Discrete time approximation of a COGARCH(p,q) model and its estimation

Stefano, M. Iacus*[1], Lorenzo Mercuri †[2], and Edit Rroji ‡[3]

[1]*Department of Economics, Management and Quantitative Methods, University of Milan, CREST Japan Science and Technology Agency, Tokyo Japan*
[2]*Department of Economics Management and Quantitative Methods, University of Milan, CREST Japan Science and Technology Agency, Tokyo Japan*
[3]*Department of Statistics and Quantitative Methods, University of Milano-Bicocca*

April 7, 2018

Abstract

In this paper, we construct a sequence of discrete time stochastic processes that converges in the Skorokhod metric to a COGARCH(p,q) model. The result is useful for the estimation of the COGARCH(p,q) on irregularly spaced time series data. The proposed estimation procedure is based on the maximization of a pseudo log-likelihood function and is implemented in the `yuima` package.

Keywords: COGARCH(p,q) Process, Skorokhod Distance, Pseudo Log-Likelihood Estimation

1 Introduction

The COGARCH(1,1) model has been introduced by Klüppelberg et al. (2004) as a continuous time counterpart of the GARCH(1,1). The continuous time model preserves the main features of the GARCH since, even in the COGARCH model, there is only one source of randomness for returns and variance dynamics. For the COGARCH(1,1) estimation different methods have been proposed. For instance, Haug et al. (2007) develop a procedure based on the matching of theoretical and empirical moments. Maller et al. (2008) use an approximation scheme for obtaining estimates of parameters through the maximization of a pseudo log-likelihood function while Müller (2010) develops a Markov Chain Monte Carlo estimation.

The COGARCH(1,1) model has been generalized to the higher order case by Chadraa (2009) and Brockwell et al. (2006). In Chadraa (2009), a procedure for the estimation of COGARCH(p,q) parameters by matching empirical and theoretical moments is also proposed. To

*Electronic address: `stefano.iacus@unimi.it`
†Electronic address: `lorenzo.mercuri@unimi.it`
‡Electronic address: `edit.rroji@unimib.it`

the best of our knowledge, the latter is the only available approach for the estimation of COGARCH(p,q) parameters.
In this paper, we construct a sequence of discrete time stochastic processes that converges in probability and in the Skorokhod metric to a COGARCH(p,q) model. Our results generalize the approach in Maller et al. (2008). Results derived for a COGARCH(p,q) model in Chadraa (2009) are used in this paper for extending the estimation procedure based on the maximization of the pseudo log-likelihood function. This estimation method is then implemented in the `yuima` package available on CRAN (See Brouste et al., 2014; Iacus et al., 2017; Iacus and Yoshida, 2018, for more details on `yuima` package).
The outline of the paper is as follows. In Section 2 we review some useful properties needed in Section 3 where we introduce a sequence of discrete time processes and prove its convergence to the COGARCH(p,q) model using the Skorokhod metric. Section 4 generalizes the maximum pseudo log-likelihood procedure proposed in Maller et al. (2008). In Section 5 we present some numerical examples about the estimation of a COGARCH(2,2) model using simulated and real data. Section 6 concludes the paper.

2 Preliminaries

In this section we review useful results.

Definition 1 *A sequence of random vector valued functions $Q_{n,\theta}$ with argument θ is uniformly convergent in probability to Q_θ if and only if:*

$$\sup_{\theta\in\Theta}\|Q_{n,\theta}-Q_\theta\| \xrightarrow{P} 0, \tag{1}$$

where $\|\cdot\|$ is the Euclidean norm.

Remark 2 *Definition 1 holds also for any vector norm $\|\cdot\|_A$ induced by an invertible matrix A, i.e.*[1] *$\|x\|_A=\|Ax\|$ where A is a non singular matrix.*

Definition 3 *Let $\|\cdot\|$ be a norm on $\mathcal{R}^n$, the induced matrix norm $\|\cdot\|_M$ as a function from $\mathcal{R}^{n\times n}$ to $\mathcal{R}_+$ defined as:*

$$\|A\|_M := \sup_{\|x\|\neq 0}\frac{\|Ax\|}{\|x\|} = \sup_{\|z\|=1}\|Az\|$$

where $A\in\mathcal{R}^{n\times n}$.

Theorem 4 *The induced matrix norm $\|\cdot\|_M$ satisfies the following properties (see Meyer, 2000):*
1) $\|Ax\| \leq \|A\|_M\|x\|$
2) $\|\alpha A\|_M \leq |\alpha|\,\|A\|_M$
3) $\|A+B\|_M \leq \|A\|_M+\|B\|_M$
4) $\|AB\|_M \leq \|A\|_M\|B\|_M$,
where $A\in\mathcal{R}^{q\times q}$, $B\in\mathcal{R}^{q\times q}$ and α is a scalar.

Any induced matrix norm satisfies the following inequality (see Chapter 7 page 117 Serre, 2002):

$$\left\|\frac{e^{At}-I}{t}-A\right\|_M \leq \frac{e^{\|At\|_M}-1-\|At\|_M}{|t|},\ t\in\mathcal{R}, \tag{2}$$

[1] In the induced vector norm we have that $\|\cdot\|$ can be a generic norm but in this paper if not otherwise stated explicitly $\|\cdot\|$ refers to the Euclidean norm.

where the matrix exponential e^A is defined as a power series:

$$e^A = \sum_{k=0}^{+\infty} \frac{A^k}{k!}.$$

Definition 5 *Let $\|\cdot\|_M$ be the induced matrix norm by the norm $\|\cdot\|$ defined on $\mathcal{R}^n$, the logarithmic norm $\mu(A)$ (see Ström, 1975, for its properties) is:*

$$\mu(A) := \lim_{t \to 0^+} \frac{\|I + At\|_M - 1}{t}.$$

Theorem 6 *For the logarithmic norm, from Lemma 1.c in Ström (1975) the following inequalities hold:*

$$\left\| e^{At} \right\|_M \leq e^{\mu(A)t} \leq e^{\|A\|_M t}$$

with $t > 0$.

Theorem 7 *Let (a_n) and (b_n) be two sequences of non negative numbers for $n = 1, \ldots, N$ such that $a_n \geq 1$ and $b_n \geq 0$. Define the sequence y_n such that*

$$y_n \leq a_n y_{n-1} + b_n \tag{3}$$

with $y_0 \geq 0$, then:

$$\max_{n=1,\ldots,N} y_n \leq \left[\prod_{k=0}^{N-1} a_{N-k} \right] y_0 + b_N + \sum_{j=1}^{N-1} \left[\prod_{h=1}^{j} a_{N+1-h} \right] b_{N-j}. \tag{4}$$

Moreover, if the relation in (3) *is satisfied with equality the sequence is non decreasing and*

$$\max_{n=1,\ldots,N} y_n = y_N \tag{5}$$

where

$$y_N = \left[\prod_{k=0}^{N-1} a_{N-k} \right] y_0 + b_N + \sum_{j=1}^{N-1} \left[\prod_{h=1}^{j} a_{N+1-h} \right] b_{N-j}. \tag{6}$$

Proof. The relations in (4) and (5) can be derived in a straightforward way using the induction principle. ■

3 Main result

On a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$ we consider a pure jump Lévy process $L = (L_t)_{t\geq 0}$ with characteristic triplet $(\gamma, 0, \Pi)$ and $L_0 = 0$. Its characteristic function reads:

$$E\left(e^{i\theta L_t}\right) = \exp\left(it\gamma\theta + t \int_{R\setminus\{0\}} \left(e^{it\theta x} - 1 - i\theta x \mathbf{1}_{|x|\leq 1} \right) \Pi(\mathrm{d}x) \right), \quad t \geq 0$$

where $\mathbf{1}$ is the indicator function; γ is a constant depending on the truncation at zero and we consider the standard truncation $\mathbf{1}_{|x|\leq 1}$ as in Maller et al. (2008); $\Pi(x)$ is the Lévy measure of the process L defined on the Borel subsets of $R\setminus\{0\}$. We assume that $E[L_1] = 0$ and $E[L_1] = 1$ (see Sato, 1999, for more details).

The COGARCH(p,q) model G_t, introduced in Brockwell et al. (2006), is defined through the following equations:

$$\begin{aligned}
\mathrm{d}G_t &= \sqrt{V_t}\mathrm{d}L_t \\
V_t &= a_0 + \mathbf{a}^\top Y_{t-} \\
\mathrm{d}Y_t &= BY_{t-}\mathrm{d}t + \mathbf{e}\left(a_0 + \mathbf{a}^\top Y_{t-}\right)\mathrm{d}\left[L,L\right]^d
\end{aligned} \tag{7}$$

where $B \in \mathcal{R}^{q\times q}$ is matrix of the form:

$$B = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -b_q & -b_{q-1} & \dots & \dots & -b_1 \end{bmatrix}$$

and **a** and **e** are vectors defined as:

$$\begin{aligned}
\mathbf{a} &= \left[a_1, \dots, a_p, a_{p+1}, \dots, a_q\right]^\top \\
\mathbf{e} &= \left[0, \dots, 0, 1\right]^\top
\end{aligned}$$

for $a_{p+1} = \ldots = a_q = 0$. As remarked in Brockwell et al. (2006), the state process Y_t in a COGARCH(p,q) model satisfies:

$$Y_t = J_{s,t}Y_s + K_{s,t} \quad s \leq t \tag{8}$$

where $J_{s,t} \in \mathcal{R}^{q\times q}$ is a random matrix and $K_{s,t} \in \mathcal{R}^{q\times 1}$ is a random vector for each pair (s,t). In particular, if the driving noise is a Compound Poisson process then the matrices and vectors in the state process in (8) have an analytical form as in Theorem 3.5 in Brockwell et al. (2006). Indeed, if we define τ_k with $k = 1, 2, 3, \ldots$ as the time of the k-th jump of the driving Compound Poisson process L_t and $Z_k := \Delta L^2_{\tau_k} = \left(L_{\tau_k} - L_{\tau_k-}\right)^2$ the square of the jump at time τ_k, then for any $t \in [\tau_k, \tau_{k+1})$ the state process Y_t is given by

$$Y_t = e^{B(t-\tau_k)}Y_{\tau_k}$$

where Y_{τ_k}, the state process at jump time τ_k, satisfies the following recurrence equation:

$$Y_{\tau_k} = C_k Y_{\tau_{k-1}} + D_k \tag{9}$$

with $\tau_0 = 0$ and Y_{τ_0} coincides with the starting condition in the definition of the COGARCH(p,q) process in (7). The random matrices C_k and the random vectors D_k in (9) are respectively:

$$\begin{aligned}
C_k &= \left(I + Z_k\mathbf{e}\mathbf{a}^\top\right)e^{B\Delta\tau_k} \\
D_k &= a_0 Z_k \mathbf{e}.
\end{aligned} \tag{10}$$

3.1 Approximation of the COGARCH(p,q) model

As in Maller et al. (2008), we construct a sequence of piecewise constant processes $G_n = \left(G_{t,n}\right)_{t\geq 0, n=1,2,\ldots}$ that converges to the COGARCH(p,q) model $G = \left(G_t\right)_{t\geq 0}$ in (7) by means of the Skorokhod distance (see Billingsley, 1968, for more details).

Unlike the COGARCH(1,1) case investigated in Maller et al. (2008), the family of processes G_n here introduced is not constructed from a GARCH(p,q) process. However, we show that the state space process defined on a discrete grid has a similar structure of a GARCH(p,q) in the sense that it depends on past squared increments of the process G_n.
To prove the convergence of G_n to G we perform the same steps as in Maller et al. (2008) except for the approximation of the variance process V_t. For the construction of the approximating process we need first to introduce an i.i.d. sequence of innovation terms computed using the first jump approximation scheme introduced in Szimayer and Maller (2007) for the underlying Lévy process L in the COGARCH(p,q) model definition.
Given a discrete grid on a finite time interval $[0, T]$, the Lévy process is approximated pathwise by considering for each subinterval in the grid, only the first jump of the process greater than some fixed minimal jump size and shifting this jump to the next point in the grid.
If both maximum size of subintervals and minimal jump size converge to zero, the first jump approximations defined on this grid converge in probability, using the Skorokhod metric, to the Lévy process (see Szimayer and Maller, 2007, for details on convergence rates of the grid and the jump size).
The discretization works as follows. For each $n \geq 0$, we consider a sequence of natural numbers N_n such that $\lim_{n\to+\infty} N_n = +\infty$ and obtain a partition of the interval $[0, T]$ defined as:

$$0 = t_{0,n} \leq t_{1,n} \leq \ldots \leq t_{N_n,n} = T. \tag{11}$$

We define Δt_n as follows:

$$\Delta t_n := \max_{i=1,\ldots,N_n} \Delta t_{i,n} \underset{n\to+\infty}{\rightarrow} 0$$

where $\Delta t_{i,n} := t_{i,n} - t_{i-1,n}$.
Using the partition in (11), we introduce the process $G_{i,n}$ as follows:

$$G_{i,n} = G_{i-1,n} + \sqrt{V_{i,n}\Delta t_{i,n}}\epsilon_{i,n}, \tag{12}$$

where the innovation $\epsilon_{i,n}$ is constructed using the first jump approximation method developed in Szimayer and Maller (2007) for a pure jump Lévy process. The definition of $V_{i,n}$ also involves $\epsilon_{i,n}$ and we will explicit it just after having explained how to construct the innovations.
Let $(m_n)_{n\in N}$ be a sequence of strictly positive real numbers satisfying the conditions:

$$m_n \leq 1 \ \forall n \geq 0 \ \text{ and } \lim_{n\to+\infty} m_n = 0.$$

We require the Lévy measure Π to satisfy the following property:

$$\lim_{n\to+\infty} \Delta t_n \bar{\Pi}^2 (m_n) = 0$$

where $\bar{\Pi}(x) := \int_{|y|>x} \Pi(\mathrm{d}x)$.
We define the stopping time process as:

$$\tau_{i,n} := \inf \{t \in [t_{i-1,n}, t_{i,n}) : |\Delta L_t| > m_n\} \tag{13}$$

if we have a jump with a size greater than m_n in the interval $[t_{i-1,n}, t_{i,n})$ otherwise $\tau_{i,n} = +\infty$.
Using the stopping time $\tau_{i,n}$, we construct a sequence of independent random variables $\left(\mathbf{1}_{\tau_{i,n}<+\infty}\Delta L_{\tau_{i,n}}\right)_{i=1,\ldots,N_n}$ with distribution:

$$F_{i,n}(\mathrm{d}x) = \frac{\Pi(\mathrm{d}x)\,\mathbf{1}_{|x|>m_n}}{\bar{\Pi}(m_n)}\left(1 - e^{\Delta t_{i,n}\bar{\Pi}(m_n)}\right)$$

for $x \neq 0$ and mass $e^{-\Delta t_{i,n}\bar{\Pi}(m_n)}$ at $x = 0$.
Under the requirement $E\left[L_1^2\right] < +\infty$, we introduce the innovation $\epsilon_{i,n}$ defined as:

$$\epsilon_{i,n} = \frac{\mathbf{1}_{\tau_{i,n}<+\infty}\Delta L_{\tau_{i,n}} - v_{i,n}}{\eta_{i,n}} \tag{14}$$

where $v_{i,n}$ and $\eta_{i,n}$ are respectively the mean and the variance of the random variables $\mathbf{1}_{\tau_{i,n}<+\infty}\Delta L_{\tau_{i,n}}$ and, by construction, the innovation $\epsilon_{i,n}$ has zero mean and unitary variance.
The variance process V_t in (7) is approximated by the process $V_{i,n}$ defined as:

$$V_{i,n} = a_0 + \mathbf{a}^\top Y_{i-1,n} \tag{15}$$

where $Y_{i,n}$ is given by:

$$Y_{i,n} = C_{i,n} Y_{i-1,n} + D_{i,n}, \tag{16}$$

with coefficients:

$$\begin{aligned} C_{i,n} &= \left(I + \epsilon_{i,n}^2 \Delta t_{i,n} \mathbf{e}\mathbf{a}^\top\right) e^{B\Delta t_{i,n}} \\ D_{i,n} &= a_0 \epsilon_{i,n}^2 \Delta t_{i,n} \mathbf{e} \end{aligned} \tag{17}$$

and initial condition $Y_{0,n} = Y_0$ for each n.
Combining (16) with (17) we get:

$$Y_{i,n} = \left(I + \epsilon_{i,n}^2 \Delta t_{i,n} \mathbf{e}\mathbf{a}^\top\right) e^{B\Delta t_{i,n}} Y_{i-1,n} + a_0 \epsilon_{i,n}^2 \Delta t_{i,n} \mathbf{e}. \tag{18}$$

Since from (12), we have $\epsilon_{i,n}^2 \Delta t_{i,n} = \frac{\left(G_{i,n} - G_{i-1,n}\right)^2}{V_{i,n}}$ and the recursion in (18):

$$Y_{i,n} = \left[I + \frac{\left(G_{i,n} - G_{i-1,n}\right)^2}{V_{i,n}} \mathbf{e}\mathbf{a}^\top\right] e^{B\Delta t_{i,n}} Y_{i-1,n} + a_0 \frac{\left(G_{i,n} - G_{i-1,n}\right)^2}{V_{i,n}} \mathbf{e}. \tag{19}$$

The representation in (19) is the main block of the pseudo log-likelihood estimation procedure discussed in Section 4.
Observe that we start from the first jump approximation of process L_t that drives G_t and get subsequently the approximation of the variance process. Since $[L, L]^d$ is itself a Lévy process it is possible to obtain a first jump approximation of the variance process by applying Theorem 3.6 in Stelzer (2009). However it is not straightforward to get the discretized version of G_t. Indeed we are interested in the couple $(G_{i,n}, V_{i,n})$ and we show that it converges to the couple (G_t, V_t) in the Skorokhod distance in our setup.
The Skorokhod distance between two processes U, V defined on $D^d[0, T]$, i.e. the space of càdlàg $\mathcal{R}^d$ valued stochastic processes on $[0, T]$, is:

$$\rho(U, V) := \inf_{\lambda \in \Lambda} \left\{ \sup_{0 \le t \le T} \|U_t - V_{\lambda_t}\| + \sup_{0 \le t \le T} |\lambda_t - t| \right\}$$

where Λ is a set of strictly increasing continuous functions with $\lambda_0 = 0$ and $\lambda_T = T$.

Theorem 8 *Let $N_{t,n}$ be a counting process defined as:*

$$N_{t,n} := \#\left\{i \in \mathcal{N} : \ \tau_{i,n}^\star \le t\right\}$$

where $t \le T$, $N_{n,0} = 0$, $N_{t,n} = N_n$ and $\tau_{i,n}^\star = \min\left\{\tau_{i,n}, t_{i,n}\right\}$ with $\tau_{i,n}$ as in (13) and $t_{i,n}$ as in (11).
Let $L = (L_t)_{t \ge 0}$ be a Lévy process with finite variation and $E\left[L_1^2\right] < +\infty$. The following results hold:

1. *The positive process*

$$H_{t,n} := \prod_{k=1}^{N_{t,n}} C^{\star}_{k,n}$$

where $C^{\star}_{k,n} := \left(1 + \epsilon^2_{k,n}\Delta t_{k,n} \left\|\mathbf{ea}^\top\right\|_M\right) e^{\|B\|_M \Delta t_{i,n}}$ *and the positive process*

$$\tilde{H}_{t,n} := \prod_{k=1}^{N_{t,n}} \tilde{C}_{k,n}$$

with $\tilde{C}_{k,n} := \left(1 + \mathbf{1}_{\tau_{k,n}<+\infty}\Delta L^2_{\tau_{k,n}} \left\|\mathbf{ea}^\top\right\|_M\right) e^{\|B\|_M \Delta t_{i,n}}$ *satisfy:*

$$\sup_{t\in[0,T]} \left|H_{t,n} - \tilde{H}_{t,n}\right| \xrightarrow{p} 0 \quad \textit{when} \quad n \to +\infty. \tag{20}$$

2. *For each fixed* n*,* $\tilde{H}_{t,n}$ *is a non decreasing strictly positive process such that* $\forall t \in [0,T]$*:*

$$\tilde{H}_{t,n} \leq H_T := e^{\|B\|_M T + \sum_{0\leq s\leq T} \ln\left(1+\Delta L^2_s \left\|\mathbf{ea}^\top\right\|_M\right)}. \tag{21}$$

Moreover, $\tilde{H}_{t,n}$ *converges uniformly in probability (hereafter ucp) on a compact interval* $[0,T]$ *to the process* $H_t := e^{\|B\|_M t + \sum_{0\leq s\leq t} \ln\left(1+\Delta L^2_s \left\|\mathbf{ea}^\top\right\|_M\right)}$.

3. *The process* $H_{t,n}$ *converges also ucp to the process* H_t*.*

Proof. We start from point 2. $\tilde{H}_{t,n}$ is a non decreasing strictly positive process since it is a product of terms $\tilde{C}_{k,n} \geq 1$ a.s. and if $s > t$ then $\tilde{H}_{n,s}$ has at least the same terms as $\tilde{H}_{t,n}$. Since:

$$\Delta L^2_s = \Delta L^2_s \mathbf{1}_{|\Delta L_s|\geq m_n} + \Delta L^2_s \mathbf{1}_{|\Delta L_s|<m_n},$$

we have:

$$\tilde{H}_{t,n} = e^{\|B\|_M T + \sum_{k=1}^{N_n} \ln\left(1+\mathbf{1}_{\tau_{k,n}<+\infty}\Delta L^2_{\tau_{k,n}} \left\|\mathbf{ea}^\top\right\|_M\right)} \leq e^{\|B\|_M T + \sum_{0\leq s\leq T} \ln\left(1+\Delta L^2_s \left\|\mathbf{ea}^\top\right\|_M\right)}.$$

For any $0 < \delta < 1$, H_t can be written as:

$$H_t = e^{\left[\|B\|_M - \ln(\delta)\right]t - X_t},$$

where $X_t = -t\ln(\delta) - \sum_{0\leq s\leq T} \ln\left(1 + \Delta L^2_s \frac{\lambda}{\delta}\right)$ with $\lambda := \left\|\mathbf{ea}^\top\right\|_M \delta > 0$.
As stated in Proposition 3.1 page 606 in Klüppelberg et al. (2004), X_t is a spectrally negative process of bounded variation on any compact interval. Consequently, process $\ln(H_t)$ is also a Lévy process with bounded variation that implies that H_t is also bounded on a compact interval $[0,T]$. Applying the second part of Proposition 5.1 page 533 in Maller et al. (2008), on the compact interval $[0,T]$ we have:

$$\sup_{t\in[0,T]} \left|\ln\left(\tilde{H}_{t,n}\right) - \ln(H_t)\right| \xrightarrow{p} 0, \quad \text{when} \quad n \to +\infty$$

and

$$\sup_{t\in[0,T]} \left|\tilde{H}_{t,n} - H_t\right| \xrightarrow{p} 0, \quad \text{when} \quad n \to +\infty.$$

In order to prove point 1, we consider:

$$\begin{aligned}
\sup_{t\in[0,T]}\left|H_{t,n}-\tilde{H}_{t,n}\right| &= \sup_{t\in[0,T]}\left|\prod_{k=1}^{N_{t,n}} C^{\star}_{k,n}-\prod_{k=1}^{N_{t,n}}\tilde{C}_{k,n}\right| \\
&\le e^{\|B\|_M T}\sup_{t\in[0,T]}\left|\prod_{k=1}^{N_{t,n}}\left(1+\epsilon^2_{k,n}\Delta t_{k,n}\left\|\mathbf{ea}^\top\right\|_M\right)-\prod_{k=1}^{N_{t,n}}\left(1+\mathbf{1}_{\tau_{k,n}<+\infty}\Delta L^2_{\tau_{k,n}}\left\|\mathbf{ea}^\top\right\|_M\right)\right| \\
&= e^{\|B\|_M T}\sup_{t\in[0,T]}\left|e^{\sum_{k=1}^{N_{t,n}}\ln\left(1+\epsilon^2_{k,n}\Delta t_{k,n}\left\|\mathbf{ea}^\top\right\|_M\right)}-e^{\sum_{k=1}^{N_{t,n}}\ln\left(1+\mathbf{1}_{\tau_{k,n}<+\infty}\Delta L^2_{\tau_{k,n}}\left\|\mathbf{ea}^\top\right\|_M\right)}\right| .
\end{aligned}$$

Using the inequality $\ln(1+x)-\ln(1+y)\le x-y$ with $x>-1$ and $y>0$, we obtain that:

$$\begin{aligned}
&\sup_{t\in[0,T]}\left|\sum_{k=1}^{N_{t,n}}\ln\left(1+\epsilon^2_{k,n}\Delta t_{k,n}\left\|\mathbf{ea}^\top\right\|_M\right)-\sum_{k=1}^{N_{t,n}}\ln\left(1+\mathbf{1}_{\tau_{k,n}<+\infty}\Delta L^2_{\tau_{k,n}}\left\|\mathbf{ea}^\top\right\|_M\right)\right| \\
&\le \sup_{t\in[0,T]}\left|\sum_{k=1}^{N_{t,n}}\left(\epsilon^2_{k,n}\Delta t_{k,n}\left\|\mathbf{ea}^\top\right\|_M-\mathbf{1}_{\tau_{k,n}<+\infty}\Delta L^2_{\tau_{k,n}}\left\|\mathbf{ea}^\top\right\|_M\right)\right| \\
&\le \left\|\mathbf{ea}^\top\right\|_M\sup_{t\in[0,T]}\sum_{k=1}^{N_{t,n}}\left|\left(\epsilon^2_{k,n}\Delta t_{k,n}-\mathbf{1}_{\tau_{k,n}<+\infty}\Delta L^2_{\tau_{k,n}}\right)\right| .
\end{aligned}$$

As shown in Maller et al. (2008), we have:

$$\sup_{t\in[0,T]}\sum_{k=1}^{N_{t,n}}\left|\left(\epsilon^2_{k,n}\Delta t_{k,n}-\mathbf{1}_{\tau_{k,n}<+\infty}\Delta L^2_{\tau_{k,n}}\right)\right|\xrightarrow{p}0,$$

that implies

$$\sup_{t\in[0,T]}\left|\sum_{k=1}^{N_{t,n}}\ln\left(1+\epsilon^2_{k,n}\Delta t_{k,n}\left\|\mathbf{ea}^\top\right\|_M\right)-\sum_{k=1}^{N_{t,n}}\left(1+\mathbf{1}_{\tau_{k,n}<+\infty}\Delta L^2_{\tau_{k,n}}\left\|\mathbf{ea}^\top\right\|_M\right)\right|\xrightarrow{p}0\,. \tag{22}$$

Since on compact intervals, $\tilde{H}_{t,n}$ is a bounded process and converges ucp to the bounded process H_t, using the result in (22), we have:

$$\sup_{t\in[0,T]}\left|e^{\sum_{k=1}^{N_{t,n}}\ln\left(1+\epsilon^2_{k,n}\Delta t_{k,n}\left\|\mathbf{ea}^\top\right\|_M\right)}-e^{\sum_{k=1}^{N_{t,n}}\left(1+\mathbf{1}_{\tau_{k,n}<+\infty}\Delta L^2_{\tau_{k,n}}\left\|\mathbf{ea}^\top\right\|_M\right)}\right|\xrightarrow{p}0\,.$$

Proof of point 3 is based on the following inequality:

$$\sup_{t\in[0,T]}\left|H_{t,n}-H_t\right|\le\sup_{t\in[0,T]}\left|H_{t,n}-\tilde{H}_{t,n}\right|+\sup_{t\in[0,T]}\left|H_t-\tilde{H}_{t,n}\right| . \tag{23}$$

Both terms of the right hand side in (23) go to zero in probability uniformly on interval $[0,T]$ when $n\to+\infty$, as shown respectively in point 1 and point 2. ■

Now, using the discrete processes $G_{i,n}$, $V_{i,n}$ and $Y_{i,n}$ respectively in (12), (15) and (16), we introduce the continuous-time càdlàg piecewise constant processes $G_{t,n}:=G_{i,n}$, $V_{t,n}:=V_{i,n}$ and $Y_{t,n}:=Y_{i,n}$ for $t\in[t_{i-1,n},t_{i,n})$.

3.2 Convergence proof for the Compound Poisson case

Now we establish a first convergence result for the case of a Compound Poisson process as driving Lévy process.

Theorem 9 *Let L_t be a Compound Poisson process with $E\left(L_1^2\right) < +\infty$. The Skorokhod distance computed between the processes $\left(G_t, V_t\right)_{t\geq 0}$ and on their constant piecewise version $\left(G_{t,n}, V_{t,n}\right)_{t\geq 0}$ converges in probability to zero, i.e.:*

$$\rho\left(\left(G_{t,n}, V_{t,n}\right)_{t\geq 0}, \left(G_t, V_t\right)_{t\geq 0}\right) \xrightarrow{P} 0 \text{ as } n \to +\infty.$$

Proof. The proof follows the same steps as in Maller et al. (2008)

1. Approximation procedure for the underlying process.
2. Approximation procedure for the variance process.
3. Approximation procedure for the COGARCH(p,q) model.
4. Convergence of the pair in the Skorokhod distance.

Steps 1, 3, 4 are exactly the same as in Maller et al. (2008). To prove that the continuous piecewise constant variance process $V_{t,n}$ converges ucp on a compact time interval to the continuous-time process V_t we first need to show that $Y_{t,n} \xrightarrow{ucp} Y_t$. The counting process $N_{t,n}$ increases by one unit in each subinterval $(t_{i-1,n}, t_{i,n}]$, $i = 1, 2, \ldots, n$, at the first time the jump is of magnitude greater or equal to m_n or at time $t_{i,n}$ if that jump does not occur.
We construct the time process $\Gamma_{t,n}$ as:

$$\Gamma_{t,n} = \sum_{i=1}^{N_{t,n}} \Delta t_{i,n}. \tag{24}$$

Now we want to show that the piecewise constant process $Y_{t,n} := Y_{i,n}$ for $t \in [t_{i-1,n}, t_{i,n})$ converges in ucp to the process $\bar{Y}_{t,n} := e^{B\left(t-\Gamma_{t,n}\right)} Y_{i,n}$ i.e.:

$$\sup_{0\leq t\leq T} \left\|\bar{Y}_{t,n} - Y_{t,n}\right\| \xrightarrow{P} 0.$$

For each $t \in [0, T]$, we have:

$$\begin{aligned}
\left\|Y_{t,n} - \bar{Y}_{t,n}\right\| &= \left\|e^{B\left(t-\Gamma_{t,n}\right)} Y_{i,n} - Y_{i,n}\right\| \\
&\leq \left\|e^{B\left(t-\Gamma_{t,n}\right)} - I\right\|_M \left\|Y_{i,n}\right\| \\
&= \left\|e^{B\left(t-\Gamma_{t,n}\right)} - I - B\left(t-\Gamma_{t,n}\right) + B\left(t-\Gamma_{t,n}\right)\right\|_M \left\|Y_{i,n}\right\| \\
&\leq \left(\left\|e^{B\left(t-\Gamma_{t,n}\right)} - I - B\left(t-\Gamma_{t,n}\right)\right\|_M + \left\|B\left(t-\Gamma_{t,n}\right)\right\|_M\right) \left\|Y_{i,n}\right\|.
\end{aligned}$$

Using the inequality in (2), we get:

$$\begin{aligned}
\left\|Y_{t,n} - \bar{Y}_{t,n}\right\| &\leq \left(e^{\left\|B\left(t-\Gamma_{t,n}\right)\right\|_M} - 1\right) \left\|Y_{i,n}\right\| \\
&\leq \left(e^{\|B\|_M \Delta t_n} - 1\right) \left\|Y_{i,n}\right\|.
\end{aligned} \tag{25}$$

Since by construction $Y_{t,n} = Y_{i,n}$ with $t \in [t_{i-1,n}, t_{i,n})$ and $Y_{t,n}$ has càdlàg paths, it follows that $\sup_{t\in[0,T]} \|Y_{t,n}\|$ is almost surely finite and

$$\sup_{t\in[0,T]} \left\|Y_{t,n} - \bar{Y}_{t,n}\right\| \leq \left(e^{\|B\|_M \Delta t_n} - 1\right) \sup_{t\in[0,T]} \|Y_{t,n}\| \xrightarrow{P} 0$$

as $n \to +\infty$.
The next step is to show the convergence ucp of $\bar{Y}_{t,n}$ to $\tilde{Y}_{t,n}$ where the last process is defined as:

$$\tilde{Y}_{t,n} = e^{B(t-\Gamma_{t,n})}\tilde{Y}_{i,n} \tag{26}$$

with:

$$\tilde{Y}_{i,n} = \tilde{C}_{i,n}\tilde{Y}_{i-1,n} + \tilde{D}_{i,n} \tag{27}$$

where the random matrix $\tilde{C}_{i,n}$ and the random vector $\tilde{D}_{i,n}$ are respectively:

$$\begin{aligned} \tilde{C}_{i,n} &= \left(I + \left(\mathbf{1}_{\tau_{i,n}<+\infty}\Delta L_{\tau_{i,n}}\right)^2 \mathbf{e}\mathbf{a}^\top\right) e^{B\Delta t_{i,n}} \\ \tilde{D}_{i,n} &= a_0 \left(\mathbf{1}_{\tau_{i,n}<+\infty}\Delta L_{\tau_{i,n}}\right)^2 \mathbf{e}. \end{aligned} \tag{28}$$

We consider

$$\sup_{t\in[0,T]} \left\|\tilde{Y}_{t,n} - \bar{Y}_{t,n}\right\| \leq e^{\|B\|_M \Delta t_n} \sup_{i=1,\dots,N_n} \left\|\tilde{Y}_{i,n} - Y_{i,n}\right\| \tag{29}$$

and observe that, for $i = 1, \dots, N_n$, we have:

$$\left\|\tilde{Y}_{i,n} - Y_{i,n}\right\| \leq \left\|\tilde{C}_{i,n}\tilde{Y}_{i-1,n} - C_{i,n}Y_{i-1,n}\right\| + \left\|\tilde{D}_{i,n} - D_{i,n}\right\|. \tag{30}$$

We analyze the second term in (30) and get:

$$\begin{aligned} \left\|\tilde{D}_{i,n} - D_{i,n}\right\| &= \left\|a_0 \left(\mathbf{1}_{\tau_{i,n}<+\infty}\Delta L_{\tau_{i,n}}\right)^2 \mathbf{e} - a_0\epsilon^2_{i,n}\Delta t_{i,n}\mathbf{e}\right\| \\ &\leq |a_0| \left|\left(\mathbf{1}_{\tau_{i,n}<+\infty}\Delta L_{\tau_{i,n}}\right)^2 - \epsilon^2_{i,n}\Delta t_{i,n}\right|. \end{aligned} \tag{31}$$

It can be shown that the first term in (30) is bounded by adding and subtracting the quantity $C_{i,n}\tilde{Y}_{i-1,n}$. Indeed:

$$\begin{aligned} \left\|\tilde{C}_{i,n}\tilde{Y}_{i-1,n} - C_{i,n}Y_{i-1,n}\right\| &= \left\|\tilde{C}_{i,n}\tilde{Y}_{i-1,n} - C_{i,n}\tilde{Y}_{i-1,n} + C_{i,n}\tilde{Y}_{i-1,n} - C_{i,n}Y_{i-1,n}\right\| \\ &\leq \left\|\tilde{C}_{i,n} - C_{i,n}\right\|_M \left\|\tilde{Y}_{i-1,n}\right\| + \|C_{i,n}\|_M \left\|\tilde{Y}_{i-1,n} - Y_{i-1,n}\right\| \\ &\leq \left\|\left[\left(\mathbf{1}_{\tau_{i,n}<+\infty}\Delta L_{\tau_{i,n}}\right)^2 - \epsilon^2_{i,n}\Delta t_{i,n}\right] \mathbf{e}\mathbf{a}^\top e^{B\Delta t_{i,n}}\right\|_M \left\|\tilde{Y}_{i-1,n}\right\| \\ &+ \|C_{i,n}\|_M \left\|\tilde{Y}_{i-1,n} - Y_{i-1,n}\right\| \\ &\leq \left|\left(\mathbf{1}_{\tau_{i,n}<+\infty}\Delta L_{\tau_{i,n}}\right)^2 - \epsilon^2_{i,n}\Delta t_{i,n}\right| \left\|\mathbf{e}\mathbf{a}^\top\right\|_M e^{\|B\|_M\Delta t_{i,n}} \left\|\tilde{Y}_{i-1,n}\right\| \\ &+ \|C_{i,n}\|_M \left\|\tilde{Y}_{i-1,n} - Y_{i-1,n}\right\|. \end{aligned} \tag{32}$$

Substituting (32) and (31) into (30) we have:

$$\begin{aligned} \left\|\tilde{Y}_{i,n} - Y_{i,n}\right\| &\leq \|C_{i,n}\|_M \left\|\tilde{Y}_{i-1,n} - Y_{i-1,n}\right\| \\ &+ \left|\left(\mathbf{1}_{\tau_{i,n}<+\infty}\Delta L_{\tau_{i,n}}\right)^2 - \epsilon^2_{i,n}\Delta t_{i,n}\right| \left(|a_0| + \left\|\mathbf{e}\mathbf{a}^\top\right\|_M e^{\|B\|_M\Delta t_{i,n}} \left\|\tilde{Y}_{i-1,n}\right\|\right). \end{aligned} \tag{33}$$

Since a.s.:

$$\left\| C_{i,n} \right\|_M \le \left(1 + \epsilon_{i,n}^2 \Delta t_{i,n} \left\| \mathbf{e}\mathbf{a}^\top \right\|_M \right) e^{\|B\|_M \Delta t_{i,n}} := C_{i,n}^\star \ge 1 \tag{34}$$

and defining

$$K_{i-1,n} := |a_0| + \left\| \mathbf{e}\mathbf{a}^\top \right\|_M e^{\|B\|_M \Delta t_{i,n}} \left\| \tilde{Y}_{i-1,n} \right\| \tag{35}$$

we have:

$$\begin{aligned} \left\| \tilde{Y}_{i,n} - Y_{i,n} \right\| \le{}& C_{i,n}^\star \left\| \tilde{Y}_{i-1,n} - Y_{i-1,n} \right\| \\ &+ \left| \left(\mathbf{1}_{\tau_{i,n} < +\infty} \Delta L_{\tau_{i,n}} \right)^2 - \epsilon_{i,n}^2 \Delta t_{i,n} \right| K_{i-1,n}. \end{aligned} \tag{36}$$

As observed in (34) $C_{i,n}^\star \ge 1$ and applying Theorem 7 to the inequality in (36) we obtain the following result:

$$\begin{aligned} \sup_{i=1,\dots,N_n} \left\| \tilde{Y}_{i,n} - Y_{i,n} \right\| \le{}& \left[\prod_{i=0}^{N_n-1} C_{N_n-i,n}^\star \right] \left\| \tilde{Y}_{0,n} - Y_{0,n} \right\| + \left| \left(\mathbf{1}_{\tau_{N_n,n} < +\infty} \Delta L_{\tau_{N_n,n}} \right)^2 - \epsilon_{N_n,n}^2 \Delta t_{N_n,n} \right| K_{N_n-1,n} \\ &+ \sum_{i=1}^{N_n-1} \left[\prod_{h=1}^{i} C_{N_n+1-h,n}^\star \right] \left| \left(\mathbf{1}_{\tau_{N_n-i,n} < +\infty} \Delta L_{\tau_{N_n-i,n}} \right)^2 - \epsilon_{N_n-i,n}^2 \Delta t_{N_n-i,n} \right| K_{N_n-1-i,n}. \end{aligned} \tag{37}$$

Moreover:

$$\left[\prod_{i=0}^{N_n-1} C_{N_n-i,n}^\star \right] \left\| \tilde{Y}_{0,n} - Y_{0,n} \right\|, \quad \ge 0 \; n \ge 1$$

with

$$E\left[\left(\prod_{i=0}^{N_n-1} C_{N_n-i,n}^\star \right) \left\| \tilde{Y}_{0,n} - Y_{0,n} \right\| \right] = E\left[\left(\prod_{i=0}^{N_n-1} C_{N_n-i,n}^\star \right) \right] \left\| \tilde{Y}_{0,n} - Y_{0,n} \right\|.$$

Since $\tilde{Y}_{0,n} = Y_{0,n}$ we have that[2]:

$$E\left[\left(\prod_{i=0}^{N_n-1} C_{N_n-i,n}^\star \right) \right] \left\| \tilde{Y}_{0,n} - Y_{0,n} \right\| = 0 \Rightarrow \left(\prod_{i=0}^{N_n-1} C_{N_n-i,n}^\star \right) \left\| \tilde{Y}_{0,n} - Y_{0,n} \right\| = 0 \text{ a.s.} \tag{38}$$

Condition (37) becomes:

$$\begin{aligned} \sup_{i=1,\dots,N_n} \left\| \tilde{Y}_{i,n} - Y_{i,n} \right\| \le{}& \left| \left(\mathbf{1}_{\tau_{N_n,n} < +\infty} \Delta L_{\tau_{N_n,n}} \right)^2 - \epsilon_{N_n,n}^2 \Delta t_{N_n,n} \right| K_{N_n-1,n} \\ &+ \sum_{i=1}^{N_n-1} \left[\prod_{h=1}^{i} C_{N_n+1-h,n}^\star \right] \left| \left(\mathbf{1}_{\tau_{N_n-i,n} < +\infty} \Delta L_{\tau_{N_n-i,n}} \right)^2 - \epsilon_{N_n-i,n}^2 \Delta t_{N_n-i,n} \right| K_{N_n-1-i,n}. \end{aligned} \tag{39}$$

[2]Here we consider $\tilde{Y}_{0,n}$ and $Y_{0,n}$ deterministic starting values for processes $\tilde{Y}_{i,n}$ and $Y_{i,n}$ respectively. Moreover both are equal to the starting point Y_0 of the state process Y_t. Even if we consider Y_0 as a random vector independent of the driving Lévy process, the conclusion in (38), i.e. $\left(\prod_{i=0}^{N_n-1} C_{N_n-i,n}^\star\right) \left\| \tilde{Y}_{0,n} - Y_{0,n} \right\| = 0$ a.s., remains valid.

Defining:

$$Q_n := \left|\left(\mathbf{1}_{\tau_{N_n,n}<+\infty}\Delta L_{\tau_{N_n,n}}\right)^2 - \epsilon^2_{N_n,n}\Delta t_{N_n,n}\right| K_{N_n-1,n}$$
$$+\sum_{i=1}^{N_n-1}\left[\prod_{h=1}^{i} C^{\star}_{N_n+1-h,n}\right]\left|\left(\mathbf{1}_{\tau_{N_n-i,n}<+\infty}\Delta L_{\tau_{N_n-i,n}}\right)^2 - \epsilon^2_{N_n-i,n}\Delta t_{N_n-i,n}\right| K_{N_n-1-i,n}.$$

we observe that Q_n can be bounded. Indeed, $\forall i = 1,\ldots,N_n$:

$$\prod_{h=1}^{i} C^{\star}_{N_n+1-h,n} \leq \prod_{h=1}^{N_n} C^{\star}_{N_n+1-h,n}$$

and, from Theorem 8 we have that $\prod_{h=1}^{N_n} C^{\star}_{N_n+1-h,n}$ converges in probability to a non negative r.v. that is a.s. bounded by:

$$e^{\|B\|_M T+\sum_{0\leq s\leq T}\ln\left(1+\Delta L_s^2\left\|\mathbf{e}\mathbf{a}^\top\right\|_M\right)}.$$

Even $\sup\limits_{i=1,\ldots,N_n} K_{i,n}$ is bounded a.s. $\forall n$. Consequently, we have:

$$Q_n \leq \left[\prod_{h=1}^{N_n} C^{\star}_{N_n+1-h,n}\right]\left[\sup_{i=1,\ldots,N_n} K_{i,n}\right]\sum_{i=1}^{N_n}\left|\left(\mathbf{1}_{\tau_{i,n}<+\infty}\Delta L_{\tau_{i,n}}\right)^2 - \epsilon^2_{i,n}\Delta t_{i,n}\right|. \tag{40}$$

Since $\lim\limits_{n\to+\infty}\sup\limits_{i=1,\ldots,N_n} K_{i,n} = M < +\infty$ a.s. and, as shown in Maller et al. (2008),

$$\sup_{t\in[0,T]}\sum_{i=1}^{N_{t,n}}\left|\left(\mathbf{1}_{\tau_{i,n}<+\infty}\Delta L_{\tau_{i,n}}\right)^2 - \epsilon^2_{i,n}\Delta t_{i,n}\right| \xrightarrow{p} 0 \tag{41}$$

as $n\to+\infty$, then $Q_n \xrightarrow{p} 0$ that implies:

$$\sup_{t\in[0,T]}\left\|\bar{Y}_{t,n}-\tilde{Y}_{t,n}\right\| \xrightarrow{p} 0 \quad \text{as} \quad n\to+\infty. \tag{42}$$

Notice that the result in (42) does not require any particular assumption on the driving noise L_t.

We observe that, if the driven noise is a Compound Poisson, we have only a finite number of jumps on a compact interval $[0,T]$. We denote with τ_k the time of the k-th jump. Since the irregular grid becomes finer as n increases and satisfies the following two conditions:

$$\Delta t_n := \max_{i=1,\ldots,N_n}\Delta t_{i,n} \underset{n\to+\infty}{\rightarrow} 0$$
$$T = \sum_{i=1}^{N_n}\Delta t_{i,n},$$

then, there exists $n^\star$ such that for $n\geq n^\star$, all jump times $\tau_k\in\{t_{0,n},t_{1,n},\ldots,t_{N_n,n}\}$. The COGARCH(p,q) state process Y_t in (9) can be defined equivalently $\forall\ n\geq n^\star$ and for any $i=1,\ldots,N_n$ as:

$$Y_{t_{i,n}} = C_{t_{i,n}}Y_{t_{i-1,n}} + D_{t_{i,n}} \tag{43}$$

with coefficients $C_{t_{i,n}}$ and $D_{t_{i,n}}$ defined as:

$$
\begin{aligned}
C_{t_{i,n}} &= \left(I + \Delta L^2_{t_{i,n}} \mathbf{e}\mathbf{a}^\top\right) e^{B\Delta t_{i,n}} \\
D_{t_{i,n}} &= a_0 \Delta L^2_{t_{i,n}} \mathbf{e}.
\end{aligned}
$$

To show the ucp convergence of process $\tilde{Y}_{t,n}$ to Y_t, we start observing that:

$$
\begin{aligned}
\sup_{t\in[0,T]} \left\| Y_t - \tilde{Y}_{t,n} \right\| &= \sup_{t\in[0,T]} \left\| e^{B(t-\Gamma_{t,n})} \left(Y_{t_{i,n}} - \tilde{Y}_{i,n} \right) \right\| \\
&\leq e^{\|B\|_M T} \sup_{i=1,\dots,N_n} \left\| Y_{t_{i,n}} - \tilde{Y}_{i,n} \right\|. \qquad (44)
\end{aligned}
$$

For fixed n each term in $\sup\limits_{i=1,\dots,N_n} \left\| Y_{t_{i,n}} - \tilde{Y}_{i,n} \right\|$ satisfies:

$$
\left\| Y_{t_{i,n}} - \tilde{Y}_{i,n} \right\| \leq \left\| C_{t_{i,n}} Y_{t_{i-1,n}} - \tilde{C}_{i,n} \tilde{Y}_{i-1,n} \right\| + \left\| D_{t_{i,n}} - \tilde{D}_{i,n} \right\|. \qquad (45)
$$

The term $\left\| D_{t_{i,n}} - \tilde{D}_{i,n} \right\|$ in (45) is bounded as follows:

$$
\left\| D_{t_{i,n}} - \tilde{D}_{i,n} \right\| \leq |a_0| \left| \Delta L^2_{t_{i,n}} - \mathbf{1}_{\tau_{i,n}<+\infty} \Delta L^2_{\tau_{i,n}} \right|. \qquad (46)
$$

Since

$$
\Delta L^2_{t_{i,n}} = \Delta L^2_{t_{i,n}} \mathbf{1}_{|\Delta L_{t_{i,n}}| \geq m_n} + \Delta L^2_{t_{i,n}} \mathbf{1}_{|\Delta L_{t_{i,n}}| < m_n}, \qquad (47)
$$

the inequality in (46) becomes:

$$
\begin{aligned}
\left\| D_{t_{i,n}} - \tilde{D}_{i,n} \right\| &\leq |a_0| \left| \Delta L^2_{t_{i,n}} \mathbf{1}_{0<|\Delta L_{t_{i,n}}|<m_n} \right| \\
&\leq m_n |a_0| \left| \mathbf{1}_{|\Delta L_{t_{i,n}}|>0} \right|. \qquad (48)
\end{aligned}
$$

Inserting (48) into (45), we have:

$$
\left\| Y_{t_{i,n}} - \tilde{Y}_{i,n} \right\| \leq \left\| C_{t_{i,n}} Y_{t_{i-1,n}} - \tilde{C}_{i,n} \tilde{Y}_{i-1,n} \right\| + m_n |a_0| \left| \mathbf{1}_{|\Delta L_{t_{i,n}}|>0} \right|. \qquad (49)
$$

We add and subtract the term $C_{t_{i,n}} \tilde{Y}_{i-1,n}$ into the quantity $\left\| C_{t_{i,n}} Y_{t_{i-1,n}} - \tilde{C}_{i,n} \tilde{Y}_{i-1,n} \right\|$. By exploiting the triangular inequality we obtain:

$$
\begin{aligned}
\left\| C_{t_{i,n}} Y_{t_{i-1,n}} - \tilde{C}_{i,n} \tilde{Y}_{i-1,n} \right\| &\leq \left\| C_{t_{i,n}} Y_{t_{i-1,n}} - C_{t_{i,n}} \tilde{Y}_{i-1,n} \right\| + \left\| C_{t_{i,n}} \tilde{Y}_{i-1,n} - \tilde{C}_{i,n} \tilde{Y}_{i-1,n} \right\| \\
&\leq \left\| C_{t_{i,n}} \right\|_M \left\| Y_{t_{i-1,n}} - \tilde{Y}_{i-1,n} \right\| + \left\| C_{t_{i,n}} - \tilde{C}_{i,n} \right\|_M \left\| \tilde{Y}_{i-1,n} \right\| \\
&\leq \left\| C_{t_{i,n}} \right\|_M \left\| Y_{t_{i-1,n}} - \tilde{Y}_{i-1,n} \right\| \\
&\quad + \left| \Delta L^2_{t_{i,n}} - \mathbf{1}_{\tau_{i,n}<+\infty} \Delta L^2_{\tau_{i,n}} \right| \left\| \mathbf{e}\mathbf{a}^\top \right\|_M e^{\|B\|_M \Delta t_{i,n}} \left\| \tilde{Y}_{i-1,n} \right\|. \qquad (50)
\end{aligned}
$$

Defining:

$$
C^{\star\star}_{t_{i,n}} := \left(1 + \Delta L^2_{t_{i,n}} \left\| \mathbf{e}\mathbf{a}^\top \right\|_M \right) e^{\|B\|_M \Delta t_{i,n}} \geq \left\| C_{t_{i,n}} \right\|_M,
$$

substituting (50) into (49) and using the same arguments as in (47) and (48), we obtain:

$$\begin{aligned}\left\|Y_{t_{i,n}} - \tilde{Y}_{i,n}\right\| \leq C^{\star\star}_{t_{i,n}} \left\|Y_{t_{i-1,n}} - \tilde{Y}_{i-1,n}\right\| + m_n \left|a_0\right| \left|\mathbf{1}_{\left|\Delta L_{t_{i,n}}\right|>0}\right| \\ + \left|\Delta L^2_{t_{i,n}} - \mathbf{1}_{\tau_{i,n}<+\infty} \Delta L^2_{\tau_{i,n}}\right| \left\|\mathbf{e}\mathbf{a}^\top\right\|_M e^{\|B\|_M \Delta t_{i,n}} \left\|\tilde{Y}_{i-1,n}\right\|.\end{aligned}$$

Using $K_{i,n}$ in (35), we have:

$$\left\|Y_{t_{i,n}} - \tilde{Y}_{i,n}\right\| \leq C^{\star\star}_{t_{i,n}} \left\|Y_{t_{i-1,n}} - \tilde{Y}_{i-1,n}\right\| + m_n K_{i-1,n} \left|\mathbf{1}_{\left|\Delta L_{t_{i,n}}\right|>0}\right|. \tag{51}$$

We introduce a stochastic recurrence equation on the grid $\{t_{i,n}\}_{i=0,\dots,N_n}$ defined as:

$$\zeta_{i,n} = C^{\star\star}_{t_{i,n}} \zeta_{i-1,n} + m_n K_{i-1,n} \mathbf{1}_{\left|\Delta L_{t_{i,n}}\right|>0}$$

with initial condition a.s. $\zeta_{0,n} := \left\|Y_{t_{0,n}} - \tilde{Y}_{0,n}\right\| = 0$. Since $\forall i$, $C^{\star\star}_{i,n} \geq 1$ and $m_n K^{\star}_{i-1,n} \mathbf{1}_{\left|\Delta L_{t_{i,n}}\right|>0} \geq 0$ a.s., $\zeta_{i,n}$ is a non decreasing process that is an upper bound for $\left\|Y_{t_{i,n}} - \tilde{Y}_{i,n}\right\|$ for each fixed i then:

$$\begin{aligned}\sup_{i=1,\dots,N_n} \left\|Y_{t_{i,n}} - \tilde{Y}_{i,n}\right\| \leq \left[\prod_{i=0}^{N_n-1} C^{\star\star}_{N_n-i,n}\right] \left\|Y_{t_{0,n}} - \tilde{Y}_{0,n}\right\| \\ + m_n \left\{\sum_{i=1}^{N_n-1} \left[\prod_{h=1}^{i} C^{\star\star}_{N_n+i-h,n}\right] \mathbf{1}_{\left|\Delta L_{t_{N_n-i,n}}\right|>0} K_{N_n-1-i,n} + \mathbf{1}_{\left|\Delta L_{t_{N_n,n}}\right|>0} K_{N_n-1,n}\right\}.\end{aligned} \tag{52}$$

The right-hand side in (52) is non-negative as a summation of non-negative terms. We split it into two parts:

$$\begin{aligned}G_n &:= \left[\prod_{i=0}^{N_n-1} C^{\star\star}_{N_n-i,n}\right] \left\|Y_{t_{0,n}} - \tilde{Y}_{0,n}\right\| \\ W_n &:= m_n \left\{\sum_{i=1}^{N_n-1} \left[\prod_{h=1}^{i} C^{\star\star}_{N_n+i-h,n}\right] \mathbf{1}_{\left|\Delta L_{t_{N_n-i,n}}\right|>0} K_{N_n-1-i,n} + \mathbf{1}_{\left|\Delta L_{t_{N_n,n}}\right|>0} K_{N_n-1,n}\right\}.\end{aligned}$$

Using the same arguments as in (38), we can say that:

$$G_n = 0 \ \texttt{a.s.} \ \forall n \geq 0$$

and

$$W_n \overset{n\to+\infty}{\to} 0.$$

Since for $n \to +\infty$, the quantity

$$\sum_{i=1}^{N_n-1} \left[\prod_{h=1}^{i} C^{\star\star}_{N_n+i-h,n}\right] \mathbf{1}_{\left|\Delta L_{t_{N_n-i,n}}\right|>0} K_{N_n-1-i,n} + \mathbf{1}_{\left|\Delta L_{t_{N_n,n}}\right|>0} K_{N_n-1,n}$$

is composed of a finite number of terms, it is also finite a.s. for the same arguments in (40). In conclusion we have:

$$\sup_{i=1,\dots,N_n} \left\|Y_{t_{i,n}} - \tilde{Y}_{i,n}\right\| \leq G_n + W_n \underset{n\to+\infty}{\to} 0$$

that implies

$$Y_{t,n} \overset{ucp}{\to} Y_t \tag{53}$$

where $Y_{t,n}$ is the constant piecewise process associated to the process $Y_{i,n}$ defined in (16). From (53) we obtain the ucp convergence of process $V_{t,n}$ to the COGARCH(p,q) variance process V_t. The remaining part of the proof follows the same steps as in Maller et al. (2008). ■

3.3 Convergence proof for a General Pure Jump Lévy process

The result can be generalized to any COGARCH(p,q) model driven by a pure jump Lévy process as reported in the following theorem.

Theorem 10 *Let L be a pure jump Lévy process with finite second moment. The Skorokhod distance between the continuous piecewise constant processes $(G_{t,n}, V_{t,n})_{t\geq 0}$ and the processes $(G_t, V_t)_{t\geq 0}$ converges in probability to zero.*

Proof. The proof follows the same steps as in Theorem 9. In particular, it is sufficient to extend the ucp convergence of the piecewise constant process

$$Y_{t,n} := Y_{i,n} \quad \text{with} \quad t \in [t_{i-1,n}, t_{i,n})$$

to the real process Y_t.
As observed previously in the proof of Theorem 9, $\sup_{t\in[0,T]} \left\| \bar{Y}_{t,n} - \tilde{Y}_{t,n} \right\| \underset{n\to+\infty}{\overset{p}{\to}} 0$ where $\tilde{Y}_{t,n}$ is defined by relations (26), (27) and (28) even if the driving noise is not a Compound Poisson process.
We have that $\tilde{Y}_{t,n}$ is the state process driven by a $\sqrt{\epsilon}-$cut Lévy process $L^{(\epsilon)}_t$ as in Definition 8.1 page 810 in Brockwell et al. (2006):

$$L^{(\epsilon)}_t := \sum_{0<s\leq t} \Delta L_s \mathbf{1}_{\Delta|L_s|\geq\sqrt{\epsilon}}$$

where $\sqrt{\epsilon} = m_n$.
The ucp convergence of process $\tilde{Y}_{t,n}$ to Y_t is established by Lemma 8.2 page 810 in Brockwell et al. (2006). ■

4 Maximum Pseudo Log-Likelihood Estimation for the COGARCH(p,q) process

On the irregular grid

$$0 = t_0 < t_1 < \ldots < t_N = T, \tag{54}$$

the increment of a COGARCH(p,q) process is defined as:

$$\Delta G_{t_i} := G_{t_i} - G_{t_{i-1}} = \int_{t_{i-1}}^{t_i} \sqrt{V_u} \mathrm{d}L_u.$$

The pseudo log-likelihood method for parameter estimation given the observations $G_{t_1}, \ldots, G_{t_N}$ is based on the assumption that $\Delta G_{t_i} \sim N\left(E\left[\Delta G_{t_i} \left|\mathcal{F}_{t_{i-1}}\right.\right], Var\left[\Delta G_{t_i} \left|\mathcal{F}_{t_{i-1}}\right.\right]\right)$.
As shown in Chadraa (2009) and recently computed using Teugels martingales in Iacus et al.

(2017), the conditional first moment and the conditional variance for the increments ΔG_{t_i} are respectively:

$$E\left[\Delta G_{t_i}\left|\mathcal{F}_{t_{i-1}}\right.\right]=0 \tag{55}$$

$$Var\left[\Delta G_{t_i}\left|\mathcal{F}_{t_{i-1}}\right.\right]=E\left[L_1\right]\left\{\frac{a_0\Delta t_i b_q}{b_q-a_1\mu}+\mathbf{a}^\top e^{\tilde{B}\Delta t_i}\tilde{B}^{-1}\left(I-e^{-\tilde{B}\Delta t_i}\right)\left[Y_{t_{i-1}}-\frac{a_0\mu}{b_q-a_1\mu}\left(\begin{array}{c}1\\0\\\vdots\\0\end{array}\right)\right]\right\} \tag{56}$$

where $\tilde{B}:=B+\mu\mathbf{e}\mathbf{a}^\top$, $\mu=\int_{\mathcal{R}}y^2\Pi\left(\mathrm{d}y\right)$ and $\Pi\left(y\right)$ is the Lévy measure of the pure jump Lévy process L. For simplicity, we require the underlying process to be centered in zero with unitary second moment $\mu=E\left(L_1^2\right)=1$.
We remark that the conditional variance on the grid (54) is predictable since it is a function of the state process at time t_{i-1}.
It is worth to notice that the conditional variance of ΔG_{t_i} in (56) can not be determined since the state process Y_t is not observable. We approximate Y_{t_i} in (56) with Y_i that satisfies the same recursion in (19) where $G_{i,n}$ is substituted by the observations G_{t_i}, then we have:

$$Y_i=\left(I+\frac{\left(G_{t_i}-G_{t_{i-1}}\right)^2}{a_0+\mathbf{a}^\top Y_{i-1}}\mathbf{e}\mathbf{a}^\top\right)e^{B\Delta t_i}Y_{i-1}+a_0\frac{\left(G_{t_i}-G_{t_{i-1}}\right)^2}{a_0+\mathbf{a}^\top Y_{i-1}}\mathbf{e}. \tag{57}$$

Using (57) we can define $\widehat{Var}\left[\Delta G_{t_i}\left|\mathcal{F}_{t_{i-1}}\right.\right]$ as in (56) with Y_i instead of Y_{t_i}.
The pseudo log-likelihood estimates are determined by solving the following optimization problem:

$$\begin{aligned}&\max_{\mathbf{a},a_0,B\in\Theta}\hat{\mathcal{L}}_N\left(\mathbf{a},a_0,B\right)\\&\texttt{s.t.}\\&\left\{\begin{array}{l}Y_i=\left(I+\frac{\left(G_i-G_{i-1}\right)^2}{a_0+\mathbf{a}^\top Y_{i-1}}\mathbf{e}\mathbf{a}^\top\right)e^{B\Delta t_i}Y_{i-1}+a_0\frac{\left(G_i-G_{i-1}\right)^2}{a_0+\mathbf{a}^\top Y_{i-1}}\mathbf{e}\\i=1,\dots,N\end{array}\right.\end{aligned} \tag{58}$$

where

$$\hat{\mathcal{L}}_N\left(\mathbf{a},a_0,B\right)=-\frac{1}{2}\sum_{i=1}^{N}\left(\frac{\left(\Delta G_{t_i}\right)^2}{\widehat{Var}\left[\Delta G_{t_i}\left|\mathcal{F}_{t_{i-1}}\right.\right]}+\ln\left(\widehat{Var}\left[\Delta G_{t_i}\left|\mathcal{F}_{t_{i-1}}\right.\right]\right)\right)-\frac{N\ln\left(2\pi\right)}{2}$$

and the set Θ contains the model parameters that ensure stationarity, existence of the mean for the state process Y_t and non-negativity of process V_t (see Brockwell et al., 2006, for a discussion on sufficient conditions for the identification of parameters in Θ).

5 Numerical analysis

In this section we analyze the estimation procedure developed in Section 4. The analysis in composed of two parts. The first is a simulation study while the second considers real data collected at high frequency level. The analysis is conducted using the R package `yuima` available on CRAN. The estimation procedure discussed in Section 4 can be performed using the `qmle` function in `yuima` that controls, through the `Diagnostic.Cogarch` function, if model parameters satisfy conditions used for the identification of set Θ in the optimization routine (see Iacus et al., 2017, for details).

5.1 Simulation study

We apply the estimation method developed in Section 4 to simulated data generated on a regular grid through the `simulate` function in `yuima` where we choose the Euler-Maruyama discretization of the SDE in (7) (see Iacus et al., 2017, for more details). In order to verify whether the pseudo maximum likelihood method gives plausible estimates, we perform a simulation-estimation study for different values of orders p and q and for two different driving Lévy processes. The first underlying process is a Compound Poisson with unitary intensity rate whose jump size is standard normally distributed. The second is a Variance Gamma process with parameters $(\lambda, \alpha, \beta, \mu)$ (see Iacus and Mercuri, 2015, for more details on this formulation).
In Table 1 we report for each model the values of parameters used for the simulation exercise. The chosen parameters ensure the underlying Lévy process is centered in zero with unitary second moment $E\left[L_1^2\right] = 1$. In the simulation analysis we consider sample paths with length $T = 3200$ and with $N = 64000$ observations

Insert Here Table 1.

Through the `Diagnostic.Cogarch` function, it is possible to verify that the COGARCH(2,2) with the given parameters is stationary with positive variance process V_t. For each trajectory, we estimate the COGARCH(p,q) parameters by solving the maximization problem in (58) where, to reduce the possibility of falling into a local maximum, the simulated annealing algorithm proposed in Bèlisle (1992) has been considered.
Table 2, for the Compound Poisson used for the driving process and Table 3, for the Variance Gamma case, report for each estimated parameter over the 1000 sample paths the mean, the median, the standard error (s.d.), the root mean squared error (RMSE) around the true value and the expected bias.

Insert Here Table 2.

Insert Here Table 3.

From Tables 2 and 3 we observe similar fitting results for the two driving processes respectively Compound Poisson and Variance Gamma. In both cases for b_1 we get the highest bias especially for COGARCH(2,2) and COGARHC(1,3). By increasing N and T we obtain a better fitting also for b_1. For example for $N = 250000$ and $T = 8000$ the bias in the COGARCH(2,2) is 0.07. It is interesting to notice also that the fitting results are influenced by the algorithm we choose in order to avoid local maxima. We observe that for small variations of parameter b_1, the `Diagnostic.Cogarch` function returns a corresponding model where stationarity and/or positivity of variance is not guaranteed. In this cases, the simulated annealing pushes the value b_1 away from the "barrier" where the two conditions above are not guaranteed. These additional penalties influenced the mean of the estimates while the median resulted to be closer to the true value. For increasing N and T these numerical issues have a smaller influence on final results. To the best of our knowledge, statistical properties of the estimators obtained by solving the problem (58) have not been stated yet. However, results reported in Tables 2 and 3 seem to be promising.

5.2 Real Data

In this section we show how to estimate a COGARCH(p,q) model on real data using the pseudo log-likelihood estimation procedure described in Section 4. To better approximate the continuous time framework we consider a dataset composed of 19230 observations in 5-minute intervals of the S&P500 index ranging from 2016-02-22 15:34:59 Central European Time (CET hereafter) to 2017-02-21 21:59:59 CET as reported in Figure 1.

Insert here Figure 1

The usage of the COGARCH(p,q) model, in this context, allows us to manage irregularly spaced data (due to weekends, holidays and breaks during the trading hours) without considering any missing data imputation method. To show the irregularity in our dataset, we report in Figure 2 the number of observations in each day.

Insert here Figure 2

Before fitting the COGARCH(p,q) model to the data, we decide to perform a data cleaning procedure as in Müller et al. (2009) by subtracting to observed data the local trend and subsequently dividing by a local volatility weight both determined using a time frame of 21 trading days. The local trend is estimated as the moving average of the observations over 21 trading days while the local volatility weight in the same period is computed by reweighting the moving average, over the same time frame, of the absolute values of the detrended observations. For formulas and details on this methodology please refer to Section 3.2 in Müller et al. (2009). This pre-processing stage is required since before applying the COGARCH model to the data we must check whether these data display a stationary volatility pattern over the whole time frame.
In the first row of Figure 3 we plot the observed log-returns. In the same figure second and third rows show respectively the estimated local trends and estimated local volatility weights. In the fourth row we plot the detrended and reweighted log-returns on which we apply the COGARCH model.

Insert here Figure 3

We estimate the COGARCH(2,2) parameters fitted to S&P500 log-prices through the maximization of the pseudo log-likelihood and then recover the estimated increments, on the original irregular grid, of the underlying Lévy process using the procedure described in Iacus et al. (2017).

We consider one day as time unit and get different values for Δt_i: for example for 5-minute intervals $\Delta t_i \approx 0.0035$, for daily closure-opening $\Delta t_i \approx 0.7326$ and from Friday closure to Monday opening $\Delta t_i \approx 2.7326$.
In Table 4 we report the estimated parameters with their standard deviation in brackets computed using the fixed block bootstrap with block length 1 week (see Carlstein, 1986, for more details) in order to maintain the time structure of the data.

Insert here Table 4.

To confirm the existence of an ARCH effect (see Tsay, 2005, for additional information), we report in Table 5 the results of the Lagrange Multiplier Test (ARCH-LM test hereafter) proposed in Engle (1982) and the autocorrelation function of the squared log-returns in Figure 4. Our observations refer to time intervals that are not equally spaced while the ARCH effect should be investigated on quantities that refer to the same interval length. To this aim, we create a fictive dataset that contains log-returns that refer to 5-minute intervals without gaps by linearly interpolating missing values.

Insert here Figure 4.

Insert here Table 5.

The same interpolation is done for missing values of the Lévy increments retrieved from the COGARCH(2,2) and plotted in Figure 5. In Table 6 are displayed the main statistics of the interpolated Lévy increments. The linear interpolation introduces additional dependence in the data but as observed from Table 7 and Figure 6, there is no ARCH effect in the Lévy increments meaning that the dependence structure is well captured by the COGARCH(2,2) model.

Insert here Figure 5.

Insert here Table 6.

Insert here Table 7.

Insert here Figure 6.

We report in Table 8 estimates and standard errors for parameters of COGARCH(p,q) models with $p, q \neq 2$. Notice that COGARCH(1,2), COGARCH(1,3) and COGARCH(2,3) models are estimated on detrended and reweighted data.

Insert here Table 8.

Looking at results of the ARCH-LM test in Table 8, we observe that the ARCH effect has been removed in all considered models. A selection procedure of the best model among those reported in Table 4 and Table 8 requires the evaluation of the likelihood associated to the observations $G_{t_0}, \ldots, G_{t_N}$ that depend on the unobservable process Y_t. A similar problem arises in the context of stochastic volatility models since volatility is unobservable. In the literature of stochastic volatility models, evaluation of the likelihood function involves numerical time-consuming techniques like for example multiple integration (see Danielsson, 1994), Monte Carlo Markov Chain (see Kim et al., 1998) or particle filtering (see Michael and Shephard, 1999). In our framework, once we estimate the COGARCH(p,q) parameters using the approach in Section 4, one can implement an algorithm for the numerical evaluation of the likelihood based on results available in literature.

6 Conclusions

This paper suggests the introduction of a sequence of discrete time processes that converges to a COGARCH(p,q) model in the Skorokhod distance. The approximation is the main block for the estimation procedure based on the maximization of the pseudo log-likelihood function.
In the empirical analysis we showed that the estimation procedure based on our approximation seems to be adequate to deal with irregularly spaced data.

References

Bèlisle, C. J. P. Convergence theorems for a class of simulated annealing algorithms on r^d. *Journal of Applied Probability*, 29(4):885–895, 1992.

Billingsley, P. *Convergence of Probability Measures*. Wiley, New York, 1968.

Brockwell, P., Chadraa, E., and Lindner, A. Continuous-time GARCH processes. *Annals of Applied Probability*, 16(2):790–826, 2006.

Brouste, A., Fukasawa, M., Hino, H., Iacus, S. M., Kamatani, K., Koike, Y., Masuda, H., Nomura, R., Ogihara, T., Y., S., Uchida, M., and N., Y. The YUIMA project: A computational framework for simulation and inference of stochastic differential equations. *Journal of Statistical Software*, 57(4):1–51, 2014.

Carlstein, E. The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The Annals of Statistics*, 14(3):1171–1179, 1986.

Chadraa, E. *Statistical Modelling with COGARCH(P,Q) Processes*, 2009. PhD Thesis.

Danielsson, J. Stochastic volatility in asset prices estimation with simulated maximum likelihood. *Journal of Econometrics*, 64(1-2):375–400, 1994.

Engle, R. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007, 1982.

Haug, S., Klüppelberg, C., Lindner, A., and Zapp, M. Method of moment estimation in the COGARCH(1,1) model. *Econometrics Journal*, 10(2):320–341, 2007.

Iacus, S. M. and Mercuri, L. Implementation of Lévy CARMA model in yuima package. *Computational Statistics*, pages 1–31, 2015.

Iacus, S. M. and Yoshida, N. *Simulation and Inference for Stochastic Processes with YUIMA*. Springer, New York, 2018.

Iacus, S. M., Mercuri, L., and Rroji, E. COGARCH(p,q): Simulation and inference with `yuima` package. *Journal of Statistical Software*, 80:1–49, 2017.

Kim, S., Shephard, N., and Chib, S. Stochastic volatility: Likelihood inference and comparison with arch models. *Review of Economic Studies*, 65(3):361–393, 1998.

Klüppelberg, C., Lindner, A., and Maller, R. A continuous-time GARCH process driven by a Lévy process: Stationarity and second-order behaviour. *Journal of Applied Probability*, 41(3):601–622, 2004.

Maller, R. A., Müller, G., and Szimayer, A. GARCH modelling in continuous time for irregularly spaced time series data. *Bernoulli*, 14(2):519–542, 2008.

Meyer, C. D. *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia, 2000.

Michael, K. P. and Shephard, N. Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Association*, 1994(446):590–599, 1999.

Müller, G. MCMC estimation of the COGARCH(1,1) model. *Journal of Financial Econometrics*, 8(4):481–510, 2010.

Müller, G., Durand, R. B., Maller, R., and Klüppelberg, C. Analysis of stock market volatility by continuous-time garch models. *Stock Market Volatility*, pages 32–48, 2009.

Sato, K.-i. *Lévy processes and infinitely divisible distributions*. Cambridge university press, 1999.

Serre, D. *Matrices : theory and applications*. Springer, New York, 2002.

Stelzer, R. First jump approximation of a Lévy-driven SDE and an application to multivariate ECOGARCH processes. *Stochastic Processes and their Applications*, 119(6):1932–1951, 2009.

Ström, T. On logarithmic norms. *SIAM Journal on Numerical Analysis*, 12(5):741–753, 1975.

Szimayer, A. and Maller, R. A. Finite approximation schemes for Lévy processes, and their application to optimal stopping problems. *Stochastic Processes and their Applications*, 117(10):1422–1447, 2007.

Tsay, R. S. *Analysis of financial time series*. Wiley series in probability and statistics. Wiley-Interscience, New Jersey, 2005.

param.	COGARCH$(1,2)$		COGARCH$(2,2)$		COGARCH$(1,3)$		COGARCH$(2,3)$	
	CP	VG	CP	VG	CP	VG	CP	VG
a_0	0.005	0.005	1.12E-04	1.12E-04	5.51E-03	5.51E-03	1.15E-04	1.15E-04
a_1	0.1	0.1	0.016	0.016	5.43E-03	5.43E-03	0.0374	0.0374
a_2			0.0478	0.0478			0.0389	0.0389
b_1	1.5	1.5	0.781	0.781	1.265	1.265	0.479	0.479
b_2	0.5	0.5	0.1084	0.1084	0.385	0.385	1.131	1.131
b_3					0.0237	0.0237	0.302	0.302
λ_{CP}	1		1		1		1	
μ_{CP}	0		0		0		0	
σ_{CP}	1		1		1		1	
λ_{VG}		1				1		1
α_{VG}		$\sqrt{2}$		$\sqrt{2}$		$\sqrt{2}$		$\sqrt{2}$
β_{VG}		0		0		0		0
μ_{VG}		0		0		0		0

Table 1: True parameters using in the simulation/estimation analysis.

			COGARCH(1,2) - CP			
True Param.	$a_0 = 0.005$	$a_1 = 0.1$	$b_1 = 1.5$	$b_2 = 0.5$		
mean	4.82E-03	0.0951	1.5841	0.4745		
median	4.86E-03	0.0997	1.5406	0.4937		
s.d.	3.53E-04	0.0270	0.1792	0.0948		
RMSE	3.95E-04	0.0275	0.1980	0.0981		
bias	-1.76E-04	-4.85E-03	0.0841	-0.0254		
			COGARCH(2,2) - CP			
True Param.	$a_0 = 1.115\text{E-}04$	$a_1 = 0.016$	$a_2 = 0.0478$	$b_1 = 0.781$	$b_2 = 0.1084$	
mean	9.69E-05	0.0160	0.0512	1.0019	0.1118	
median	9.63E-05	0.0151	0.0501	0.8106	0.1101	
s.d.	7.82E-06	9.35E-03	0.0110	0.4242	0.0164	
RMSE	1.65E-05	9.35E-03	0.0115	0.4782	0.0168	
bias	-1.46E-05	6.41E-05	3.41E-03	2.21E-01	3.46E-03	
			COGARCH(1,3) - CP			
True Param.	$a_0 = 5.51\text{E-}03$	$a_1 = 5.43\text{E-}03$	$b_1 = 1.265$	$b_2 = 0.385$	$b_3 = 0.0237$	
mean	5.42E-03	5.68E-03	1.4267	0.4388	0.0307	
median	5.39E-03	4.97E-03	1.3445	0.4002	0.0237	
s.d.	1.24E-03	4.62E-03	0.4186	0.3214	0.0652	
RMSE	1.24E-03	4.63E-03	0.4488	0.3259	0.0656	
bias	-9.34E-05	2.59E-04	0.1617	0.0538	7.03E-03	
			COGARCH(2,3) - CP			
True Param.	$a_0 = 1.15\text{e-}04$	$a_1 = 0.0374$	$a_2 = 0.0389$	$b_1 = 0.479$	$b_2 = 1.131$	$b_3 = 0.302$
mean	2.46E-03	0.0381	0.0387	0.5062	1.1222	0.3412
median	1.12e-04	0.0374	0.0389	0.4590	1.1209	0.3020
s.d.	0.0744	0.0219	6.22E-03	0.1691	0.0804	0.1631
RMSE	0.0744	0.0219	6.22E-03	0.1713	0.0808	0.1677
bias	2.35E-03	7.91E-04	-1.88E-04	0.0272	-8.72E-03	0.0392

Table 2: Summary statistics of estimated parameters of COGARCH(p,q) models driven by a Compound Poisson process over 1000 simulated sample paths. The Compound Poisson process has unitary intensity rate with jump size standard normally distributed.

			COGARCH$(1,2)$ - VG			
True Param.	$a_0 = 0.005$	$a_1 = 0.1$	$b_1 = 1.5$	b_2 =0.5		
mean	4.97E-03	0.0950	1.6047	0.4720		
median	4.96E-03	0.0998	1.5453	0.4994		
s.d.	3.66E-04	0.0281	0.2728	0.0926		
RMSE	3.67E-04	0.0285	0.2922	0.0967		
bias	-2.97E-05	-4.93E-03	0.1047	-0.0279		
			COGARCH$(2,2)$ - VG			
True Param.	a_0 = 1.115E-04	$a_1 = 0.016$	$a_2 = 0.0478$	$b_1 = 0.781$	$b_2 = 0.1084$	
mean	1.10E-04	0.0179	0.0500	1.0049	0.1104	
median	1.09E-04	0.0171	0.0489	0.7703	0.1094	
s.d.	1.07E-05	0.0102	0.0119	0.4525	0.0120	
RMSE	1.07E-05	0.0104	0.0121	0.5049	0.0122	
bias	-8.82E-07	0.0019	0.0022	0.2239	0.0020	
			COGARCH$(1,3)$ - VG			
True Param.	a_0 = 5.51E-03	a_1 = 5.43E-03	$b_1 = 1.265$	$b_2 = 0.385$	$b_3 = 0.0237$	
mean	5.51E-03	6.06E-03	1.4454	0.4524	0.0311	
median	5.51E-03	5.33E-03	1.3390	0.3976	0.0240	
s.d.	8.65E-04	4.76E-03	0.4756	0.3260	0.0508	
RMSE	8.65E-04	4.80E-03	0.5087	0.3329	0.0513	
bias	-1.26E-06	6.32E-04	0.1804	0.0674	7.49E-03	
			COGARCH$(2,3)$ - VG			
True Param.	a_0 = 1.149e-04	$a_1 = 0.0374$	$a_2 = 0.0389$	$b_1 = 0.479$	$b_2 = 1.131$	$b_3 = 0.302$
mean	5.49E-04	0.0372	0.0374	0.4808	1.1200	0.3174
median	1.14E-04	0.0374	0.0389	0.4503	1.1205	0.3019
s.d	0.0189	1.85E-03	0.0412	0.1199	0.0912	0.0880
RMSE	0.0189	1.85E-03	0.0412	0.1200	0.0919	0.0893
bias	4.34E-04	-1.11E-04	-1.44E-04	1.81E-03	-0.0110	0.0154

Table 3: Summary statistics of estimated parameters of COGARCH(p,q) models driven by a Variance Gamma process over 1000 simulated sample paths. The parameters of the Variance Gamma process are $\left(1, \sqrt{2}, 0, 0\right)$.

		Estimated parameters		
a_0	a_1	a_2	b_1	b_2
1.115e-04	1.624e-02	0.0479	0.7812	0.1084
(6.758e-07)	(2.052e-05)	(1.275e-04)	(1.682e-04)	(2.054e-04)

Table 4: Estimated parameters and their standard deviation for fitted COGARCH(2,2) to S&P500 log-prices.

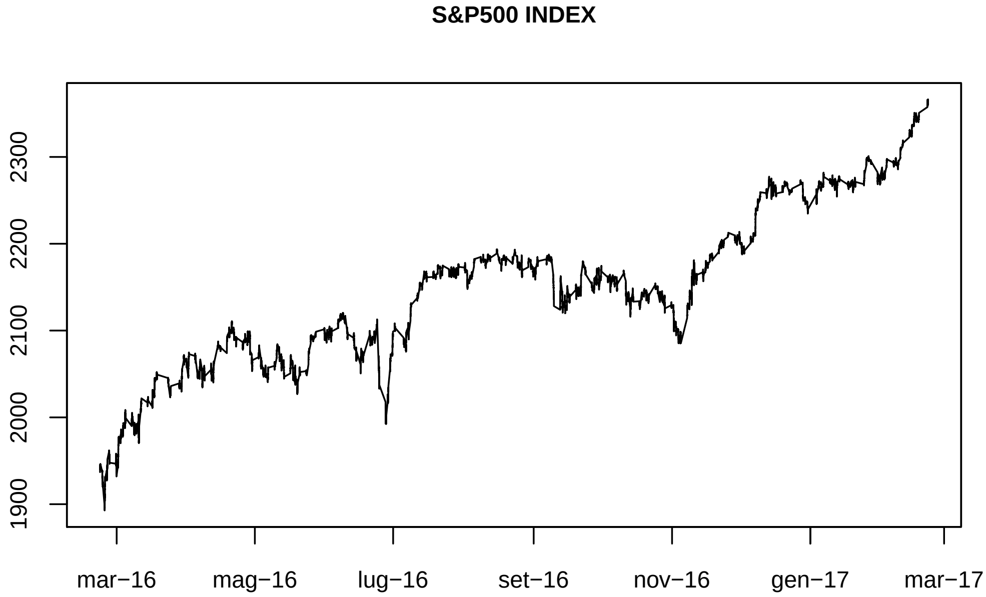

Figure 1: Observations in 5-minute intervals of the S&P500 Index from 2016-02-22 15:34:59 CET to 2017-02-21 21:59:59 CET.

LM Stat.	DF	p-value
56.953	5	5.170e-11
62.663	7	4.427e-11
64.837	10	4.355e-10
66.504	12	1.435e-09

Table 5: Lagrange Multiplier test for detecting ARCH effect in S&P500 log-returns.

Lévy increments				
mean	median	s.d.	skew	kurt
-4.453e-06	1.135e-04	3.966e-02	-6.6667e-01	70.9932
min	1st Qu.	3rd Qu.	max	Interquart
-1.1323	-1.5384e-02	1.5693e-02	6.9596e-01	3.107e-02

Table 6: Main statistics of underlying Lévy increments in a COGARCH(2,2) model fitted to S&P500 log-returns.

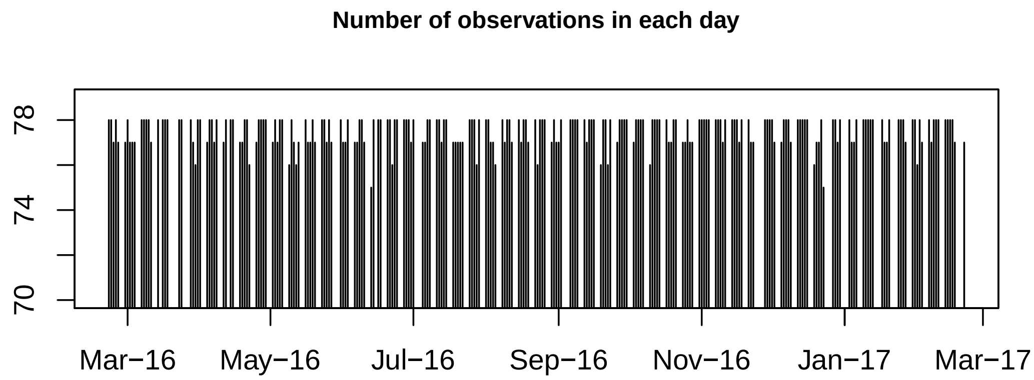

Figure 2: Number of observations in each day.

ARCH effect		
LM stat	DF	p-value
8.5417	5	0.1288
10.657	7	0.1543
11.424	10	0.3255
12.057	12	0.4411

Table 7: Lagrange Multiplier test for detecting ARCH effect in recovered increments of underlying Lévy process.

	COGARCH$(1,2)$		COGARCH$(1,3)$		COGARCH$(2,3)$	
	Est.	s.d.	Est.	s.d.	Est.	s.d.
a_0	3.24E-03	1.74E-06	1.00E-0.4	1.57e-05	1.15E-04	1.70e-06
a_1	0.09172	0.03250	0.00543	3.44E-03	0.03743	0.32E-03
a_2					0.03893	0.59E-03
b_1	1.58062	0.16508	1.26596	0.05415	0.47995	0.24553
b_2	0.51917	0.03051	0.38577	0.04574	1.13110	0.33072
b_3			0.02379	9.02E-03	0.30261	0.14972
DF	LM-Test	p-value	LM-Test	p-value	LM-Test	p-value
5	9.5866	0.08783	10.44140	0.06365	9.55835	0.08876
7	11.3349	0.12466	12.28151	0.09167	11.31661	0.12539
10	11.6900	0.30633	12.70160	0.24083	11.65903	0.30852
12	11.6084	0.45631	12.89741	0.37654	11.83280	0.45919

Table 8: Comparison among different COGARCH(p,q) models on real data.

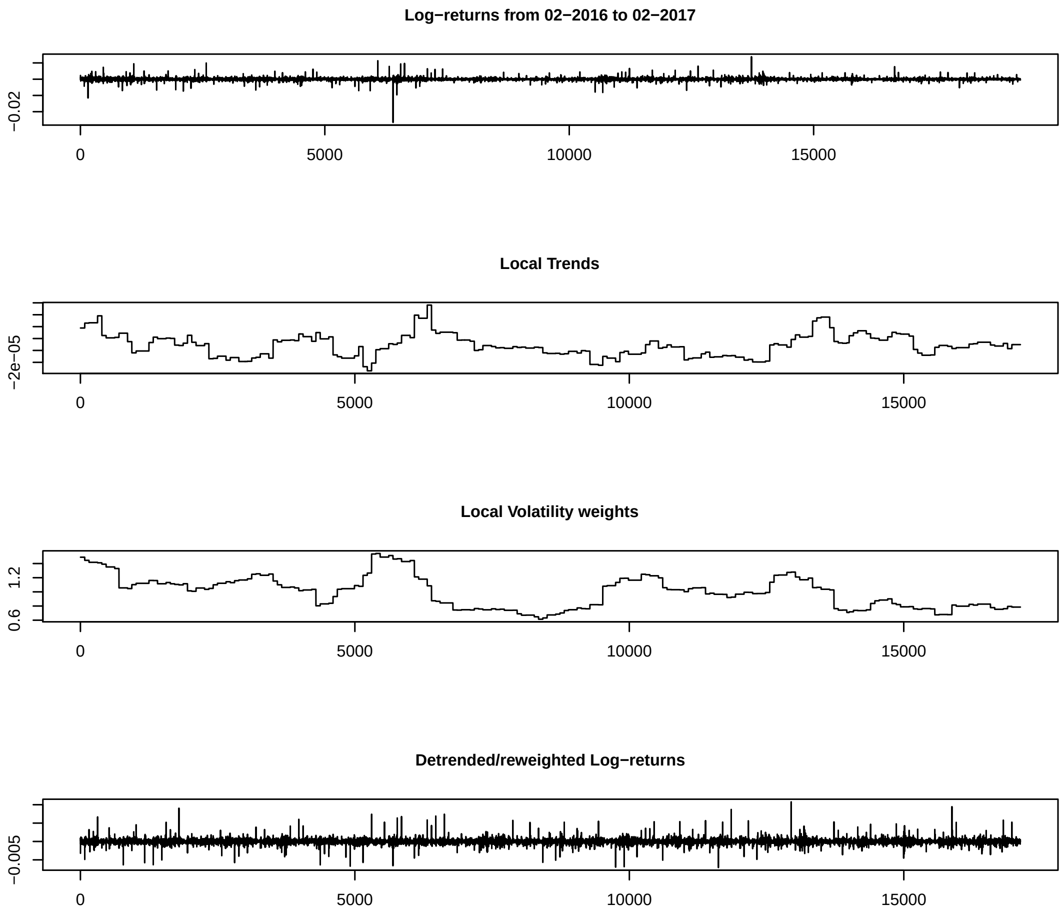

Figure 3: 5 minutes log-returns of S&P500 index, local trends, local volatility weights and detrended/reweighted log-returns.

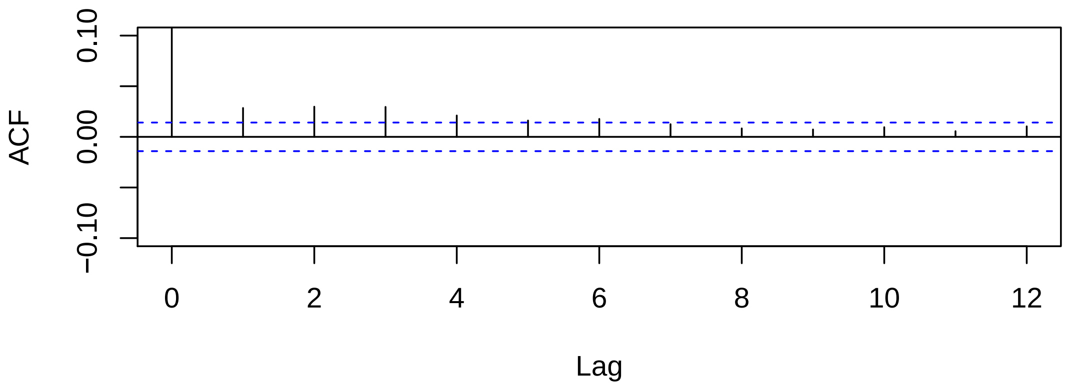

Figure 4: Autocorrelation of squared logreturns of the S&P500 index.

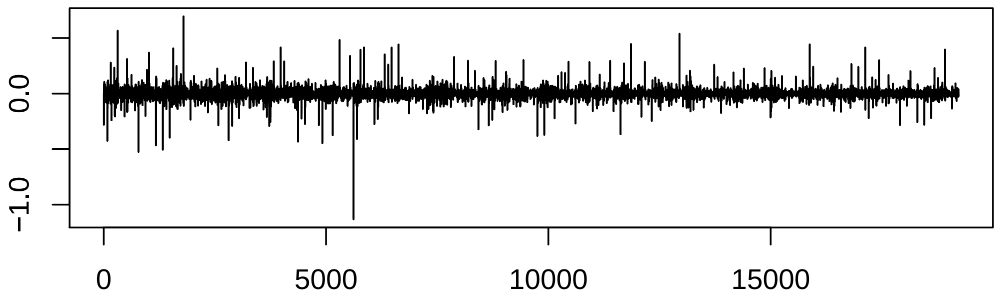

Figure 5: Lévy increments obtained fitting a COGARCH(2,2) model to S&P 500 log-returns.

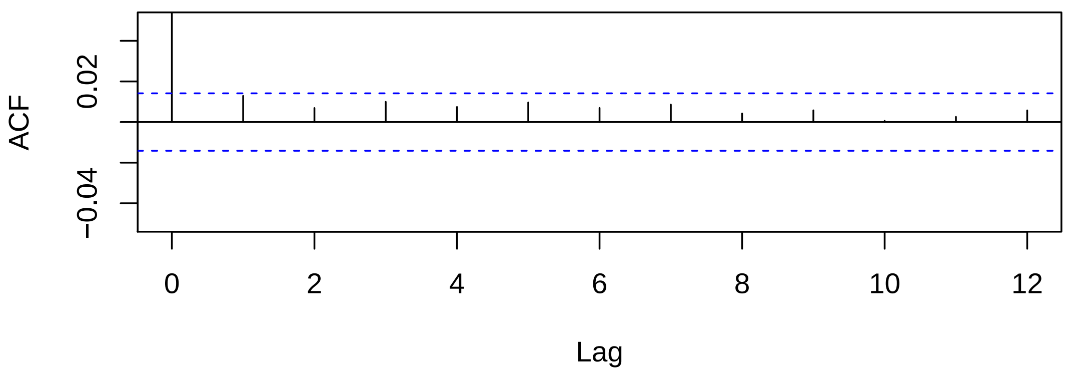

Figure 6: ACF for underlying Lévy increments of a COGARCH(2,2) model fitted to S&P 500 log-returns.