that

$$\sup_{t_{k-1}<u\le t_k} |\sigma^2_{n,k-1}(\vartheta) - \sigma^2_u(\vartheta)| \le \varphi e^{\eta\Delta t_k} \int_{(t_{k-1},t_k)} \sigma^2_s d\left(\sum_{0<i\le s} h(\Delta L_i)\right) +$$
$$(1-e^{-\eta\Delta t_k})\varphi \int_{(-\infty,t_{k-1})} e^{-\eta(t_{k-1}-s)}\sigma^2_s d\left(\sum_{0<i\le s} h(\Delta L_i)\right).$$

For $t_{k-1} < s \le t_k$

$$\begin{aligned}
\sigma^2_s(\vartheta) &= \sigma^2_{n,k-1}(\vartheta) + (\sigma^2_0(\vartheta) - \beta/\eta)(e^{-\eta s} - e^{-\eta t_{k-1}}) + \\
&+ \varphi e^{-\eta s} \int_{(0,s)} e^{\eta u}\sigma^2_u d\left(\sum_{0<i\le u} h(\Delta L_i)\right) - \varphi e^{-\eta t_{k-1}} \int_{(0,t_{k-1})} e^{\eta u}\sigma^2_u d\left(\sum_{0<i\le u} h(\Delta L_i)\right) \\
&\le \sigma^2_{n,k-1}(\vartheta) + \varphi e^{-\eta s} \int_{(0,s)} e^{\eta u}\sigma^2_u d\left(\sum_{0<i\le u} h(\Delta L_i)\right) + \\
&\quad - \varphi e^{-\eta t_{k-1}} \int_{(0,t_{k-1})} e^{\eta u}\sigma^2_u d\left(\sum_{0<i\le u} h(\Delta L_i)\right) \\
&\le \sigma^2_{n,k-1}(\vartheta) + \varphi e^{-\eta t_{k-1}} \int_{(0,s)} e^{\eta u}\sigma^2_u d\left(\sum_{0<i\le u} h(\Delta L_i)\right) + \\
&\quad - \varphi e^{-\eta t_{k-1}} \int_{(0,t_{k-1})} e^{\eta u}\sigma^2_u d\left(\sum_{0<i\le u} h(\Delta L_i)\right) \\
&\le \sigma^2_{n,k-1}(\vartheta) + \varphi \int_{(t_{k-1},t_k)} e^{\eta u}\sigma^2_u d\left(\sum_{0<i\le u} h(\Delta L_i)\right).
\end{aligned}$$

Therefore

$$\sup_{t_{k-1}<s\le t_k} \sigma^2_s(\vartheta) \le \sigma^2_{n,k-1}(\vartheta) + \varphi \int_{(t_{k-1},t_k)} e^{\eta u}\sigma^2_u d\left(\sum_{0<i\le u} h(\Delta L_i)\right).$$

By Lemma 3.2.4

$$\begin{aligned}
\mathbb{E}\left(\sup_{t_{k-1}<s\le t_k} \sigma^2_s(\vartheta)\right)^2 &= \mathbb{E}\sup_{t_{k-1}<s\le t_k} \sigma^4_s(\vartheta) \le \\
&\le \mathbb{E}(\sigma^4_{n,k-1}(\vartheta)) + \varphi^2\mathbb{E}\left(\int_{(t_{k-1},t_k)} e^{\eta u}\sigma^2_u d\left(\sum_{0<i\le u} h(\Delta L_i)\right)\right)^2 + \\
&\quad + 2\varphi\mathbb{E}(\sigma^2_{n,k-1}(\vartheta)) \int_{(t_{k-1},t_k)} e^{\eta u}\sigma^2_u d\left(\sum_{0<i\le u} h(\Delta L_i)\right) \\
&\le \mathbb{E}(\sigma^4_{n,k-1}(\vartheta)) + \varphi^2\mathbb{E}\left(\int_{(t_{k-1},t_k)} e^{\eta u}\sigma^2_u d\left(\sum_{0<i\le u} h(\Delta L_i)\right)\right)^2 + \\
&\quad + 2\varphi \left\|\sigma^2_{n,k-1}(\vartheta)\right\|_2 \left\|\int_{(t_{k-1},t_k)} e^{\eta u}\sigma^2_u d\left(\sum_{0<i\le u} h(\Delta L_i)\right)\right\|_2 < \infty
\end{aligned}$$

and

$$\left\| \sup_{t_{k-1}<u\le t_k} |\sigma^2_{n,k-1}(\vartheta) - \sigma^2_u(\vartheta)| \right\|_2 \le \varphi e^{\eta \Delta t_k} \left\| \int_{(t_{k-1},t_k)} \sigma^2_s \mathrm{d}\left(\sum_{0<i\le s} h(\Delta L_i) \right) \right\|_2 + \\ + (1-e^{-\eta\Delta t_k})\varphi \left\| \int_{(-\infty,t_{k-1})} e^{-\eta(t_{k-1}-s)} \sigma^2_s \mathrm{d}\left(\sum_{0<i\le s} h(\Delta L_i) \right) \right\|_2 < \infty.$$

Then

$$\max_{m\le k\le N} \left\| \sup_{t_{k-1}<u\le t_k} |\sigma^2_{n,k-1}(\vartheta) - \sigma^2_u(\vartheta)| \right\|_2 = o(1).$$

One can also observe that

$$\begin{aligned}
&\mathbb{E}\left| \frac{1}{t_N} \sum_{k=m}^{N} \left(\frac{\sigma^2_{n,k-1}}{\sigma^2_{n,k-1}(\vartheta)} + \log \sigma^2_{n,k-1}(\vartheta) \right) \Delta t_k - \frac{1}{t_N} \int_{t_m}^{t_N} \left(\frac{\sigma^2_s}{\sigma^2_s(\vartheta)} + \log \sigma^2_s(\vartheta) \right) \mathrm{d}s \right| = \\
&= \mathbb{E}\left| \frac{1}{t_N} \sum_{k=m}^{N} \int_{t_{k-1}}^{tk} \left(\frac{\sigma^2_{n,k-1}}{\sigma^2_{n,k-1}(\vartheta)} + \log \sigma^2_{n,k-1}(\vartheta) \right) \mathrm{d}s + \right. \\
&\left. - \frac{1}{t_N} \sum_{k=m}^{N} \int_{t_{k-1}}^{t_k} \left(\frac{\sigma^2_s}{\sigma^2_s(\vartheta)} + \log \sigma^2_s(\vartheta) \right) \mathrm{d}s + \frac{1}{t_N} \int_{t_{m-1}}^{t_m} \left(\frac{\sigma^2_s}{\sigma^2_s(\vartheta)} + \log \sigma^2_s(\vartheta) \right) \mathrm{d}s \right| \\
&\le \mathbb{E}\left| \frac{1}{t_N} \sum_{k=m}^{N} \int_{t_{k-1}}^{tk} \left(\frac{\sigma^2_{n,k-1}}{\sigma^2_{n,k-1}(\vartheta)} - \frac{\sigma^2_s}{\sigma^2_s(\vartheta)} + \log \sigma^2_{n,k-1}(\vartheta) - \log \sigma^2_s(\vartheta) \right) \mathrm{d}s \right| + \\
&+ \frac{1}{t_N} \mathbb{E}\left| \int_{t_{m-1}}^{t_m} \left(\frac{\sigma^2_s}{\sigma^2_s(\vartheta)} + \log \sigma^2_s(\vartheta) \right) \mathrm{d}s \right| \\
&\le \mathbb{E} \frac{1}{t_N} \sum_{k=m}^{N} \sup_{t_{k-1}<s\le t} \left(\left| \frac{\sigma^2_{n,k-1}}{\sigma^2_{n,k-1}(\vartheta)} - \frac{\sigma^2_s}{\sigma^2_s(\vartheta)} \right| + \left| \log \sigma^2_{n,k-1}(\vartheta) - \log \sigma^2_s(\vartheta) \right| \right) \Delta_{t_k} + \\
&+ \frac{1}{t_N} \mathbb{E} \sup_{t_{k-1}<s\le t} \left| \frac{\sigma^2_s}{\sigma^2_s(\vartheta)} + \log \sigma^2_s(\vartheta) \right| \Delta_{t_m}.
\end{aligned}$$

This upper bound converges to 0 since

$$\begin{aligned}
\sup_{t_{k-1}<s\le t_k} \left| \frac{\sigma^2_{n,k-1}}{\sigma^2_{n,k-1}(\vartheta)} - \frac{\sigma^2_s}{\sigma^2_s(\vartheta)} \right| &\le \sup_{t_{k-1}<s\le t_k} \left| \frac{\sigma^2_{n,k-1}}{\sigma^2_{n,k-1}(\vartheta)} - \frac{\sigma^2_{n,k-1}}{\sigma^2_s(\vartheta)} \right| + \\
&+ \sup_{t_{k-1}<s\le t_k} \left| \frac{\sigma^2_{n,k-1}}{\sigma^2_s(\vartheta)} - \frac{\sigma^2_s}{\sigma^2_s(\vartheta)} \right| \\
&= \sup_{t_{k-1}<s\le t_k} \sigma^2_{n,k-1} \left| \frac{1}{\sigma^2_{n,k-1}(\vartheta)} - \frac{1}{\sigma^2_s(\vartheta)} \right| + \sup_{t_{k-1}<s\le t_k} \frac{1}{\sigma^2_s(\vartheta)} \left| \sigma^2_{n,k-1} - \sigma^2_s \right| \\
&\le \sup_{t_{k-1}<s\le t_k} \frac{\sigma^2_{n,k-1}}{\min(\sigma^2_{n,k-1}(\vartheta), \sigma^2_s(\vartheta))^2} |\sigma^2_{n,k-1}(\vartheta) - \sigma^2_s(\vartheta)| + \\
&+ \sup_{t_{k-1}<s\le t_k} \frac{1}{\sigma^2_s(\vartheta)} |\sigma^2_{n,k-1} - \sigma^2_s|.
\end{aligned}$$

Given that $\sigma_t^2(\vartheta) \geq \beta/\eta$ a.s. for every t taking the expectation

$$\begin{aligned}
&\mathbb{E} \sup_{t_{k-1}<s\leq t_k} \left| \frac{\sigma_{n,k-1}^2}{\sigma_{n,k-1}^2(\vartheta)} - \frac{\sigma_s^2}{\sigma_s^2(\vartheta)} \right| \leq \\
&\leq C\mathbb{E} \sup_{t_{k-1}<s\leq t_k} \sigma_{n,k-1}^2 |\sigma_{n,k-1}^2(\vartheta) - \sigma_s^2(\vartheta)| + C\mathbb{E} \sup_{t_{k-1}<s\leq t_k} |\sigma_{n,k-1}^2 - \sigma_s^2| \\
&\leq C \max_{m\leq k\leq N} \left\|\sigma_{n,k-1}^2\right\|_2 \left\| \sup_{t_{k-1}<s\leq t_k} \sigma_{n,k-1}^2 |\sigma_{n,k-1}^2(\vartheta) - \sigma_s^2(\vartheta)| \right\|_2 + \\
&\quad + C \max_{m\leq k\leq N} \mathbb{E} \sup_{t_{k-1}<s\leq t_k} |\sigma_{n,k-1}^2 - \sigma_s^2| \to 0
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E} \sup_{t_{k-1}<s\leq t_k} |\log \sigma_{n,k-1}^2(\vartheta) - \log \sigma_s^2(\vartheta)| \leq \\
&\leq \mathbb{E} \sup_{t_{k-1}<s\leq t_k} \frac{1}{\min(\sigma_{n,k-1}^2(\vartheta), \sigma_s^2(\vartheta))} |\sigma_{n,k-1}^2(\vartheta) - \sigma_s^2(\vartheta)| \\
&\leq C \max_{m\leq k\leq N} \mathbb{E} \sup_{t_{k-1}<s\leq t_k} |\sigma_{n,k-1}^2(\vartheta) - \sigma_s^2(\vartheta)| \to 0.
\end{aligned}$$

In this way the series can converge to 0 since one obtains

$$o(1)\frac{1}{t_N}\sum_{k=m}^{N} \Delta t_k = o(1)\frac{t_N - t_m}{t_N} \to 0.$$

Furthermore

$$\begin{aligned}
&\mathbb{E} \sup_{t_{k-1}<s\leq t_k} \left| \frac{\sigma_s^2}{\sigma_s^2(\vartheta)} + \log \sigma_s^2(\vartheta) \right| \Delta_{t_m} \leq \\
&\leq C \left\| \sup_{t_{k-1}<s\leq t_k} \sigma_s^2 \right\|_2 \Delta_{t_m} + \mathbb{E} \sup_{t_{k-1}<s\leq t_k} \frac{1}{\min(\sigma_s^2(\vartheta), 1)} |\sigma_s^2(\vartheta) - 1| \Delta_{t_m} \\
&\leq C \left(\left\| \sup_{t_{k-1}<s\leq t_k} \sigma_s^2 \right\|_2 + \mathbb{E} \sup_{t_{k-1}<s\leq t_k} |\sigma_s^2(\vartheta) - 1| \right) \Delta_{t_m} \to 0.
\end{aligned}$$

We also have

$$\begin{aligned}
&\frac{1}{t_N}\sum_{k=m}^{N} \left(\frac{\sigma_{n,k-1}^2}{\sigma_{n,k-1}^2(\vartheta)} + \log \sigma_{n,k-1}^2(\vartheta) \right) \Delta t_k - \mathbb{E}\left(\frac{\sigma_0^2}{\sigma_0^2(\vartheta)} + \log \sigma_0^2(\vartheta) \right) = \\
&= \frac{1}{t_N}\sum_{k=m}^{N} \left(\frac{\sigma_{n,k-1}^2}{\sigma_{n,k-1}^2(\vartheta)} + \log \sigma_{n,k-1}^2(\vartheta) \right) \Delta t_k - \frac{1}{t_N}\int_{t_m}^{t_N} \left(\frac{\sigma_s^2}{\sigma_s^2(\vartheta)} + \log \sigma_s^2(\vartheta) \right) \mathrm{d}s + \\
&\quad + \frac{1}{t_N}\int_{t_m}^{t_N} \left(\frac{\sigma_s^2}{\sigma_s^2(\vartheta)} + \log \sigma_s^2(\vartheta) \right) \mathrm{d}s - \mathbb{E}\left(\frac{\sigma_0^2}{\sigma_0^2(\vartheta)} + \log \sigma_0^2(\vartheta) \right).
\end{aligned}$$

Thanks to the previous calculations

$$\frac{1}{t_N}\sum_{k=m}^{N} \left(\frac{\sigma_{n,k-1}^2}{\sigma_{n,k-1}^2(\vartheta)} + \log \sigma_{n,k-1}^2(\vartheta) \right) \Delta t_k - \frac{1}{t_N}\int_{t_m}^{t_N} \left(\frac{\sigma_s^2}{\sigma_s^2(\vartheta)} + \log \sigma_s^2(\vartheta) \right) \mathrm{d}s \xrightarrow{L_1} 0$$

and by the ergodic theorem

$$\frac{1}{t_N}\int_{t_m}^{t_N}\left(\frac{\sigma_s^2}{\sigma_s^2(\vartheta)}+\log\sigma_s^2(\vartheta)\right)\mathrm{d}s-\mathbb{E}\left(\frac{\sigma_0^2}{\sigma_0^2(\vartheta)}+\log\sigma_0^2(\vartheta)\right)\xrightarrow{\mathrm{P}}0.$$

That ends the proof.

□

Lemma 3.2.8. *There exists a constant $c>0$ such that for all large n*

$$\min_{m\le k\le N}\inf_{\vartheta\in\Theta}\tilde{\sigma}_{n,k}^2(\vartheta)\wedge\min_{m\le k\le N}\inf_{\vartheta\in\Theta}\frac{\tilde{\rho}_{n,k}^2(\vartheta)}{\Delta t_k}>c\quad a.s..$$

Proof.

$$\begin{aligned}\tilde{\sigma}_{n,k}^2(\vartheta)&\ge\beta\sum_{i=0}^{k-1}\Delta_{t_{k-i}}\mathrm{e}^{-\eta(t_k-t_{k-i})}\ge\beta_*\sum_{i=0}^{k-1}\Delta_{t_{k-i}}\mathrm{e}^{-\eta(t_k-t_{k-i})}\\&=\frac{\beta_*}{\eta}\sum_{i=0}^{k-1}\eta\Delta_{t_{k-i}}\mathrm{e}^{\eta\Delta_{t_{k-i}}}\mathrm{e}^{-\eta(t_k-t_{k-i-1})}\ge\frac{\beta_*}{\eta^*}\sum_{i=0}^{k-1}(\mathrm{e}^{\eta\Delta_{t_{k-i}}}-1)\mathrm{e}^{-\eta(t_k-t_{k-i-1})}\\&=\frac{\beta_*}{\eta^*}(1-\mathrm{e}^{-\eta t_k})\ge\frac{\beta_*}{\eta^*}(1-\mathrm{e}^{-\eta t_N})\ge\frac{\beta_*}{\eta^*}(1-\epsilon)\end{aligned}$$

since there exists $\tilde{n}\in\mathbb{N}$ such that $\mathrm{e}^{-\eta t_N}<\epsilon$ for every $n\ge\tilde{n}$. We choose $\epsilon=1/2$, then

$$\min_{m\le k\le N}\inf_{\vartheta\in\Theta}\tilde{\sigma}_{n,k}^2(\vartheta)\ge\frac{\beta_*}{2\eta^*}>0\quad\text{a.s..}$$

$$\begin{aligned}\frac{\tilde{\rho}_{n,k}^2(\vartheta)}{\Delta t_k}&=\left(\tilde{\sigma}_{n,k-1}^2(\vartheta)-\frac{\beta}{\eta-\varphi(1+\gamma^2)}\right)\frac{\mathrm{e}^{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}-1}{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}+\frac{\beta\Delta_{t_k}}{(\eta-\varphi(1+\gamma^2))\Delta_{t_k}}\\&=\tilde{\sigma}_{n,k-1}^2(\vartheta)\frac{\mathrm{e}^{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}-1}{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}+\frac{\beta}{\eta-\varphi(1+\gamma^2)}\cdot\\&\quad\cdot\left(1-\frac{\mathrm{e}^{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}-1}{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}\right).\end{aligned}$$

We just deal with

$$\tilde{\sigma}_{n,k-1}^2(\vartheta)+\frac{\beta}{(\eta-\varphi(1+\gamma^2))}\left(1-\frac{\mathrm{e}^{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}-1}{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}\right).$$

We know that

$$\begin{aligned}\inf_{\vartheta\in\Theta}(f(\vartheta)+g(\vartheta))&=-\sup_{\vartheta\in\Theta}(-f(\vartheta)-g(\vartheta))\ge-\sup_{\vartheta\in\Theta}(-f(\vartheta))-\sup_{\vartheta\in\Theta}(-g(\vartheta))\\&=\inf_{\vartheta\in\Theta}f(\vartheta)+\inf_{\vartheta\in\Theta}g(\vartheta)\end{aligned}$$

so that

$$\min_{m\le k\le N}\inf_{\vartheta\in\Theta}(f(\vartheta)+g(\vartheta)) \ge \min_{m\le k\le N}(\inf_{\vartheta\in\Theta} f(\vartheta)+\inf_{\vartheta\in\Theta} g(\vartheta)) \ge \min_{m\le k\le N}\inf_{\vartheta\in\Theta} f(\vartheta)+\min_{m\le k\le N}\inf_{\vartheta\in\Theta} g(\vartheta).$$

Then

$$\min_{m\le k\le N}\inf_{\vartheta\in\Theta}\tilde\sigma^2_{n,k-1}(\vartheta)+\min_{m\le k\le N}\inf_{\vartheta\in\Theta}\frac{\beta}{(\eta-\varphi(1+\gamma^2))}\left(1-\frac{\mathrm{e}^{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}-1}{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}\right)\ge$$
$$\ge\frac{\beta_*}{2\eta^*}-\frac{\beta_*}{4\eta^*}=\frac{\beta_*}{4\eta^*}>0\quad\text{a.s.}$$

as

$$\left(1-\frac{\mathrm{e}^{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}-1}{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}\right)\to 0.$$

□

Lemma 3.2.9.

$$\mathbb{E}\left(\int_{(0,\tilde h]}G_{s-}\sigma_s\mathrm{d}L_s\right)^2=O(\tilde h^2),\quad \tilde h\to 0.$$

Proof.

$$\mathbb{E}(\sigma_t^2-\sigma_0^2)^2=2\mathbb{E}(\sigma_0^4)-2\mathbb{E}(\sigma_t^2\sigma_0^2)=\left(\frac{4\beta^2}{\Psi(1)\Psi(2)}-\frac{2\beta^2}{\Psi^2(1)}\right)(1-\mathrm{e}^{t\Psi(1)})=O(t)$$

as $t\to 0$. Further with $\tilde h>0$

$$\begin{aligned}\mathbb{E}\left(\int_{(0,\tilde h]}G_{s-}\sigma_s\mathrm{d}L_s\right)^2&=\int_{(0,\tilde h]}\mathbb{E}(G^2_{s-}\sigma^2_s)\mathrm{d}s\\&=\int_{(0,\tilde h]}\mathbb{E}[G^2_{s-}(\sigma^2_s-\sigma^2_0)]\mathrm{d}s+\int_{(0,\tilde h]}\mathbb{E}(G^2_{s-}\sigma^2_0)\mathrm{d}s.\end{aligned}$$

$$|\mathbb{E}[G^2_{s-}(\sigma^2_s-\sigma^2_0)]|\le\mathbb{E}^{1/2}(G^4_{s-})\mathbb{E}^{1/2}(G^2_{s-}\sigma^2_0)^2=O(s)$$

as $s\to 0$ since

$$\mathbb{E}(G^4_{s-})=\mathbb{E}(L_1^4)\int_0^s\mathbb{E}(\sigma_0^4)\mathrm{d}u=\mathbb{E}(L_1^4)\mathbb{E}(\sigma_0^4)s.$$

$$\begin{aligned}\mathbb{E}(G^2_{s-}\sigma^2_0)&=\mathbb{E}[\mathbb{E}(G^2_{s-}|\mathcal{F}_0)\sigma_0^2]=\mathbb{E}[\sigma_0^2\int_0^s\mathbb{E}(\sigma_u^2|\mathcal{F}_0)\mathrm{d}u]\\&=\mathbb{E}\left[\sigma_0^2\int_0^s(\sigma_0^2-\mathbb{E}(\sigma_0^2))\mathrm{e}^{u\Psi(1)}\mathrm{d}u+\sigma_0^2\int_0^s\mathbb{E}(\sigma_u^2)\mathrm{d}u\right]\\&=\mathbb{E}\left(\sigma_0^2\left(\sigma_0^2-\mathbb{E}(\sigma_0^2)\frac{\mathrm{e}^{s\Psi(1)-1}}{\Psi(1)}\right)\right)+\mathbb{E}^2(\sigma_0^2)s=O(s)\end{aligned}$$

for $s\to 0$. Then there exist δ, M such that if we choose $s<\tilde h<\delta$

$$\int_{(0,\tilde h]}\mathbb{E}[G^2_{s-}(\sigma^2_s-\sigma^2_0)]\mathrm{d}s+\int_{(0,\tilde h]}\mathbb{E}(G^2_{s-}\sigma^2_0)\mathrm{d}s\le 2M\int_0^{\tilde h}s\mathrm{d}s=M\tilde h^2.$$

□

Lemma 3.2.10. *Suppose that* $\mathrm{e}^{-\eta_* t_m} = O(\Delta^{1/2})$, *then*

$$\max_{m\le k\le N} \left\|\tilde\sigma^2_{n,k}(\vartheta) - \sigma^2_{n,k}(\vartheta)\right\|_2 = O(\Delta^{1/2}).$$

Proof.

$$\sigma^2_{n,k}(\vartheta) - \tilde\sigma^2_{n,k}(\vartheta) = \beta/\eta + (\sigma_0^2(\vartheta)-\beta/\eta)\mathrm{e}^{-\eta t_k} + \varphi \mathrm{e}^{-\eta t_k}\int_{(0,t_k)} \mathrm{e}^{\eta s}\sigma_s^2 \mathrm{d}\left(\sum_{0<i\le s} h(\Delta L_i)\right) + $$
$$-\beta\sum_{i=0}^{k-1}\Delta_{t_{k-i}}\mathrm{e}^{-\eta(t_k-t_{k-i})} - \mathrm{e}^{-\eta t_k}\tilde\sigma^2_{n,0}(\vartheta) - \varphi\mathrm{e}^{-\eta t_k}\sum_{i=1}^{k}\mathrm{e}^{\eta t_{k-i}} h(Y_{n,k-i+1})$$

with

$$h(Y_{n,k-i+1}) = \begin{cases}(1-\gamma)^2 Y^2_{n,k-i+1} & Y_{n,k-i+1} > 0\\ (1+\gamma)^2 Y^2_{n,k-i+1} & Y_{n,k-i+1}\le 0.\end{cases}$$

We observe that

$$\mathrm{e}^{-\eta t_k}\tilde\sigma^2_{n,0}(\vartheta) = \mathrm{e}^{-\eta t_k}\frac{\beta}{\eta-\varphi(1+\gamma^2)} \le \mathrm{e}^{-\eta_* t_m}\frac{\beta}{\eta-\varphi(1+\gamma^2)} = O(\Delta^{1/2}),$$
$$\max_{m\le k\le N}\left\|(\sigma_0^2(\vartheta)-\beta/\eta)\mathrm{e}^{-\eta t_k}\right\|_2 = \max_{m\le k\le N}\left\|(\sigma_0^2(\vartheta)-\beta/\eta)\right\|_2 \mathrm{e}^{-\eta t_k} \le$$
$$\le \max_{m\le k\le N}\left\|(\sigma_0^2(\vartheta)-\beta/\eta)\right\|_2 \mathrm{e}^{-\eta_* t_m} = O(\Delta^{1/2})$$

and

$$\beta/\eta\left(1-\sum_{i=0}^{k-1}\eta\Delta t_{k-i}\mathrm{e}^{-\eta(t_k-t_{k-i})}\right) = \beta/\eta\left(1-\sum_{i=0}^{k-1}\eta\Delta t_{k-i}\mathrm{e}^{\eta\Delta_{t_{k-1}}}\mathrm{e}^{\eta(t_{k-i-1}-t_k)}\right)\le$$
$$\le \beta/\eta\left(1-\sum_{i=0}^{k-1}\left(\mathrm{e}^{\eta\Delta_{t_{k-i}}}-1\right)\mathrm{e}^{\eta(t_{k-i-1}-t_k)}\right) = \beta/\eta\mathrm{e}^{-\eta t_k} = O(\Delta^{1/2}).$$

Integration by parts gives the following equation.

$$\sum_{i=1}^{k}\mathrm{e}^{-\eta(t_k-t_{k-i})}Y^2_{n,k-i+1} =$$
$$= \sum_{i=1}^{k}\mathrm{e}^{-\eta(t_k-t_{k-i})}\left[2\int_{(t_{k-i},t_{k-i+1}]}(G_{u-}-G_{t_{k-i}})\sigma_u \mathrm{d}L_u + \int_{(t_{k-i},t_{k-i+1}]}\sigma_u^2\mathrm{d}[L]_u\right].$$

We only have to deal with

$$\mathrm{e}^{-\eta t_k}\int_{(0,t_k)}\mathrm{e}^{\eta s}\sigma_s^2\mathrm{d}\left(\sum_{0<i\le s}h(\Delta L_i)\right) - \mathrm{e}^{-\eta t_k}\sum_{i=1}^{k}\mathrm{e}^{\eta t_{k-i}}\int_{(t_{k-i},t_{k-i+1}]}\sigma_u^2\mathrm{d}[L]_u,$$

i.e. with

$$\mathrm{e}^{-\eta t_k}\sum_{(0<\le t_k)}\mathrm{e}^{\eta s}\sigma_s^2(\Delta L_s)^2 - \mathrm{e}^{-\eta t_k}\sum_{i=1}^{k}\mathrm{e}^{\eta t_{k-i}}\int_{(t_{k-i},t_{k-i+1}]}\sigma_u^2\mathrm{d}[L]_u, \tag{3.5}$$

and with

$$\sum_{i=1}^{k} \mathrm{e}^{-\eta(t_l - t_{k-i})} \int_{(t_{k-i},t_{k-i+1}]} (G_{u-} - G_{t_{k-i}})\sigma_u \mathrm{d}L_u.$$

$$\sum_{i=1}^{k} (\mathrm{e}^{\eta\Delta_{t_{k-i+1}}} - 1)\mathrm{e}^{-\eta(t_k - t_{k-i})} \int_{(t_{k-i},t_{k-i+1}]} \sigma_u^2 \mathrm{d}[L]_u =$$

$$= \sum_{i=1}^{k} \mathrm{e}^{\eta(t_{k-i+1}-t_k)} \sum_{t_{k-i}<u\le t_{k-i+1}} \sigma_u^2(\Delta L_u)^2 - \sum_{i=1}^{k} \mathrm{e}^{-\eta(t_k - t_{k-i})} \sum_{t_{k-i}<u\le t_{k-i+1}} \sigma_u^2(\Delta L_u)^2$$

$$= \sum_{t_{k-1}<u\le t_k} \sigma_u^2(\Delta L_u)^2 + \mathrm{e}^{\eta(t_{k-1}-t_k)} \sum_{t_{k-2}<u\le t_{k-1}} \sigma_u^2(\Delta L_u)^2 + \cdots$$

$$\cdots + \mathrm{e}^{\eta(t_1 - t_k)} \sum_{0<u\le t_1} \sigma_u^2(\Delta L_u)^2 - \sum_{i=1}^{k} \mathrm{e}^{-\eta(t_k - t_{k-i})} \sum_{t_{k-i}<u\le t_{k-i+1}} \sigma_u^2(\Delta L_u)^2,$$

then

$$e^{-\eta t_k} \sum_{(0<s\le t_k)} \mathrm{e}^{\eta s}\sigma_s^2(\Delta L_s)^2 - \mathrm{e}^{-\eta t_k} \sum_{i=1}^{k} \mathrm{e}^{\eta t_{k-i}} \int_{(t_{k-i},t_{k-i+1}]} \sigma_u^2 \mathrm{d}[L]_u \le$$

$$\le \sum_{i=1}^{k} (\mathrm{e}^{\eta\Delta_{t_{k-i+1}}} - 1)\mathrm{e}^{-\eta(t_k - t_{k-i})} \int_{(t_{k-i},t_{k-i+1}]} \sigma_u^2 \mathrm{d}[L]_u$$

$$= \sum_{i=1}^{k} (\mathrm{e}^{\eta\Delta_{t_{k-i+1}}} - 1)\mathrm{e}^{-\eta(t_k - t_{k-i})} \int_{(t_{k-i},t_{k-i+1}]} \sigma_u^2 \mathrm{d}u +$$

$$+ \sum_{i=1}^{k} (\mathrm{e}^{\eta\Delta_{t_{k-i+1}}} - 1)\mathrm{e}^{-\eta(t_k - t_{k-i})} \int_{(t_{k-i},t_{k-i+1}]} \sigma_u^2 \mathrm{d}([L]_u - u)$$

where the second term is a sum of martingale differences. Indeed for $r < u$

$$\mathbb{E}\left(\int_{(t_{k-i},t_{k-i+1}]} \sigma_u^2 \mathrm{d}[L]_u \middle| \mathcal{F}_r \right) = \int_{\mathbb{R}} x^2 \nu_L(\mathrm{d}x) \int_{t_{k-i}}^{t_{k-i+1}} \mathbb{E}_r(\sigma_u^2)\mathrm{d}u =$$

$$= \int_{t_{k-i}}^{t_{k-i+1}} \mathbb{E}_r(\sigma_u^2)\mathrm{d}u = \mathbb{E}\left(\int_{(t_{k-i},t_{k-i+1}]} \sigma_u^2 \mathrm{d}u \middle| \mathcal{F}_r \right).$$

The L^2-norm of (3.5) is $O(\Delta^{1/2})$ uniformly in $m \le k \le N$.

$$\mathbb{E}\left(\int_{(t_{k-i},t_{k-i+1}]} \sigma_u^2 \mathrm{d}([L]_u - u) \right)^2 = \mathbb{E}\left(\int_{(0,\Delta_{t_{k-i+1}}]} \sigma_u^2 \mathrm{d}([L]_u - u) \right)^2$$
$$= \mathbb{E}(\sigma_0^4)\Delta_{t_{k-i+1}} = O(\Delta).$$

$$\left\| \sum_{i=1}^{k} (\mathrm{e}^{\eta\Delta_{t_{k-i+1}}} - 1)\mathrm{e}^{-\eta t_k}\mathrm{e}^{\eta t_{k-i}} \int_{(t_{k-i},t_{k-i+1}]} \sigma_u^2 \mathrm{d}([L]_u - u) \right\|_2 \le$$

$$\le \sum_{i=1}^{k} |\mathrm{e}^{\eta\Delta_{t_{k-i+1}}} - 1|\mathrm{e}^{\eta t_{k-i}}\mathrm{e}^{-\eta t_k} \left\| \int_{(t_{k-i},t_{k-i+1}]} \sigma_u^2 \mathrm{d}([L]_u - u) \right\|_2 = O(\Delta) \le O(\Delta^{1/2}).$$

$$\left\|\sum_{i=1}^{k}(\mathrm{e}^{\eta\Delta_{t_{k-i+1}}}-1)\mathrm{e}^{-\eta t_k}\mathrm{e}^{\eta t_{k-i}}\int_{(t_{k-i},t_{k-i+1}]}\sigma_u^2\mathrm{d}u\right\|_2 \leq$$
$$\leq \sum_{i=1}^{k}|\mathrm{e}^{\eta\Delta_{t_{k-i+1}}}-1|\mathrm{e}^{\eta t_{k-i}}\mathrm{e}^{-\eta t_k}\left\|\int_{(t_{k-i},t_{k-i+1}]}\sigma_u^2\mathrm{d}u\right\|_2 = O(\Delta) \leq O(\Delta^{1/2}).$$

because

$$\mathbb{E}\left(\int_{(t_{k-i},t_{k-i+1}]}\sigma_u^2\mathrm{d}u\right)^2 = \mathbb{E}(\sigma_0^4)\int_{\mathbb{R}}x^4\nu_L(\mathrm{d}x)\Delta_{t_{k-i+1}} + B\Delta_{t_{k-i+1}}^2 +$$
$$+ A(1-\mathrm{e}^{-\Psi(1)\Delta_{t_{k-i+1}}}-\mathrm{e}^{\Psi(1)\Delta_{t_{k-i+1}}}+1) = O(\Delta).$$

Moreover

$$\sum_{i=1}^{k}\mathrm{e}^{-\eta(t_k-t_{k-i})}\int_{(t_{k-i},t_{k-i+1}]}(G_{u-}-G_{t_{k-1}})\sigma_u\mathrm{d}L_u$$

is also a sum of martingale differences

$$\mathbb{E}\left(\int_{(t_{k-i},t_{k-i+1}]}(G_{u-}-G_{t_{k-1}})\sigma_u\mathrm{d}L_u\right)^2 = \int_{(t_{k-i},t_{k-i+1}]}\mathbb{E}(G_{u-t_{k-i}}^2\sigma_u^2)\mathrm{d}u$$
$$= \int_{(0,\Delta_{t_{k-i+1}}]}\mathbb{E}(G_s^2\sigma_s^2)\mathrm{d}s = O(\Delta^2)$$

such that

$$\left\|\sum_{i=1}^{k}\mathrm{e}^{-\eta(t_k-t_{k-i})}\int_{(t_{k-i},t_{k-i+1}]}(G_{u-}-G_{t_{k-1}})\sigma_u\mathrm{d}L_u\right\|_2 \leq$$
$$\leq \sum_{i=1}^{k}\mathrm{e}^{-\eta t_k}\mathrm{e}^{\eta t_{k-i}}\left\|\int_{(t_{k-i},t_{k-i+1}]}(G_{u-}-G_{t_{k-1}})\sigma_u\mathrm{d}L_u\right\|_2 = O(\Delta^{3/2}) \leq O(\Delta^{1/2}).$$

□

Lemma 3.2.11.

$$\max_{m\leq k\leq N}\left\|\sup_{\vartheta\in\Theta}\frac{1}{\tilde{\rho}_{n,k}^2(\vartheta)}\left|\frac{\partial}{\partial\vartheta}\tilde{\rho}_{n,k}^2(\vartheta)\right|\right\|_2 < \infty.$$

Proof. Thanks to Lemma 3.2.8

$$\sup_{\vartheta\in\Theta}\frac{1}{\tilde{\rho}_{n,k}^2(\vartheta)}\left|\frac{\partial}{\partial\vartheta}\tilde{\rho}_{n,k}^2(\vartheta)\right| \leq$$
$$\leq C\sup_{\vartheta\in\Theta}\left(\left|\frac{\partial}{\partial\vartheta}\tilde{\sigma}_{n,k-1}^2(\vartheta)+O(\Delta_{t_k})\tilde{\sigma}_{n,k-1}^2(\vartheta)\right|\right)$$
$$\leq C\left(1+\sum_{i=1}^{k-1}\mathrm{e}^{-\eta_*/2(t_{k-1}-t_{k-i-1})}h(Y_{n,k-i-1})\right).$$

This bound and Lemma 3.2.10 imply the result. □

Proposition 3.2.12.

$$\sup_{\vartheta\in\Theta}\left|\frac{1}{t_N}\mathcal{L}_N(\vartheta)-\Upsilon(\vartheta)\right| = o_P(1).$$

Proof.

$$\begin{aligned}
\frac{1}{t_N}\mathcal{L}_N(\vartheta) &= \frac{1}{t_N}\sum_{k=m}^{N} l_{n,k}(\vartheta)\Delta_{t_k}\\
&= \frac{1}{t_N}\sum_{k=m}^{N}[l_{n,k}(\vartheta)-\mathbb{E}(l_{n,k}(\vartheta)|\mathcal{F}_{n,k-1})]\Delta_{t_k} + \frac{1}{t_N}\sum_{k=m}^{N}\mathbb{E}(l_{n,k}(\vartheta)|\mathcal{F}_{n,k-1})\Delta_{t_k}\\
&= \frac{1}{t_N}\sum_{k=m}^{N}[l_{n,k}(\vartheta)-\mathbb{E}(l_{n,k}(\vartheta)|\mathcal{F}_{n,k-1})]\Delta_{t_k} - \frac{1}{t_N}\sum_{k=m}^{N}\left(\frac{\rho^2_{n,k}}{\tilde{\rho}_{n,k}(\vartheta)}+\log\frac{\tilde{\rho}^2_{n,k}(\vartheta)}{\Delta_{t_k}}\right)\Delta_{t_k}
\end{aligned}$$

with

$$\rho^2_{n,k}(\vartheta) = \left(\sigma^2_{n.k-1}(\vartheta)-\frac{\beta}{\eta-\varphi(1+\gamma^2)}\right)\left(\frac{\mathrm{e}^{\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}-1}{\varphi(1+\gamma^2)-\eta}\right)+\frac{\beta\Delta_{t_k}}{\eta-\varphi(1+\gamma^2)}$$

and

$$\rho^2_{n,k} := \rho^2_{n,k}(\vartheta^\circ).$$

The first term is a sum of martingale differences which converges to 0 in probability because

$$\mathbb{E}[l_{n,k}(\vartheta)-\mathbb{E}(l_{n,k}(\vartheta)|\mathcal{F}_{n,k-1})] = 0.$$

We first prove the pointwise convergence in probability, i.e.

$$\frac{1}{t_N}\mathcal{L}(\vartheta)\xrightarrow{\mathrm{p}}\Upsilon(\vartheta).$$

$$\begin{aligned}
&\left|\frac{\rho^2_{n,k}(\vartheta)}{\Delta_{t_k}\sigma^2_{n,k-1}(\vartheta)}-1\right| =\\
&= \left|\left(\sigma^2_{n,k-1}(\vartheta)-\frac{\beta}{\eta-\varphi(1+\gamma^2)}\right)\left(\frac{\mathrm{e}^{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}-1}{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}\right)\frac{1}{\sigma^2_{n,k-1}(\vartheta)}+\right.\\
&\quad\left.+\frac{\beta}{(\eta-\varphi(1+\gamma^2))\sigma^2_{n,k-1}(\vartheta)}-1\right|\\
&= \left|\frac{\beta}{(\eta-\varphi(1+\gamma^2))\sigma^2_{n,k-1}(\vartheta)}\left(1-\frac{\mathrm{e}^{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}-1}{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}\right)+\frac{\mathrm{e}^{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}-1}{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}-1\right|\\
&\leq \left|\frac{\mathrm{e}^{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}-1}{(\varphi(1+\gamma^2)-\eta)\Delta_{t_k}}-1\right|(1+C) = O(\Delta).
\end{aligned}$$

Hence

$$\max_{m\leq k\leq N}\sup_{\vartheta\in\Theta}\left|\frac{\rho^2_{n,k}(\vartheta)}{\Delta_{t_k}\sigma^2_{n,k-1}(\vartheta)}-1\right| = O(\Delta). \tag{3.6}$$

Moreover

$$\max_{m\le k\le N}\left\|\frac{\rho^2_{n,k}}{\tilde\rho^2_{n,k}(\vartheta)}-\frac{\rho^2_{n,k}}{\rho^2_{n,k}(\vartheta)}+\log\tilde\rho^2_{n,k}(\vartheta)-\log\rho^2_{n,k}(\vartheta)\right\|_1\le$$
$$\le C\max_{m\le k\le N}(\left\|\sigma^2_{n,k-1}\right\|_2+1)\left\|\tilde\sigma^2_{n,k-1}(\vartheta)-\sigma^2_{n,k-1}(\vartheta)\right\|_2\to 0.$$

We need to show that

$$-\frac{1}{t_N}\sum_{k=m}^{N}\left(\frac{\rho^2_{n,k}}{\tilde\rho^2_{n,k}(\vartheta)}+\log\frac{\tilde\rho^2_{n,k}(\vartheta)}{\Delta_{t_k}}\right)\Delta_{t_k}\xrightarrow{\text{p}}\Upsilon(\vartheta)\quad\forall\vartheta\in\Theta.$$

One knows that

$$\max_{m\le k\le N}\left(\frac{\rho^2_{n,k}}{\tilde\rho^2_{n,k}(\vartheta)}-\frac{\rho^2_{n,k}}{\rho^2_{n,k}(\vartheta)}+\log\tilde\rho^2_{n,k}(\vartheta)-\log\rho^2_{n,k}(\vartheta)\right)\xrightarrow{\text{p}}0,$$

then

$$\frac{1}{t_N}\sum_{k=m}^{N}\left(\frac{\rho^2_{n,k}}{\tilde\rho^2_{n,k}(\vartheta)}+\log\frac{\tilde\rho^2_{n,k}(\vartheta)}{\Delta_{t_k}}\right)\Delta_{t_k}-\frac{1}{t_N}\sum_{k=m}^{N}\left(\frac{\rho^2_{n,k}}{\rho^2_{n,k}(\vartheta)}+\log\frac{\rho^2_{n,k}(\vartheta)}{\Delta_{t_k}}\right)\Delta_{t_k}\xrightarrow{\text{p}}0.$$

From (3.6)

$$\max_{m\le k\le N}\frac{\rho^2_{n,k}(\vartheta)}{\Delta_{t_k}\sigma^2_{n,k-1}(\vartheta)}\xrightarrow{\text{p}}1$$

so that

$$\frac{1}{t_N}\sum_{k=m}^{N}\left(\frac{\rho^2_{n,k}}{\rho^2_{n,k}(\vartheta)}+\log\frac{\rho^2_{n,k}(\vartheta)}{\Delta_{t_k}}\right)\Delta_{t_k}-\frac{1}{t_N}\sum_{k=m}^{N}\left(\frac{\sigma^2_{n,k-1}}{\sigma^2_{n,k-1}(\vartheta)}+\log\sigma^2_{n,k-1}(\vartheta)\right)\Delta_{t_k}\xrightarrow{\text{p}}0$$

and

$$\frac{1}{t_N}\sum_{k=m}^{N}\left(\frac{\rho^2_{n,k}}{\tilde\rho^2_{n,k}(\vartheta)}+\log\frac{\tilde\rho^2_{n,k}(\vartheta)}{\Delta_{t_k}}\right)\Delta_{t_k}-\frac{1}{t_N}\sum_{k=m}^{N}\left(\frac{\sigma^2_{n,k-1}}{\sigma^2_{n,k-1}(\vartheta)}+\log\sigma^2_{n,k-1}(\vartheta)\right)\Delta_{t_k}\xrightarrow{\text{p}}0$$

as

$$\max_{m\le k\le N}\left(\frac{\rho^2_{n,k}}{\Delta_{t_k}\sigma^2_{n,k-1}}\frac{\Delta_{t_k}\sigma^2_{n,k-1}(\vartheta)}{\rho^2_{n,k}(\vartheta)}-1\right)\xrightarrow{\text{p}}0\quad\text{and}\quad\frac{1}{t_N}\sum_{k=m}^{N}\frac{\sigma^2_{n,k-1}}{\sigma^2_{n,k-1}(\vartheta)}\Delta_{t_k}\to M<\infty.$$

All these results imply

$$\frac{1}{t_N}\sum_{k=m}^{N}\left(\frac{\rho^2_{n,k}}{\Delta_{t_k}\sigma^2_{n,k-1}}\frac{\Delta_{t_k}\sigma^2_{n,k-1}}{\Delta_{t_k}\sigma^2_{n,k-1}(\vartheta)}\frac{\Delta_{t_k}\sigma^2_{n,k-1}(\vartheta)}{\rho^2_{n,k}(\vartheta)}-\frac{\sigma^2_{n,k-1}}{\sigma^2_{n,k-1}(\vartheta)}+\log\frac{\rho^2_{n,k}(\vartheta)}{\Delta_{t_k}\sigma^2_{n,k-1}(\vartheta)}\right)\Delta_{t_k}\le$$
$$\le\frac{1}{t_N}\sum_{k=m}^{N}\max_{m\le k\le N}\left[\frac{\sigma^2_{n,k-1}}{\sigma^2_{n,k-1}(\vartheta)}\left(\frac{\rho^2_{n,k}}{\Delta_{t_k}\sigma^2_{n,k-1}}\frac{\Delta_{t_k}\sigma^2_{n,k-1}(\vartheta)}{\rho^2_{n,k}(\vartheta)}-1\right)+\log\frac{\rho^2_{n,k}(\vartheta)}{\Delta_{t_k}\sigma^2_{n,k-1}(\vartheta)}\right]\Delta_{t_k}$$
$$\xrightarrow{\text{p}}0.$$

Then

$$\frac{1}{t_N}\mathcal{L}_N(\vartheta) - \Upsilon(\vartheta) \xrightarrow{\text{p}} 0$$

since by Proposition 3.2.6

$$-\frac{1}{t_N}\sum_{k=m}^{N}\left(\frac{\sigma^2_{n,k-1}}{\sigma^2_{n,k-1}(\vartheta)} + \log\sigma^2_{n,k-1}(\vartheta)\right)\Delta_{t_k} - \Upsilon(\vartheta) \xrightarrow{\text{p}} 0.$$

We now verify the uniform convergence. By mean value theorem

$$\begin{aligned}
&\sup_{|\vartheta_1-\vartheta_2|<\tilde h}|\mathcal{L}_N(\vartheta_1)-\mathcal{L}_N(\vartheta_2)| = \sup_{|\vartheta_1-\vartheta_2|<\tilde h}\left|\sum_{k=m}^{N}(l_{n,k}(\vartheta_1)-l_{n,k}(\vartheta_2))\Delta_{t_k}\right| \le \\
&\le \sum_{k=m}^{N}\sup_{|\vartheta_1-\vartheta_2|<\tilde h}|l_{n,k}(\vartheta_1)-l_{n,k}(\vartheta_2)|\Delta_{t_k} \le \sup_{|\vartheta_1-\vartheta_2|<\tilde h}\sum_{k=m}^{N}|l_{n,k}(\vartheta_1)-l_{n,k}(\vartheta_2)|\Delta_{t_k} \\
&= \sum_{k=m}^{N}\sup_{|\vartheta_1-\vartheta_2|<\tilde h}\left|\frac{\partial}{\partial\vartheta}l_{n,k}(\vartheta)\right||\vartheta_1-\vartheta_2|\Delta_{t_k} \le \sum_{k=m}^{N}\sup_{|\vartheta_1-\vartheta_2|<\tilde h}\left|\frac{\partial}{\partial\vartheta}l_{n,k}(\vartheta)\right|\Delta_{t_k}\tilde h \\
&\le \sum_{k=m}^{N}\sup_{\vartheta\in\Theta}\left|\frac{\partial}{\partial\vartheta}l_{n,k}(\vartheta)\right|\Delta_{t_k}\tilde h,
\end{aligned}$$

and

$$\begin{aligned}
\max_{m\le k\le N}\mathbb{E}\sup_{\vartheta\in\Theta}\left|\frac{\partial}{\partial\vartheta}l_{n,k}(\vartheta)\right| &\le \max_{m\le k\le N}\mathbb{E}\sup_{\vartheta\in\Theta}\left(\frac{Y^2_{n,k}}{\tilde\rho^2_{n,k}(\vartheta)}+1\right)\left|\frac{1}{\tilde\rho^2_{n,k}(\vartheta)}\frac{\partial}{\partial\vartheta}\tilde\rho^2_{n,k}(\vartheta)\right| \\
&\le C\max_{m\le k\le N}\mathbb{E}\left(\frac{Y^2_{n,k}}{\Delta_{t_k}}+1\right)\sup_{\vartheta\in\Theta}\left|\frac{1}{\tilde\rho^2_{n,k}(\vartheta)}\frac{\partial}{\partial\vartheta}\tilde\rho^2_{n,k}(\vartheta)\right| \\
&= C\max_{m\le k\le N}\mathbb{E}\left(\frac{\mathbb{E}(Y^2_{n,k}|\mathcal{F}_{t_{k-1}})}{\Delta_{t_k}}+1\right)\sup_{\vartheta\in\Theta}\left|\frac{1}{\tilde\rho^2_{n,k}(\vartheta)}\frac{\partial}{\partial\vartheta}\tilde\rho^2_{n,k}(\vartheta)\right| \\
&\le C\max_{m\le k\le N}\mathbb{E}(\sigma^2_{n,k-1}+1)\sup_{\vartheta\in\Theta}\left|\frac{1}{\tilde\rho^2_{n,k}(\vartheta)}\frac{\partial}{\partial\vartheta}\tilde\rho^2_{n,k}(\vartheta)\right| \\
&\le C\max_{m\le k\le N}\left\|\sup_{\vartheta\in\Theta}\left|\frac{1}{\tilde\rho^2_{n,k}(\vartheta)}\frac{\partial}{\partial\vartheta}\tilde\rho^2_{n,k}(\vartheta)\right|\right\|_2 < \infty.
\end{aligned}$$

as

$$\begin{aligned}
\left|\frac{\partial}{\partial\vartheta}l_{n,k}(\vartheta)\right| &= \left|\frac{-Y^2_{n,k}}{\tilde\rho^4_{n,k}(\vartheta)}\frac{\partial}{\partial\vartheta}\tilde\rho^2_{n,k}(\vartheta) + \frac{1}{\tilde\rho^2_{n,k}(\vartheta)}\frac{\partial}{\partial\vartheta}\tilde\rho^2_{n,k}(\vartheta)\right| \\
&\le \frac{Y^2_{n,k}}{\tilde\rho^2_{n,k}(\vartheta)}\left|\frac{1}{\tilde\rho^2_{n,k}(\vartheta)}\frac{\partial}{\partial\vartheta}\tilde\rho^2_{n,k}(\vartheta)\right| + \left|\frac{1}{\tilde\rho^2_{n,k}(\vartheta)}\frac{\partial}{\partial\vartheta}\tilde\rho^2_{n,k}(\vartheta)\right| \\
&= \left(\frac{Y^2_{n,k}}{\tilde\rho^2_{n,k}(\vartheta)}+1\right)\left|\frac{1}{\tilde\rho^2_{n,k}(\vartheta)}\frac{\partial}{\partial\vartheta}\tilde\rho^2_{n,k}(\vartheta)\right|
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\mathbb{E}(Y_{n,k}^2|\mathcal{F}_{t_{k-1}})}{\Delta_{t_k}} = \\
&= \left(\sigma_{n,k-1}^2 - \frac{\beta^\circ}{\eta^\circ - \varphi^\circ(1+\gamma^{\circ 2})}\right) \frac{\mathrm{e}^{\Delta_{t_k}(\varphi^\circ(1+\gamma^{\circ 2})-\eta^\circ)} - 1}{\Delta_{t_k}(\varphi^\circ(1+\gamma^{\circ 2})-\eta^\circ)} + \frac{\beta^\circ}{\eta^\circ - \varphi^\circ(1+\gamma^{\circ 2})} \\
&= \sigma_{n,k-1}^2 \frac{\mathrm{e}^{\Delta_{t_k}(\varphi^\circ(1+\gamma^{\circ 2})-\eta^\circ)} - 1}{\Delta_{t_k}(\varphi^\circ(1+\gamma^{\circ 2})-\eta^\circ)} + \frac{\beta^\circ}{\eta^\circ - \varphi^\circ(1+\gamma^{\circ 2})}\left(1 - \frac{\mathrm{e}^{\Delta_{t_k}(\varphi^\circ(1+\gamma^{\circ 2})-\eta^\circ)} - 1}{\Delta_{t_k}(\varphi^\circ(1+\gamma^{\circ 2})-\eta^\circ)}\right) \\
&\leq C\sigma_{n,k-1}^2.
\end{aligned}$$

It follows that

$$\lim_{\tilde{h}\to 0}\limsup_{n\to\infty} \mathbb{E}\frac{1}{t_N}\sup_{|\vartheta_1-\vartheta_2|<\tilde{h}} |\mathcal{L}(\vartheta_1) - \mathcal{L}(\vartheta_2)| = 0. \tag{3.7}$$

Taking a $\vartheta \in \Theta$ and a finite subcover

$$\left\{B_{\tilde{h}}(\vartheta_i) := \left\{\vartheta \in \Theta : \vartheta_i - \tilde{h} < \vartheta < \vartheta_i + \tilde{h}\right\} : i = 1, \cdots, I\right\},$$

by the triangle inequality, we know that

$$\left|\frac{1}{t_N}\mathcal{L}(\vartheta) - \Upsilon(\vartheta)\right| \leq \left|\frac{1}{t_N}\mathcal{L}(\vartheta) - \frac{1}{t_N}\mathcal{L}(\vartheta_i)\right| + \left|\frac{1}{t_N}\mathcal{L}(\vartheta_i) - \Upsilon(\vartheta_i)\right| + |\Upsilon(\vartheta_i) - \Upsilon(\vartheta)|.$$

Choose ϑ_i so that $\vartheta \in B_{\tilde{h}}(\vartheta_i)$.

$$\left|\frac{1}{t_N}\mathcal{L}(\vartheta) - \frac{1}{t_N}\mathcal{L}(\vartheta_i)\right| \leq \frac{1}{t_N}\sup_{|\vartheta_1-\vartheta_2|<\tilde{h}} |\mathcal{L}(\vartheta_1) - \mathcal{L}(\vartheta_2)|$$

since $|\vartheta - \vartheta_i| < \tilde{h}$. Moreover

$$|\Upsilon(\vartheta_i) - \Upsilon(\vartheta)| \leq \sup_{|\vartheta_1-\vartheta_2|<\tilde{h}} |\Upsilon(\vartheta_1) - \Upsilon(\vartheta_2)|,$$

as $\Upsilon(\vartheta)$ is uniformly continuous, we know that this bound can be made arbitrarly small by choosing $\tilde{h}$ to be small. This bound can be made less than $\epsilon/3$ for any $\tilde{h} < \tilde{h}_1$.

$$\left|\frac{1}{t_N}\mathcal{L}(\vartheta_i) - \Upsilon(\vartheta_i)\right| \leq \max_{i=1,\cdots,I}\left|\frac{1}{t_N}\mathcal{L}(\vartheta_i) - \Upsilon(\vartheta_i)\right|.$$

Putting these results together, we have that for any $\tilde{h} < \tilde{h}_1$

$$\sup_{\vartheta\in\Theta}\left|\frac{1}{t_N}\mathcal{L}(\vartheta) - \Upsilon(\vartheta)\right| \leq \frac{1}{t_N}\sup_{|\vartheta_1-\vartheta_2|<\tilde{h}} |\mathcal{L}(\vartheta_1) - \mathcal{L}(\vartheta_2)| + \max_{i=1,\cdots,I}\left|\frac{1}{t_N}\mathcal{L}(\vartheta_i) - \Upsilon(\vartheta_i)\right| + \epsilon/3.$$

So for any $\tilde{h} < \tilde{h}_1$

$$\begin{aligned}
&P\left(\sup_{\vartheta\in\Theta}\left|\frac{1}{t_N}\mathcal{L}(\vartheta) - \Upsilon(\vartheta)\right| > \epsilon\right) \leq \\
&\leq P\left(\frac{1}{t_N}\sup_{|\vartheta_1-\vartheta_2|<\tilde{h}} |\mathcal{L}(\vartheta_1) - \mathcal{L}(\vartheta_2)| + \max_{i=1,\cdots,I}\left|\frac{1}{t_N}\mathcal{L}(\vartheta_i) - \Upsilon(\vartheta_i)\right| > 2\epsilon/3\right) \\
&\leq P\left(\frac{1}{t_N}\sup_{|\vartheta_1-\vartheta_2|<\tilde{h}} |\mathcal{L}(\vartheta_1) - \mathcal{L}(\vartheta_2)| > \epsilon/3\right) + P\left(\max_{i=1,\cdots,I}\left|\frac{1}{t_N}\mathcal{L}(\vartheta_i) - \Upsilon(\vartheta_i)\right| > \epsilon/3\right).
\end{aligned}$$

Now we show that we can take n sufficienlty large so that the two last probabilities can be made small. (3.7) implies there exists $N_1(\epsilon,\delta)$ so that for $n > N_1(\epsilon,\delta)$, $\tilde{h} < \tilde{h}_1$

$$P\left(\frac{1}{t_N}\sup_{|\vartheta_1-\vartheta_2|<\tilde{h}}|\mathcal{L}(\vartheta_1)-\mathcal{L}(\vartheta_2)| > \epsilon/3\right) < \delta/2.$$

For the $\tilde{h}$ considered so far, find the finite subcover

$$\{B_{\tilde{h}}(\vartheta_i) : i = 1,\cdots,I\}$$

so that

$$P\left(\max_{i=1,\cdots,I}\left|\frac{1}{t_N}\mathcal{L}(\vartheta_i)-\Upsilon(\vartheta_i)\right| > \epsilon/3\right) = P\left(\bigcup_{i=1}^{I}\left\{\left|\frac{1}{t_N}\mathcal{L}(\vartheta_i)-\Upsilon(\vartheta_i)\right| > \epsilon/3\right\}\right) \leq$$
$$\leq \sum_{i=1}^{I} P\left(\left|\frac{1}{t_N}\mathcal{L}(\vartheta_i)-\Upsilon(\vartheta_i)\right| > \epsilon/3\right).$$

We know that for each ϑ_i and for any $\delta > 0$ there exists $N_{2i}(\epsilon,\delta)$ so that for $n > N_{2i}(\epsilon,\delta)$

$$P\left(\left|\frac{1}{t_N}\mathcal{L}(\vartheta_i)-\Upsilon(\vartheta_i)\right| > \epsilon/3\right) < \frac{\delta}{2I}.$$

Let $N_2(\epsilon,\delta) = \max_{i=1,\cdots,I} N_{2i}(\epsilon,\delta)$, then for $n > N_2(\epsilon,\delta)$

$$\sum_{i=1}^{I} P\left(\left|\frac{1}{t_N}\mathcal{L}(\vartheta_i)-\Upsilon(\vartheta_i)\right| > \epsilon/3\right) < \delta/2.$$

Combining the results there exists an $N(\epsilon,\delta) = \max(N_1(\epsilon,\delta), N_2(\epsilon,\delta))$ so that for every $n > N(\epsilon,\delta)$

$$P\left(\sup_{\vartheta\in\Theta}\left|\frac{1}{t_N}\mathcal{L}(\vartheta)-\Upsilon(\vartheta)\right| > \epsilon\right) < \delta.$$

□

3.3 Alternative method: method of moments

Let assume log-returns are observed discretely with regular time spaces of fixed length Δ. For $i \in \mathbb{N}$ we denote the stationary increments of the GJR-COGARCH by

$$G_{i\Delta}^{(\Delta)} = G_{(i+1)\Delta} - G_{i\Delta}.$$

The following theorem (see [5]) shows how the parameters of the model might be estimated.

Theorem 3.3.1. *Let L a pure jump Lévy process with $\mathbb{E}(L_1) = 0$, $\mathbb{E}(L_1^2) = 1$, $\mathbb{E}(L_1^4) < \infty$ and Lévy measure such that $\int_{\mathbb{R}} x^3\nu_L(\mathrm{d}x) = 0$ and $S := \int_{\mathbb{R}} x^4\nu_L(\mathrm{d}x)$ is known. Assume $\Psi(2) < 0$ and let $(G_{i\Delta}^{(\Delta)})_{i\in\mathbb{N}}$ be the stationary increment process of the integrated*

GJR-COGARCH process with parameters β, η, φ and γ. Let μ, Γ, k and p constants such that

$$\mathbb{E}(G_{i\Delta}^{(\Delta)^2}) = \mu, \quad \mathbb{V}\mathrm{ar}(G_{i\Delta}^{(\Delta)^2}) = \Gamma$$

$$\rho(\tilde{h}) := \mathbb{C}\mathrm{orr}((G_{i\Delta}^{(\Delta)^2}), (G_{i\Delta+\tilde{h}}^{(\Delta)^2})) = k\mathrm{e}^{-\Delta\tilde{h}p}, \quad \tilde{h} \in \mathbb{N}.$$

Set $E := (1-\mathrm{e}^{-\Delta p})(\mathrm{e}^{\Delta p}-1)$,

$$M_1 := \Gamma - \frac{6k\Gamma}{E}(p\Delta - 1 + \mathrm{e}^{-\Delta p}) - 2\mu^2, \quad M_2 := 1 - \frac{\mu^2 S}{\Delta M_1}, \quad M_3 := \frac{\Delta k \Gamma p^2 S}{M_1 E}.$$

Then $M_1, M_2, M_3 > 0$. Further set

$$\tilde{\gamma}_{1,2} := \frac{-M_3 - 4pS}{2pS - M_2} \pm \frac{\sqrt{8pSM_2^2M_3 + 32p^2S^2M_2^2 + 2pSM_2M_3^2 - 8pSM_2^3}}{M_2(2pS - M_2)} \in \mathbb{R}.$$

Set additionally, for $i = 1, 2$, $H_i := \tilde{\gamma}_i^2 + 4\tilde{\gamma}_i - 4$ and

$$M_4^i := \frac{p^2}{\tilde{\gamma}_i^2} + 2\frac{\Delta k \Gamma p^3}{\tilde{\gamma}_i M_1 E H_i},$$

and choose the unique $\tilde{\gamma} \in \{\tilde{\gamma}_1, \tilde{\gamma}_2\}$ such that $M_4^i > 0$ and

$$\sqrt{M_4^i} H_i S \tilde{\gamma}_i = -M_2 \tilde{\gamma}_i^2 + M_3 \tilde{\gamma}_i + H_i S p.$$

Then $\tilde{\gamma} \in [1, 2)$ and the parameters β, η, φ and γ are uniquely determined by

$$\beta = \frac{p\mu}{\Delta}, \quad \varphi = -\frac{p}{\tilde{\gamma}} + \sqrt{\frac{p^2}{\tilde{\gamma}^2} + 2\frac{\Delta k \Gamma p^3}{\tilde{\gamma} M_1 E(\tilde{\gamma} + 4\tilde{\gamma} - 4)}},$$

$$\gamma = \sqrt{\tilde{\gamma} - 1}, \quad \eta = p + \varphi\tilde{\gamma}.$$

We are now ready to summarize the estimation algorithm.

(1) Calculate the moment estimator of μ

$$\hat{\mu}_n := \frac{1}{n}\sum_{i=1}^{n}(G_{i\Delta}^{(\Delta)})^2.$$

(2) For fixed $d \geq 2$ and $\tilde{h} = 0, \cdots, d$, calculate the estimator of the empirical autocovariances $\hat{\gamma}_n := (\hat{\gamma}_n(0), \hat{\gamma}_n(1), \cdots, \hat{\gamma}_n(d))^T$ as

$$\hat{\gamma}_n(\tilde{h}) := \frac{1}{n}\sum_{i=1}^{n-\tilde{h}}((G_{(i+\tilde{h})\Delta}^{(\Delta)})^2 - \hat{\mu}_n)((G_{i\Delta}^{(\Delta)})^2 - \hat{\mu}_n)$$

and compute the empirical autocorrelations $\hat{\rho}_n := (\hat{\gamma}_n(1)/\hat{\gamma}_n(0), \cdots, \hat{\gamma}_n(d)/\hat{\gamma}_n(0))^T$.

(3) Compute the least squares estimator for (p, k) as

$$(\hat{p}_n, \hat{k}_n) := \min_{(p,k)\in\mathbb{R}_+^2} \sum_{\tilde{h}=1}^{d} (\log(\hat{\rho}(\tilde{h})) - \log k + \Delta p\tilde{h})^2.$$

Define the mapping $H : \mathbb{R}_+^{d+2} \to \mathbb{R}$ by

$$H(\hat{\rho}_n, p, k) := \sum_{\tilde{h}=1}^{d} (\log \hat{\rho}_n(\tilde{h}) - \log k + \Delta\tilde{h}p)^2.$$

In order to obtain least squares estimators we compute partial derivatives.

$$\frac{\partial}{\partial k} H(\hat{\rho}_n, p, k) = -\frac{2}{k} \sum_{\tilde{h}=1}^{d} (\log \hat{\rho}_n(\tilde{h}) - \log k + \Delta\tilde{h}p) = 0 \quad \Leftrightarrow$$

$$\sum_{\tilde{h}=1}^{d} (\log \hat{\rho}_n(\tilde{h}) + \Delta\tilde{h}p) = d \log k.$$

Then, if $\overline{\log \hat{\rho}_n} := \frac{1}{d}\sum_{\tilde{h}=1}^{d} \log \hat{\rho}_n(\tilde{h})$,

$$k = \exp\left(\overline{\log \hat{\rho}_n} + \Delta p \frac{(d+1)}{2}\right). \tag{3.8}$$

$$\frac{\partial}{\partial p} H(\hat{\rho}_n, p, k) = -2\Delta \sum_{\tilde{h}=1}^{d} \tilde{h}(\log \hat{\rho}_n(\tilde{h}) - \log k + \Delta\tilde{h}p) = 0.$$

Using (3.8) we obtain

$$\Delta \frac{d(d+1)}{2} \overline{\log \hat{\rho}_n} - \Delta \sum_{\tilde{h}=1}^{d} \tilde{h} \log \hat{\rho}_n(\tilde{h}) = p\left(\frac{\Delta d(d+1)(2d+1)}{6} - \frac{\Delta^2 d(d+1)^2}{4}\right).$$

Hence

$$\hat{p}_n^* = \frac{-\sum_{\tilde{h}=1}^{d} (\log \hat{\rho}_n(\tilde{h}) - \overline{\log \hat{\rho}_n})(\tilde{h} - \frac{d+1}{2})}{\frac{d(d+1)(2d+1)}{6} - \frac{\Delta d(d+1)^2}{4}}.$$

For the stationary model the parameter p has to be strictly positive, but the unrestricted minimum of $H(\hat{\rho}_n, p, k)$ could give a negative estimate for p. As a remedy, we define the estimator of p as

$$\hat{p}_n := \max(\hat{p}_n^*, 0)$$

and we take $\hat{p}_n = 0$ as an indication that data are not stationary. Therefore

$$\hat{k}_n = \exp\left(\overline{\log \hat{\rho}_n} + \frac{\hat{p}_n^* \Delta(d+1)}{2}\right).$$

(4) Compute the estimator $\hat{\vartheta}_n^M := (\hat{\beta}_n^M, \hat{\eta}_n^M, \hat{\varphi}_n^M, \hat{\gamma}_n^M)$ where

$$\hat{\beta}_n^M = \hat{p}_n \hat{\mu}_n, \quad \hat{\varphi}_n^M = -\frac{\hat{p}_n}{\hat{\tilde{\gamma}}_n} + \sqrt{\frac{\hat{p}_n^2}{\hat{\tilde{\gamma}}_n^2} + 2\frac{\hat{k}_n \hat{\Gamma}_n \hat{p}_n^3}{\hat{\tilde{\gamma}}_n \hat{M}_{1n} \hat{E}_n (\hat{\tilde{\gamma}}_n^2 + 4\hat{\tilde{\gamma}}_n - 4)}},$$

$$\hat{\gamma}_n^M = \sqrt{\hat{\tilde{\gamma}}_n - 1}, \quad \hat{\eta}_n^M = \hat{p}_n + \hat{\varphi}_n^M \hat{\tilde{\gamma}}_n,$$

where $\hat{\Gamma}_n := \hat{\gamma}_n(0)$ and $\hat{E}_n$, $\hat{M}_{1n}$ and $\hat{\tilde{\gamma}}_n$ are the empirical versions of E, M_1 and $\tilde{\gamma}$ obtained by replacing p, k, μ, Γ with their estimators.

Remark 48. We could also base the least squares estimation on the autocovariance function, but it turned out that the estimators chosen are more accurate. This is because k is independent of β in contrast with $\tilde{k} := \mathbb{C}\mathrm{ov}((G_{i\Delta}^{(\Delta)})^2, (G_{(i+1)\Delta}^{(\Delta)})^2)\mathrm{e}^p$. Moreover

Theorem 3.3.2. *Under the same conditions as in Theorem 3.3.1 we obtain strong consistency for the estimator, i.e.*

$$\hat{\vartheta}_n^M \overset{\text{a.s.}}{\to} \vartheta_0.$$

Proof. The GJR-COGARCH volatility is a generalized Ornstein-Uhlenbeck process. The result of Fasen [17] makes σ^2 exponentially β-mixing. This implies that it is strongly mixing (or α-mixing) with an exponentially decreasing rate. Following [23] one can prove that $(G_{ir}^{(r)})_{i\in\mathbb{N}}$ is α-mixing with an exponentially decreasing rate as well. The volatility process is strictly stationary, then the return process is also strictly stationary and together with the strong mixing property this implies that $(G_{ir}^{(r)})_{i\in\mathbb{N}}$ is ergodic. And by Birkhoff's ergodic theorem we have strong consistency of the empirical moments and autocovariance function of $((G_{i\Delta}^{(\Delta)})^2)_{i\in\mathbb{N}}$. The parameter vector is a continuous function of the first two moments of the GJR-COGARCH and of p and k. Then, consistency of the moments implies consistency of the estimates for $(\beta, \eta, \varphi, \gamma)$. $\square$

Remark 49. In Theorem 3.3.1 $\rho(\tilde{h}) > 0$ for every $\tilde{h} \in \mathbb{N}$ and we obtain that M_1, M_2, M_3 and M_4^i are strictly positive. However the corresponding sample estimates could be negative. As we showed estimator is consistent, so, for large samples, the empirical estimates will be positive. Analogously for $\tilde{\gamma} \in [1, 2)$, the sample version might be less than 1 or greater than 2. However, consistency makes $\hat{\tilde{\gamma}}$ in $[1, 2)$ for large samples. All that was considered in the numerical algorithm developed.

3.4 Monte Carlo simulation study

In this section the PML estimation method is applied to simulated data sets and compared with the method of moments. We simulated 1000 GJR-COGARCH(1,1) data sets with $\Delta t_k = 0.5$ for every k between 0 and 10000. As driving Lévy process L we chose a variance gamma with $\tau = \sigma = 1$ and $\mu = 0$ and for the true GJR-COGARCH parameters we took $\beta = 0.04$, $\eta = 0.053$, $\varphi = 0.038$ and $\gamma = 0.6$. For an accurate simulation the grid for the Euler method is 500 times finer with respect to the final grid of the observation. For each simulated sample we estimated the parameters with both methods. The following tables highlight the results.

MM	$\hat{\beta}$	$\hat{\eta}$	$\hat{\varphi}$	$\hat{\gamma}$
mean	0.0641	0.0359	0.0267	0.4177
bias	0.0241	-0.0171	-0.0113	-0.1823
relative bias	0.6025	-0.3226	-0.2974	-0.3039
MSE	0.0059	0.0003	0.0002	0.00007

Table 3.1: Estimated mean, bias, relative bias, mean square error and mean absolute error of the MM estimators. The number of simulated samples is 1000 and the true values are $\beta = 0.04$, $\eta = 0.053$, $\varphi = 0.038$ and $\gamma = 0.6$.

PML	$\hat{\beta}$	$\hat{\eta}$	$\hat{\varphi}$	$\hat{\gamma}$
mean	0.05137	0.04888	0.03537	0.55428
bias	0.01137	-0.00412	-0.00263	-0.04572
relative bias	0.28425	-0.07774	-0.06921	-0.0762
MSE	0.00028	0.00001	0.00007	0.12535

Table 3.2: Estimated mean, bias, relative bias, mean square error and mean absolute error of the PML estimators. The number of simulated samples is 1000 and the true values are $\beta = 0.04$, $\eta = 0.053$, $\varphi = 0.038$ and $\gamma = 0.6$.

Results and kernel densities show that PML estimators for β, η and φ are more efficient than the corresponding MM estimators. We obtain the contrary for the estimator for γ: a few outliers make the PML estimator less accurate. Moreover PML estimates are always less biased. Therefore, it seems that the convergence rate to reach the consistency is faster if we estimate parameters with PML method.

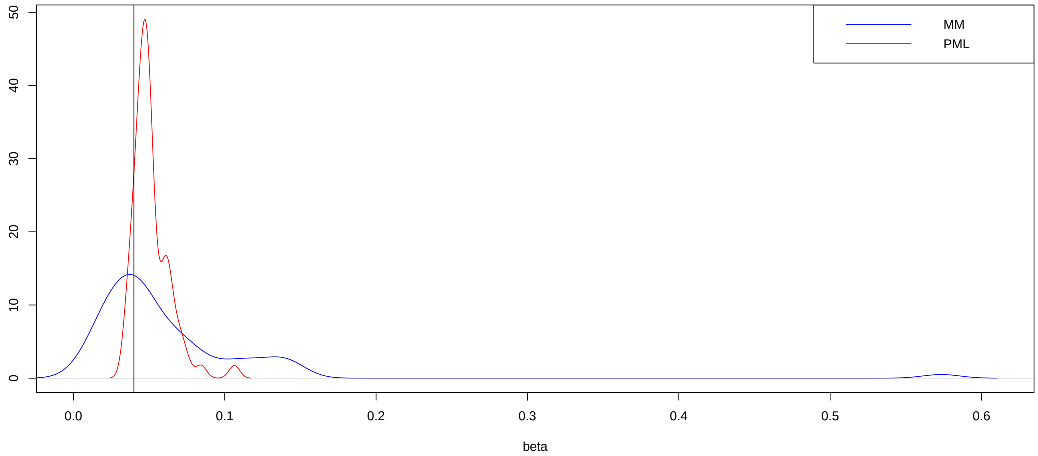

Figure 3.1: Kernel density of the PML estimator for β.

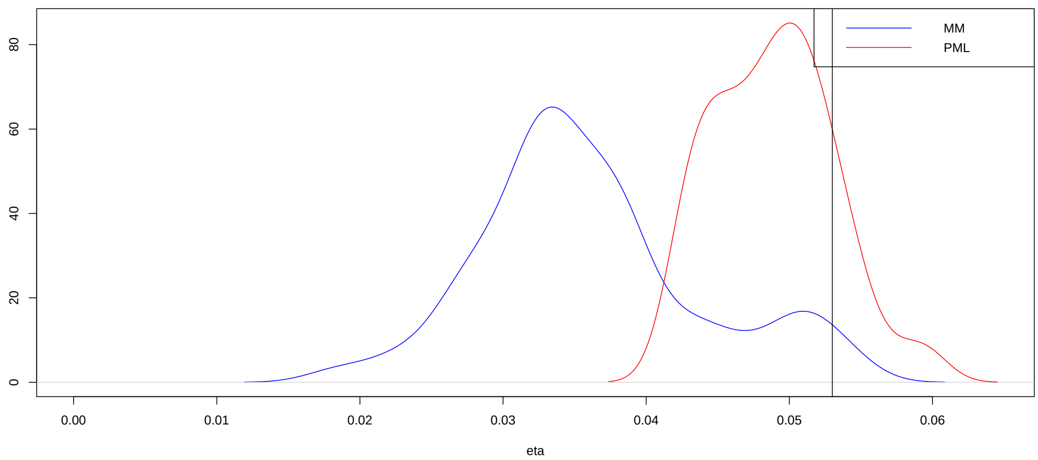

Figure 3.2: Kernel density of the PML estimator for η.

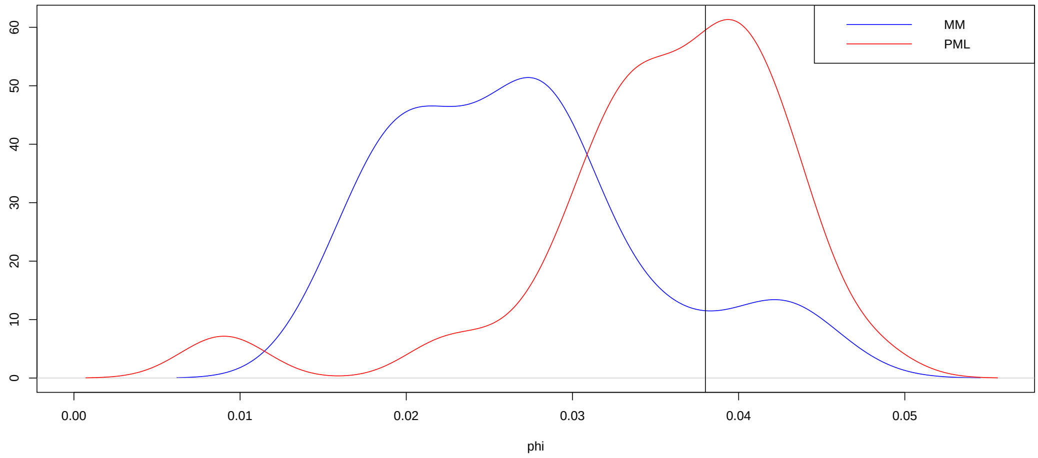

Figure 3.3: Kernel density of the PML estimator for φ.

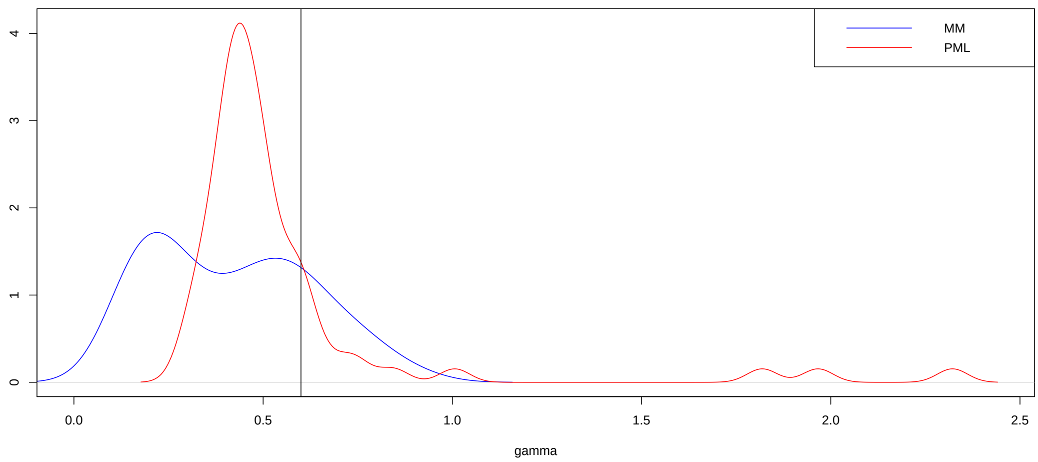

Figure 3.4: Kernel density of the PML estimator for γ.

Conclusion and outlook

Analysing high-frequency financial data and modelling the so-called stylized facts are nowadays based on continuous time models.
In this thesis, after an introduction to Lévy processes and related stochastic calculus, we addressed to stochastic volatility models in continuous time. First of all COGARCH process was introduced and its properties studied. Then, in order to capture the leverage effect asymmetric COGARCH models were analysed as well. Simulations of variance gamma COGARCH and GJR-COGARCH were proposed to study sample paths behavior. Thanks to these numerical analysis we could understand how the index, capturing the asymmetry, affects trajectories. Probabilistic properties have been studied for the continuous time APGARCH too. Finally, inferential procedures were proposed for estimating the GJR-COGARCH model parameters. In particular, we focused on a new version of the pseudo-maximum likelihood and its asymptotic properties. Results about consistency have been proved and confirmed by means of numerical studies. A Monte Carlo simulation study with 1000 GJR-COGARCH samples compared pseudo-maximum likelihood method with the method of moments and we obtained that PML estimates are less biased and more efficient than MM estimates. Asymptotic properties of the PML estimator are still subject of research; asymptotic distribution and rates of convergence are in progress. Moreover, the interest in multivariate continuous time models with stochastic volatility has increased in recent time, therefore an extension of the PML method to such multivariate processes are taken into consideration.

Bibliography

[1] Anh, V.V., Heyde, C.C. and Leonenko, N.N. Dynamic models of long-memory processes driven by Lévy noise. *Journal of Applied Probability*, 39: 730-747, 2002.

[2] Applebaum, D. *Lévy processes and Stochastic Calculus*. Cambridge studies in advanced mathematics; 116. Cambridge University Press, Cambridge, 2nd edition, 2009.

[3] Barndorff-Nielsen, O. and Shepard, N. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society*, 63: 167-241, 2001.

[4] Barndorff-Nielsen, O. and Shepard, N. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society*, 64: 253-280, 2002.

[5] Behme, A., Klüppelberg, C. and Mayr, K. Asymmetric cogarch processes. To appear in *Journal of Applied Probability*, 51A. Available at `http://arxiv.org/abs/1403.7068`

[6] Berkes, I. and Horváth, L. The efficiency of the estimators of the parameters in GARCH processes. *Annals of Statistics*, 32: 633-655, 2004.

[7] Bertoin, J. *Lévy Processes*. Cambridge tracts in mathematics; 121. Cambridge University Press, Cambridge, 1996.

[8] Bertoin J. and Yor M. Exponential functionals of Lévy processes. *Probability Surveys*, 2: 191-212, 2005.

[9] Bibbona, E. and Negri, I. *Higher moments and prediction based estimation for the COGARCH(1,1) model*. Submitted for publication. Available at `http://arxiv.org/abs/1401.7819`

[10] Black, F. and Scholes M. The Valuation of Option Contracts and a Test of Market Efficiency. *Journal of Finance*, 27: 399-417, 1972.

[11] Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31: 307-327, 1986.

[12] Brockwell, P., Chadraa, E. and Lindner, A. Continuous-Time GARCH Processes. *The annals of applied probability*, 16: 790-826, 2006.

[13] Ding, Z., Granger, C. and Engle, R. A long memory property of stock market returns and a new model. J*ournal of Empirical Finance*, 1: 83-106, 1993.

[14] Drost, F.C. and Weker, B.J.M. Closing the GARCH gap: Continuous time GARCH modelling. *Journal of Econometrics*, 74: 31-57, 1996.

[15] Engle, R. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50: 987-1007, 1982.

[16] Erickson, K.B. and Maller, R.A. Generalised Ornstein-Uhlenbeck processes and the convergence of Lévy integrals. Séminaire des Probabilités. *Lecture notes in Mathematics*. Springer, 1857: 70-94, 2005.

[17] Fasen, V. Asymptotic results for the sample autocovariance function and extremes of integrated generalized Ornstein-Uhlenbeck processes. *Bernoulli*, 16: 51-79, 2010.

[18] Fasen, V., Klüppelberg, C. and Lindner, A. Extremal behavior of stochastic volatility models. *Stochastic Finance*. Springer, 2006.

[19] Francq, C. and Zakoïan J.M. Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10: 605-637, 2004.

[20] Glosten, L., Jagannathan, R. and Runkle, D. On the relation between expected return on stocks. *Journal of Finance*, 48: 1779-1801, 1993.

[21] Hall, P. and Yao, Q. Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica*, 71: 285-317, 2003.

[22] Haug, S. and Czado C. An exponential continuous time GARCH process. *Journal of Applied Probability*, 44: 960-976, 2007.

[23] Haug, S., Klüppelberg, C., Lindner, A. and Zapp, M. Method of moment estimation in the COGARCH(1,1) model. *The Econometrics Journal*, 10: 320-341, 2007.

[24] Iacus, S. M. *Simulation and Inference for Stochastic Differential Equations*. Springer Series in Statistics. Springer, New York, 2008.

[25] Kallsen, J. and Vesenmayer, B. COGARCH as a continuous-time limit of GARCH(1,1). *Stochastic Processes and their Applications*, 19: 74-98, 2009.

[26] Kim, M. and Lee, S. On the maximum likelihood estimation for irregularly observed data from COGARCH(1,1) models. *REVSTAT - Statistical Journal*, 11: 135-168, 2013.

[27] Klüppelberg, C., Lindner, A and Maller R. A continuous time GARCH process driven by a Lévy process: stationarity and second order behavior. *Journal of Applied Probability*, 41: 601-622, 2004.

[28] Klüppelberg, C., Lindner, A and Maller R. *Continuous time volatility modelling: COGARCH versus Ornstein-Uhlenbeck models*. From Stochastic Calculus to Mathematical Finance. The Shiryaev Festschrift. Springer, 2006.

[29] Klüppelberg, C., Maller R. and Szimayer, A. The COGARCH: A Review, with News on Option Pricing and Statistical Inference. In *Surveys in Stochastic Processes. Proc. of the 33rd SPA Conference in Berlin*, pages 29-58. EMS Series of Congress Reports, EMS Publishing House, 2011.

[30] Kyprianou, A. *Introductory Lectures on Fluctuations of Lévy Processes with Applications*. Springer, Berlin, Heidelberg, 2006.

[31] Lee, O. A continuous time asymmetric power GARCH process driven by a Lévy process. *Journal of the Korean Data and Information Science Society*, 21: 1311-1317, 2010.

[32] Lee, O. V-uniform ergodicity of a continuous time asymmetric power GARCH(1,1) model. *Statistics and Probability Letters*, 82: 812-817, 2012.

[33] Madan, D.B., Carr, P. and Chang, E. The Variance Gamma Process and Option Pricing. *European Finance Review*, 2: 79-105, 1998.

[34] Madan, D.B. and Seneta, E. The Variance Gamma (V.G.) model for share market returns. *Journal of Business*, 63: 511-524, 1990.

[35] Maller, R. and Szimayer, A. Finite approximation schemes for Lévy processes, and their application to optimal stipping problems. *Stochastic Processes and their Applications*, 117: 1422-1447, 2007.

[36] Maller, R., Müller, G. and Szimayer, A. GARCH modelling in continuous time for irregularly spaced time series data. *Bernoulli*, 14: 519-542, 2008.

[37] Mandelbrot, B. The Variation of Certain Speculative Prices. *The Journal of Business*, 36: 394-419, 1963.

[38] Mayr, K. *Der asymmetrische COGARCH: seine Definition, Approximation und Schätzung*. Masterarbeit, Technische Universität München, 2013.

[39] McKenzie, M. and Mitchell, H. Generalized asymmetric power ARCH modelling of extremes rate volatility. *Applied Financial Economics*, 12: 555-564, 2002.

[40] Mikosh, T. *Non-Life Insurance Mathematics. An Introduction with Stochastic Processes*; 13. Springer-Verlag, 2nd edition, 2006.

[41] Müller, G. MCMC estimation of the COGARCH(1,1) model. *Journal of Financial Econometrics*, 8: 481-510, 2010.

[42] Müller, G., Durand, R., Maller, R. and Kluppelberg, C. Analysis of stock market volatility by continuous-time garch models. In *Stock Market Volatility*, pages 31-50. Chapman Hall/Taylor and Francis, London, 2009.

[43] Nelson, D.B. Arch models as diffusion approximations. *Journal of Econometrics*, 45: 7-38, 1990.

[44] Nelson, D.B. Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 59: 347-370, 1991.

[45] Penzer, J., Wang, M. and Yao, Q. Approximating volatilities by asymmetric power GARCH functions. *Australian & New Zealand Journal of Statistics*, 51: 201-225, 2009.

[46] Protter, P. *Stochastic Integration and Differential Equations.* Number 21 in Stochastic Modelling and Applied Probability. Springer, 2nd edition, 2005.

[47] Rabemananjara, R. and Zakoian, J. Threshold ARCH models and asymmetries in volatility. *Journal of Applied Econometrics*, 8: 31-49, 1993.

[48] Sato, K. *Lévy Processes and Infinitely Divisible Distributions.* Cambridge studies in advanced mathematics; 68. Cambridge University Press, Cambridge, 1999.

[49] Schoutens, W. *Lévy Processes in Finance: Pricing Financial Derivatives.* John Wiley & Sohns Ltd, Chichester, 2003.

[50] Sørensen, M. Prediction-based estimating functions: review and new developments. *Brazilian Journal of Probability and Statistics*, 25: 362-391, 2011.

[51] Stelzer, R. Multivariate COGARCH(1,1) processes. *Bernoulli*, 16: 80-115, 2010.

[52] Straumann, D. *Estimation in Conditionally Heteroscedastic Time Series Models.* Lecture Notes in Statistics, 181. Springer-Verlag, Berlin, Heidelberg, 2005.

[53] Zakoian, J. Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18: 931-955, 1994.