DISSERTATION

STATISTICAL MODELING WITH COGARCH(p, q) PROCESSES

Submitted by

Erdenebaatar Chadraa

Statistics Department

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2009

UMI Number: 3400988

COLORADO STATE UNIVERSITY

August 28, 2009

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY ERDENEBAATAR CHADRAA ENTITLED STATISTICAL MODELING WITH COGARCH(p, q) PROCESSES BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

P. J. Brockwell

Peter J. Brockwell (Adviser)

Wang Haonan

Haonan Wang (Committee Member)

Chihoon Lee (Committee Member)

M. R. Azimi-Sadjadi

Mahmood Azimi-Sadjadi (Committee Member)

F. Jay Breidt (Department Head)

ABSTRACT OF DISSERTATION

STATISTICAL MODELING WITH COGARCH(p,q) PROCESSES

In this paper, a family of continuous time GARCH processes, generalizing the COGARCH(1, 1) process of Klüppelberg, et. al. (2004), is introduced and studied. The resulting COGARCH(p,q) processes, $q \geq p \geq 1$, exhibit many of the characteristic features of observed financial time series, while their corresponding volatility and squared increment processes display a broader range of autocorrelation structures than those of the COGARCH(1, 1) process. We establish sufficient conditions for the existence of a strictly stationary non-negative solution of the equations for the volatility process and, under conditions which ensure the finiteness of the required moments, determine the autocorrelation functions of both the volatility and squared increment processes. The volatility process is found to have the autocorrelation function of a continuous-time ARMA process while the squared increment process has the autocorrelation function of an ARMA process.

To estimate the parameters of the COGARCH(2, 2) processes, the least-squares method is used. We give conditions under which the volatility and the squared increment processes are strongly mixing, from which it follows that the least-squares estimators are strongly consistent and asymptotically normal. Finally, the model is fitted to a high frequency dataset.

Erdenebaatar Chadraa
Statistics Department
Colorado State University
Fort Collins, CO 80523
Fall 2009

ACKNOWLEDGEMENTS

First, I would like to thank my adviser Dr. Peter Brockwell for his continued guidance, inspiration and support and for uncountable discussions and revisions of this dissertation. I would also like to thank my graduate committee members, Dr. Haonan Wang, Dr. Chihoon Lee, and Dr. Mahmood Azimi-Sadjadi for their suggestions and advise.

I thank Dr. Peter Brockwell, Dr. Richard Davis and Dr. Hari Iyer for helping me during a hard time I encountered a few years ago.

I am eternally thankful to my parents for supporting me through all my adventures. I know, in particular, the great distance has been hard for you.

I would like to thank my wife Tugsgerel for her love, support and understanding over the years. I thank my daughter Michelle and son Irvin for their love and patience with my busy schedule during this graduate study. Their sweet support has always been truly touching. This dissertation would not have been possible without them. I dedicate it to my family.

Contents

1 Preliminaries 1

1.1 Introduction . . . 1

1.2 ARMA processes . . . 4

1.3 Lévy processes . . . 7

1.4 CARMA processes . . . 10

1.5 GARCH processes . . . 16

2 COGARCH processes 19

2.1 The COGARCH$(1,1)$ process . . . 19

2.2 The COGARCH(p,q) equations . . . 20

2.3 Stationarity conditions . . . 26

2.4 Second order properties of the volatility process . . . 36

2.5 Positivity conditions for the volatility . . . 48

2.6 The autocorrelation of the squared increments . . . 51

2.7 An Example . . . 58

3 Parameter estimation 61

3.1 Preliminary estimation . . . 61

3.2 Least-squares estimators . . . 63

3.3 Strong consistency and asymptotic normality of the least-squares estimators . . . 66

3.4 The estimation algorithm . . . 75

3.5 A simulation study . . . 76

3.6 Estimating the volatility . . . 80

3.7 Real data analysis . . . 84

3.7.1 Filtering the data . . . 86

3.7.2 Fitting COGARCH$(2,2)$ to squared 30-minute returns . . . 88

3.7.3 Volatility estimation and goodness-of-fit 89

4 Conclusions **92**

4.1 Summary 92

4.2 Future problems 93

References **94**

Appendix **98**

List of Figures

1	A sample path of a standard Brownian Motion	9
2	A Sample path of a compound Poisson process with standard normal jumps	10
3	The simulated compound-Poisson driven COGARCH process	59
4	The sample autocorrelation functions of the volatilities and of the squared COGARCH increments	60
5	ACF of the squared increment process in the COGARCH(1,2) case	73
6	Preliminary estimators of ϕ	77
7	Least-squares estimators of the ARMA parameters	79
8	Least-squares estimators of the COGARCH parameters	79
9	COGARCH$(1,2)$ estimated volatility	83
10	Dow Jones 5-minutes data recorded from February 12th, 2003 to May 12th, 2006	84
11	Dow Jones unfiltered log returns, their squares and ACF	85
12	Dow Jones estimated seasonality coefficients	87
13	Dow Jones filtered log returns, their squares and ACF	87
14	Estimated volatilities and residuals for the 30-minute log returns	90
15	ACF of the squared 30-minute log returns and the residuals	91

List of Tables

3.1 Least-squares estimates of COGARCH parameters 80

1 Preliminaries

1.1 Introduction

In financial econometrics, discrete-time GARCH (i.e. generalised autoregressive conditionally heteroscedastic) processes are widely used to model the returns at regular intervals on stocks, currency investments and other assets. Specifically, a GARCH process $(\xi_n)_{n\in\mathbb{N}}$ typically represents the increments, $\ln P_n - \ln P_{n-1}$, of the logarithms of the asset price at times $1, 2, 3, \ldots$. These models capture many of the so called *stylized features* of such data, e.g. tail heaviness, volatility clustering and dependence without correlation. For GARCH processes with finite fourth moments, the autocorrelation functions of both the squared process and of the associated volatility process are those of ARMA (autoregressive moving average) processes. The squared GARCH$(1,1)$ process, for example, has the autocorrelation function (ACF) of an ARMA$(1,1)$ process and the corresponding volatility has the autocorrelation function of an AR(1) process.

Various attempts have been made to capture the stylized features of financial time series using continuous-time models. The interest in continuous-time models stems from their use in modelling irregularly spaced data, their use in financial applications such as option-pricing and the current wide-spread availability of high-frequency data. In continuous time it is natural to model the logarithm of the asset price itself, i.e. $G_t = \ln P_t$, rather than its increments as in discrete time.

Notable among these attempts is the GARCH diffusion approximation of Nelson (1990). (See also Duan (1997) and Drost and Werker (1996).) Although the GARCH process is driven by a single noise sequence, the diffusion limit is driven by two independent Brownian motions $(W_t^{(1)})_{t\geq 0}$ and $(W_t^{(2)})_{t\geq 0}$. For example, the GARCH$(1,1)$ diffusion limit satisfies

$$dG_t = \sigma_t \, dW_t^{(1)}, \quad d\sigma_t^2 = \theta(\gamma - \sigma_t^2)dt + \rho\sigma_t^2 \, dW_t^{(2)}, \quad t \geq 0. \tag{1.1}$$

The behaviour of this diffusion limit is therefore rather different from that of the GARCH process itself since the volatility process $(\sigma_t^2)_{t\geq 0}$ evolves independently of the driving process $(W_t^{(1)})_{t\geq 0}$ in the first of the equations (1.1).

Another approach is via the stochastic volatility model of Barndorff-Nielsen and Shephard (2001a, 2001b) in which the volatility process σ^2 is an Ornstein–Uhlenbeck (O-U) process driven by a non-decreasing Lévy process and G satisfies an equation of the form $dG_t = \mu dt + \sigma_t \, dW_t$, where W is a Brownian motion independent of the Lévy process. The autocorrelation function of the Lévy-driven O-U volatility process has the form $\rho(h) = \exp(-c|h|)$ for some $c > 0$, but this class can be extended by specifying the volatility to be a superposition of O-U processes as in Barndorff–Nielsen (2001), or a Lévy-driven CARMA (continuous-time ARMA) process as in Brockwell (2004). As in Nelson's diffusion, the process G is again driven by two independent noise processes and the volatility process σ^2 evolves independently of the process W in the equation for G.

A continuous-time analog of the GARCH(1,1) process, denoted COGARCH(1,1), has recently been constructed and studied by Klüppelberg et al. (2004). Their construction is based on the explicit representation of the volatility of the discrete-time GARCH$(1,1)$ process to obtain a continuous-time analog. Since no such representation exists for higher-order discrete-time GARCH processes, a different approach is needed to construct higher-order continuous-time analogs. In this thesis we do this by specifying a system of Lévy-driven stochastic differential equations for the processes G and σ^2. If the volatility process σ^2 is strictly stationary we refer to the process G as a COGARCH(p,q) process. In the special case $p = q = 1$ we recover the COGARCH$(1,1)$ process of Klüppelberg et al. (2004). In general we obtain a class of processes G with uncorrelated increments but for which the corresponding volatility and squared increment processes exhibit a broad range of autocorrelation functions. The volatility process has the autocorrelation function of a continuous-time ARMA process.

It is not clear how the approach outlined above leading to the COGARCH$(1,1)$ process can be generalised to higher order GARCH processes in continuous time. In particular, the recursion corresponding to (2.1) (see below) cannot be solved easily and generalised to a continuous time setting. In this thesis we adopt a different but related approach which allows us to define a continuous time GARCH process of order (p,q) with $1 \leq p \leq q$. The process is driven by a single Lévy process and, when $p = q = 1$, it reduces to the COGARCH$(1,1)$ process. It will therefore be referred to as a COGARCH(p,q) process. While the COGARCH$(1,1)$ process is restricted to have decreasing ACF, for higher orders this is not necessarily the case and we can obtain damped oscillatory behaviour.

The dissertation is organised as follows:

In Chapter 1, we present some background information on ARMA, CARMA and discrete-time GARCH processes.

Most of Chapter 2 is the published work of Brockwell, Chadraa and Lindner (2006). In Section 2.2, we specify a system of stochastic differential equations for the COGARCH(p,q) process G and its associated volatility process, which we shall denote by V. This is directly motivated by the corresponding structure of the discrete-time GARCH(p,q) process. We then show that the solution of these equations coincides with that of the COGARCH$(1,1)$ equations if $p = q = 1$. Notation and definitions used throughout the paper are given at the end of Section 2.2.

In Section 2.3, we give sufficient conditions for the existence of a strictly stationary volatility process. In the COGARCH$(1,1)$ case, these are exactly the necessary and sufficient conditions obtained by Klüppelberg et al. (2004). More detailed results are given in the special case when the driving Lévy process is compound-Poisson. The proofs rely on the fact that the state vector of the COGARCH(p,q) process, sampled at uniformly spaced discrete times, satisfies a multivariate random recurrence equation.

In Section 2.4, we focus on the autocorrelation structure of the stationary volatility process. Just as the discrete-time GARCH volatility process has the autocorrelation function of an ARMA process, the COGARCH volatility process has the autocorrelation function of a CARMA process.

Section 2.5 deals with conditions which ensure positivity of the volatility, while the autocorrelation structure of the squared increments of the COGARCH process itself is obtained in Section 2.6. The results are illustrated with simulations in Section 2.7.

In Section 3.3, we show that when the driving Lévy process is compound Poisson and $p = q = 2$, then the state process and the squared increment of the COGARCH process are strongly mixing with exponential rate.

In Chapter 3, we propose a least-squares estimation algorithm for the parameters of a COGARCH$(2, 2)$ process, making use of the property that the autocorrelation function of the squared increments of the COGARCH(p, q) process is that of an ARMA(q, q) process. The squared increment process of the COGARCH$(2, 2)$ is strongly mixing which ensures that the least-squares estimators (LSE) are strongly consistent and are asymptotically normal. Finally, the COGARCH$(2, 2)$ model is fitted to a high-frequency data set.

1.2 ARMA processes

The process $\{Y_n, n = 0, 1, \ldots\}$ is said to be an *ARMA(p, q) process* with real-valued parameters $\{\phi_1, \ldots, \phi_p; \theta_1, \ldots, \theta_q; \sigma\}$ if it is a stationary solution of the equations

$$Y_n - \phi_1 Y_{n-1} - \cdots - \phi_p Y_{n-p} = \sigma(Z_n + \theta_1 Z_{n-1} + \cdots + \theta_q Z_{n-q}), \qquad (1.2)$$

where $\sigma > 0$, $\phi_p \neq 0$, $\theta_q \neq 0$, $\{Z_n\} \sim$WN(0,1), i.e. a sequence of uncorrelated random variables with mean zero and variance 1 and the polynomials $\phi(z) := 1 - \phi_1 z - \ldots - \phi_p z^p$ and $\theta(z) := 1 + \theta_1 z + \cdots + \theta_q z^q$ have no common zeroes. The process $\{Y_n\}$ is said to be an ARMA(p, q) process with mean μ if $\{Y_n - \mu\}$ is an ARMA(p, q) process. If we suppose in addition that the sequence $\{Z_n\}$ is an independent and identically distributed sequence then $\{Y_n\}$ is said to be a *strict* ARMA process. In this case $\{Y_n\}$ is a *strictly stationary* process, i.e. for each positive integer j and for each $(n_1, \ldots, n_j)$, the joint distribution of $(Y_{n_1+h}, \ldots, Y_{n_j+h})$ is independent of h.

It is more convenient to use the more concise form of (1.2)

$$\phi(B)Y_n = \sigma\theta(B)Z_n, \tag{1.3}$$

where B is the backward shift operator ($B^j Y_n = Y_{n-j}$, $B^j Z_n = Z_{n-j}$, $j = 0, \pm 1, \ldots$). The time series $\{Y_n\}$ is said to be an *autoregressive process of order* p (or AR(p)) if $\theta(z) \equiv 1$, and a *moving-average process of order* q (or MA(q)) if $\phi(z) \equiv 1$. A stationary solution $\{Y_n\}$ of equations (1.2) exists (and is also the unique stationary solution) if and only if

$$\phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p \neq 0 \quad \text{for all } |z| = 1. \tag{1.4}$$

An ARMA(p, q) process $\{Y_n\}$ is *causal* if there exist constants $\{\psi_j\}$ such that $\sum_{j=0}^{\infty} |\psi_j| < \infty$ and

$$Y_n = \sum_{j=0}^{\infty} \sigma\psi_j Z_{n-j} \tag{1.5}$$

for all n. Causality is equivalent to the condition

$$\phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p \neq 0 \quad \text{for all } |z| \leq 1. \tag{1.6}$$

An ARMA(p, q) process $\{Y_n\}$ is *invertible* if there exist constants $\{\pi_j\}$ such that $\sum_{j=0}^{\infty} |\pi_j| < \infty$ and

$$\sigma Z_n = \sum_{j=0}^{\infty} \pi_j Y_{n-j}$$

for all t. Invertibility is equivalent to the condition

$$\theta(z) = 1 + \theta_1 z + \ldots + \theta_q z^q \neq 0 \quad \text{for all } |z| \leq 1. \tag{1.7}$$

The process $\{Y_n\}$ has an equivalent *state-space representation*, given by

$$Y_n = \sigma\boldsymbol{\theta}'\mathbf{X}_n \tag{1.8}$$

and

$$\mathbf{X}_{n+1} - \mathbf{\Phi}\mathbf{X}_n = \mathbf{e}Z_{n+1} \tag{1.9}$$

where

$$\mathbf{\Phi} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \phi_r & \phi_{r-1} & \phi_{r-2} & \cdots & \phi_1 \end{bmatrix}, \quad \mathbf{e} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{\theta} = \begin{bmatrix} \theta_{r-1} \\ \theta_{r-2} \\ \vdots \\ \theta_1 \\ \theta_0 \end{bmatrix},$$

$r := \max(p, q+1)$, $\theta_0 := 1$, $\theta_j := 0$ for $j > q$ and $\phi_j := 0$ for $j > p$. The causality condition (1.6) implies that the eigenvalues of the matrix $\mathbf{\Phi}$ are all less than 1 in absolute value so that the state-vector has the stationary solution

$$\mathbf{X}_n = \sum_{j=0}^{\infty} \mathbf{\Phi}^j \, \mathbf{e} \, Z_{n-j} \tag{1.10}$$

which, with (1.8), immediately yields $EY_n = 0$ and

$$\gamma_Y(h) := E[Y_{n+h}Y_n] = \sigma^2 \boldsymbol{\theta}' \mathbf{\Phi}^{|h|} \Xi \, \boldsymbol{\theta} \tag{1.11}$$

where $\Xi = E[\mathbf{X}_n\mathbf{X}_n'] = \sum_{j=0}^{\infty} \mathbf{\Phi}^j \mathbf{e}\mathbf{e}' \mathbf{\Phi}'^j$.

Parallel to the *time domain* representation (1.5) of $\{Y_n\}$, there is a *spectral* or *frequency domain* representation,

$$Y_n = \int_{-\pi}^{\pi} \sigma \frac{\theta(e^{-i\omega})}{\phi(e^{-i\omega})} e^{i\omega n} dZ(\omega), \tag{1.12}$$

where $\{Z(\omega), -\pi \le \omega \le \pi\}$ is an orthogonal increment process with mean zero and $E|dZ(\omega)|^2 = d\omega/(2\pi)$, and the ACVF has the corresponding spectral representation

$$\gamma_Y(h) = \int_{-\pi}^{\pi} f(\omega) e^{ih\omega} d\omega,$$

where the spectral density function f is given by

$$f(\omega) = \frac{\sigma^2}{2\pi} \left| \frac{\theta(e^{-i\omega})}{\phi(e^{-i\omega})} \right|^2, \quad -\pi \le \omega \le \pi.$$

Because the spectral density of an ARMA process is a ratio of trigonometric polynomials, it is often called a *rational spectral density.*

In the case when $q < p$ and the reciprocals, $\xi_1, \ldots, \xi_p$, of the zeroes of the polynomial $\phi(z)$ are distinct, the ACVF can be written as

$$\gamma_Y(h) = -\sigma^2 \sum_{j=1}^{p} \frac{\theta(\xi_j)\theta(\xi_j^{-1})}{\phi(\xi_j)\phi'(\xi_j^{-1})} \xi_j^{|h|+1}. \tag{1.13}$$

1.3 Lévy processes

The volatility of a COGARCH(p, q) process has the autocovariance structure of a CARMA process. This section is to provide the basic facts about the Lévy processes which will drive the CARMA processes. For more information on Lévy processes we refer to the books by Applebaum (2004), Bertoin (1996) and Sato (1999).

Definition 1.1. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \le t \le \infty}, P)$ be a filtered probability space, where $\mathcal{F}_0$ contains all the P-null sets of $\mathcal{F}$ and $(\mathcal{F}_t)$ is right-continuous. An adapted process $L := \{L_t,\, t \ge 0\}$ with $L_0 = 0$ a.s. is a real valued Lévy process if

(*L1*) L has independent increments, i.e., $L_t - L_s$ is independent of $\mathcal{F}_s, 0 \le s < t \le \infty$,

(*L2*) L has stationary increments, i.e., $L_t - L_s$ has the same distribution as $L_{t-s}, 0 \le s < t \le \infty$,

(*L3*) L is continuous in probability, i.e., for all $\varepsilon > 0$ and all $t \ge 0$,

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

Every Lévy process has a unique modification which is càdlàg (right-continuous with left limits) and also a Lévy process. We shall therefore assume that our Lévy process has these properties. The characteristic function of L, $\varphi_t(\theta) = E(\exp(i\theta L_t))$, has the form,

$$\varphi_t(\theta) = \exp(t\xi(\theta)), \quad \theta \in \mathbb{R},$$

where $\xi(\theta)$ is often called *characteristic exponent* or *Lévy symbol* and satisfies the following Lévy -Khinchin formula,

$$\xi(\theta) = i\gamma_L\theta - \tau_L^2\frac{\theta^2}{2} + \int_{\mathbb{R}} \left(e^{i\theta x} - 1 - i\theta x \mathbf{1}_{|x|\leq 1}\right) d\nu_L(x). \tag{1.14}$$

The triplet $\gamma_L \in \mathbb{R}$, $\tau_L^2 \geq 0$ and ν_L uniquely determines the distribution of L and is called the *characteristic triplet* of L.

The measure ν_L on $\mathbb{R}$ is called the Lévy measure. As usual, the Lévy measure ν_L is required to satisfy

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

and $\nu_L(0) = 0$. From the Lévy -Khinchin formula, we see that, in general, a Lévy process can be decomposed into three parts: a constant drift part, a Brownian motion part, and a pure jump part. If A is a Borel subset of $\{x : |x| > \epsilon\}$ for some $\epsilon > 0$, then the number of jumps with sizes in A, occurring in any time interval of length $t > 0$, has the Poisson distribution with mean $t\nu(A)$. If ν is a finite measure, i.e. $\nu(\mathbb{R}_0) = \int_{\mathbb{R}_0} \nu(dx) < \infty$, then almost all paths of L have a finite number of jumps on every compact interval and the process is said to have *finite activity*. Otherwise, if $\nu(\mathbb{R}_0) = \infty$, then an infinite number of jumps occur in any interval of positive length with probability one and the process is said to have *infinite activity*. As we shall see in the examples below, Poisson processes and compound Poisson processes have finite activity.

Example 1.2 (Brownian motion). The Lévy process B is a *Brownian motion* if ν_B is a zero measure, $EB_t = \gamma_B t$ and $\mathrm{Var}(B_t) = \tau_B^2 t$. Hence, the Lévy symbol of a Brownian motion is given by

$$\xi(\theta) = i\gamma_B\theta - \tau_B^2\frac{\theta^2}{2},$$

with the characteristic triplet $(\gamma_B, \tau_B^2, 0)$. When $\gamma_B = 0$ and $\tau_B^2 = 1$, it is called a *standard Brownian motion*. A simulated sample path of a standard Brownian motion is shown in Figure 1. For a detailed discussion of Brownian motion, we refer to Karatzas and Shreve (1991).

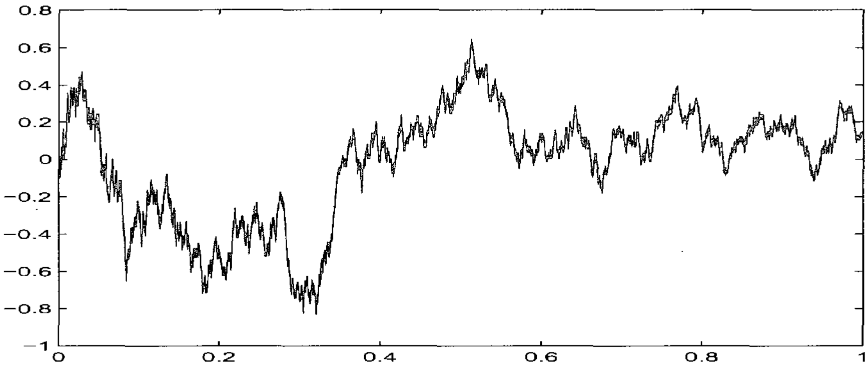

Figure 1: A simulated sample path of a standard Brownian motion.

Example 1.3 (Poisson process). Let $(\Gamma_n)_{n\in\mathbb{N}}$ be a strictly increasing sequence of stopping times, then the *counting process* N defined by

$$N_t := \sum_{i=1}^{\infty} I_{[\Gamma_i,\infty)}(t)$$

for each $t \geq 0$ is an adapted process and if further N has independent and stationary increment, it is called a *Poisson process*. For each $t \geq 0$ and some $\lambda > 0$, N_t is Poisson distributed with parameter λt. The parameter λ is called the *jump rate* of N. The differences $T_{n+1} := \Gamma_{n+1} - \Gamma_n$ are called *sojourn times*; T_n measures the duration that the Poisson process sojourns in state n. The Lévy symbol of a Poisson process N_t with jump rate λ is given by

$$\xi(\theta) = \int_{\mathbb{R}} (e^{i\theta x} - 1 - i\theta x I_{(-1,1)}(x))\lambda\delta_1(dx) = \lambda(e^{i\theta} - 1),$$

and the characteristic triplet is $(0, 0, \lambda\delta_1)$, where δ_1 denotes the Dirac measure with total mass 1 concentrated at the point 1. From the characteristic function, it immediately follows that $EN_t = \lambda t$ and $\text{Var}N_t = \lambda t$. The sample path of a Poisson process is piece-wise constant with discontinuities of size one at random points $(\Gamma_n)_{n\in\mathbb{N}}$.

Example 1.4 (Compound Poisson process). Let $(Y_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables with distribution function F_Y, independent of a Poisson process

$N = (N_t)_{t \geq 0}$. The *compound Poisson* process L_t is defined as

$$L_t = \sum_{i=1}^{N_t} Y_i, \quad t \geq 0.$$

The Lévy symbol of the compound Poisson process L_t is given by

$$\xi(\theta) = \int_{\mathbb{R}} (e^{i\theta x} - 1)\lambda F_Y(dx) = i\theta\gamma_L + \int_{\mathbb{R}} (e^{i\theta x} - 1 - i\theta I_{(-1,1)}(x))\lambda F_Y(dx),$$

where $\gamma_L = \lambda \int_{|x|<1} x F_Y(dx)$. Hence the characteristic triplet is $(\gamma_L, 0, \lambda F_Y)$. The sample path of a compound Poisson process is piece-wise constant with discontinuities at random points $(\Gamma_n)_{n \in \mathbb{N}}$, but with random jumps with distribution F_Y. A simulated sample path of a compound Poisson process with parameter $\lambda = 25$ and standard normal jumps is shown in Figure 2.

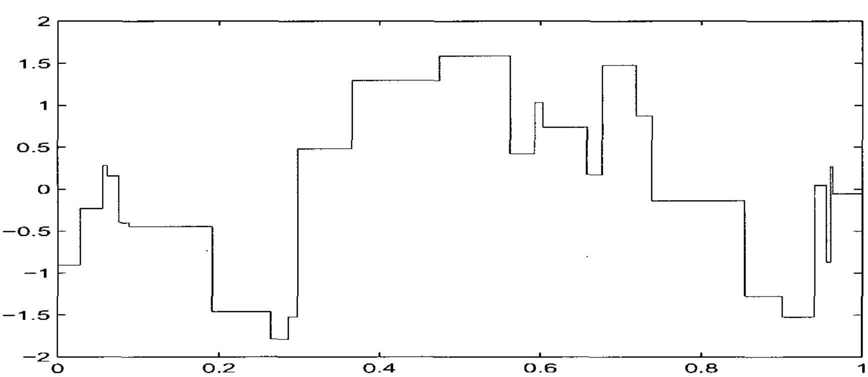

Figure 2: Simulated sample path of a compound Poisson process with parameter $\lambda = 25$ and standard normal jumps.

1.4 CARMA processes

A natural analogue, in continuous time, of the stochastic difference equation (1.3) is the stochastic differential equation,

$$a(D)Y(t) = \sigma b(D)DL(t), \quad t \geq 0, \tag{1.15}$$

where σ is a strictly positive scale parameter, D denotes differentiation with respect to t, $a(z) := z^p + a_1 z^{p-1} + \cdots + a_p$, $b(z) := b_0 + b_1 z + \cdots + b_{p-1} z^{p-1}$, and the coefficients b_j satisfy $b_q = 1$ and $b_j = 0$ for $q < j < p$. To avoid trivial and easily eliminated complications we shall assume that $a(z)$ and $b(z)$ have no common factor.

The continuous-time analogue of the driving noise terms Z_n in equation (1.2) are the increment of the process L. We shall assume that L is a Lévy process on $(-\infty, \infty)$, i.e. a process with homogeneous independent increments, continuous in probability, with càdlàg sample-paths and $L(0) = 0$. We shall also restrict attention to second-order Lévy processes, i.e. those satisfying the condition $EL(1)^2 < \infty$, and suppose, without further loss of generality, that $\mathrm{Var}L(t) = t$ and $EL(t) = \mu t$ for some $\mu \in \mathbb{R}$. The increments of L on disjoint intervals of equal length are then independent and identically distributed random variables with finite variance and some infinitely divisible distribution which could, for example, be Gaussian, gamma, compound Poisson, inverse Gaussian or one of many other possibilities.

Since the derivatives on the right of equation (1.15) do not exist in the usual sense, we write the equation in state-space form,

$$Y(t) = \sigma \mathbf{b}'\mathbf{X}(t), \tag{1.16}$$

and

$$d\mathbf{X}(t) - A\mathbf{X}(t)dt = \mathbf{e}\, dL(t), \tag{1.17}$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_p & -a_{p-1} & -a_{p-2} & \dots & -a_1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-2} \\ b_{p-1} \end{bmatrix}, \quad \mathbf{e} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

with solution satisfying

$$\mathbf{X}(t) = e^{A(t-s)}\mathbf{X}(s) + \int_s^t e^{A(t-u)}\mathbf{e}dL(u), \quad \text{for all } t > s. \tag{1.18}$$

In the Gaussian case L is Brownian motion, equation (1.17) is interpreted as an Itô equation and the integral in equation (1.18) is defined as in Protter (2004).

If we restrict attention to causal solutions, i.e. if we make the assumption that $\mathbf{X}(s)$ is independent of $\{L(t) - L(s), t > s\}$ for every s, then necessary and sufficient conditions for the existence of a strictly stationary process $\mathbf{X}$ satisfying equation (1.18) (see Brockwell et. al. (2005)) are

$$\Re(\lambda_r) < 0, \quad r = 1, \ldots, p, \tag{1.19}$$

and, under these conditions, the stationary solution must satisfy

$$\mathbf{X}(t) \text{ is distributed as } \int_0^\infty e^{Au}\mathbf{e}dL(u) \text{ for all } t, \tag{1.20}$$

where $\lambda_1, \ldots, \lambda_p$ are the eigenvalues of A (which are the zeroes of the autoregressive polynomial $a(z)$). Condition (1.20) specifies the stationary marginal distribution of $\mathbf{X}(t)$ and condition (1.19) is the continuous-time analogue of the causality condition (1.6).

If we assume that the conditions (1.19) and (1.20) hold, and let $s \to -\infty$ in equation (1.18) we find that

$$\mathbf{X}(t) = \int_{-\infty}^t e^{A(t-u)}\mathbf{e}dL(u). \tag{1.21}$$

Conversely if $\mathbf{X}(t)$ is defined by equation (1.21) then $\mathbf{X}$ is a strictly stationary causal process satisfying equation (1.18).

We now define the strictly stationary causal CARMA process by equation (1.16) with $\mathbf{X}$ given by equation (1.21). Thus

$$Y(t) = \int_{-\infty}^t \sigma\mathbf{b}'e^{A(t-u)}\mathbf{e}dL(u). \tag{1.22}$$

From this equation we find that $EY(t) = \sigma\mu b_0/a_p$ (where $\mu = EL(1)$) and

$$\gamma_Y(h) := \mathrm{Cov}[Y(t+h), Y(t)] = \sigma^2\mathbf{b}'e^{A|h|}\mathbf{\Sigma}\mathbf{b}, \tag{1.23}$$

where $\mathbf{\Sigma} = E[\mathbf{X}(t)\mathbf{X}(t)'] = \int_0^\infty e^{Au}\mathbf{e}\mathbf{e}'e^{A'y}dy.$

Although the expression (1.23) for the autocovariance function has an awkward appearance, it is possible to evaluate the matrices e^{Ah} and Σ explicitly in terms of the eigenvalues of the matrix A using its Jordan decomposition. The eigenvalues of A, as already indicated, are just the roots of the equation $a(z) = 0$ and the right eigenvector of A corresponding to the eigenvalue λ is $[1 \;\; \lambda \;\; \cdots \;\; \lambda^{p-1}]'$.

From the resulting expression for the autocovariance function of the process $\{Y(t)\}$ the spectral density

$$f_Y(\omega) := \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega h} \gamma_Y(h)\, dh$$

is found as

$$f_Y(\omega) = \frac{1}{2\pi} \frac{|b(i\omega)|^2}{|a(i\omega)|^2}, \quad -\infty < \omega < \infty. \tag{1.24}$$

A much simpler form of (1.23) can be derived from (1.24) by contour integration. Thus, substituting from (1.24) into the relation

$$\gamma_Y(h) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega h} f_Y(\omega)\, d\omega,$$

and changing the variable of integration from ω to $z = i\omega$, we find that

$$\gamma_Y(h) = S, \quad h \geq 0,$$

where S is the sum of residues of $e^{zh}[b(z)b(-z)]/[a(z)a(-z)]$ in the left half of the complex plane. This gives the general expression

$$\gamma_Y(h) = \sum_{\{\lambda : a(\lambda)=0\}} \frac{1}{(m-1)!} \left[\frac{d^{m-1}}{dz^{m-1}} \frac{(z-\lambda)^m e^{z|h|} b(z) b(-z)}{a(z)a(-z)} \right]_{z=\lambda}, \tag{1.25}$$

where m is the multiplicity of the root λ of $a(z) = 0$. In the case when the roots are distinct, equation (1.25) simplifies to

$$\gamma_Y(h) = \sum_{r=1}^{p} \frac{b(\lambda_r)b(-\lambda_r)}{a'(\lambda_r)a(-\lambda_r)} e^{\lambda_r |h|}. \tag{1.26}$$

For the process to be *minimum phase* the roots of $b(z) = 0$ must have real parts less that or equal to zero. (This corresponds to *invertibility* for discrete time ARMA processes.) For a one-to-one correspondence between the second order properties of $\{Y(t)\}$ and the parameters $a_1, \ldots, a_p, b_1, \ldots, b_q$, it is necessary to restrict $b(z)$ to satisfy the minimum phase condition and to be non-negative in a neighborhood of $z = 0$. Every CARMA process has such a representation, and in order to avoid trivial ambiguities, we shall identify CARMA process whose coefficients in this minimum phase representation are the same.

Example 1.5. The Gaussian CARMA(1,0) (or CAR(1)) process is the simplest continuous-time ARMA process. It is defined formally as a stationary solution of the first-order stochastic differential equation

$$(D + a)Y(t) = bDW(t), \tag{1.27}$$

where $a > 0$ and $\{W(t)\}$ is a standard Brownian motion. From(1.16) and(1.21) with $\mathbf{b} = b$ and with L replaced by standard Brownian motion W on $(-\infty, \infty)$, we have

$$Y(t) = b \int_{-\infty}^{t} e^{-a(t-u)} dW(u). \tag{1.28}$$

For $s \le t$

$$Y(t) = e^{-a(t-s)} Y(s) + b \int_{s}^{t} e^{-a(t-u)} dW(u). \tag{1.29}$$

This shows that the process is Markovian, i.e. that the distribution of $Y(t)$ given $Y(u), u \le s$, is the same as the distribution of $Y(t)$ given $Y(s)$. It also shows that the conditional mean and variance of $Y(t)$ given $Y(s)$ are

$$E[Y(t)|Y(s)] = e^{-a(t-s)} Y(s)$$

and

$$\mathrm{Var}[Y(t)|Y(s)] = \frac{b^2}{2a}\left(1 - e^{-2a(t-s)}\right).$$

Now we can use the Markov property and the moments of the stationary distribution to write down the Gaussian likelihood of observations $y(t_1), \ldots, y(t_n)$ of a CAR(1) process satisfying (1.27). This is the joint density of $(Y(t_1), \ldots, Y(t_n))'$ at $(y(t_1), \ldots, y(t_n))'$, which can be expressed as the product of the stationary density at

$y(t_1)$ and the transition densities of $Y(t_i)$ given $Y(t_{i-1}) = y(t_{i-1})$, $i = 2, \ldots, n$. The joint density is therefore given by

$$g(y(t_1), \ldots, y(t_n); a, b) = \prod_{i=1}^{n} \frac{1}{\sqrt{v_i}} f\Big(\frac{y(t_i) - m_i}{\sqrt{v_i}}\Big), \tag{1.30}$$

where f is the standard normal density, $m_1 = 0$, $v_1 = b^2/(2a)$, and for $i > 1$,

$$m_i = e^{-a(t_i - t_{i-1})} y(t_{i-1})$$

and

$$v_i = \frac{b^2}{2a}\big(1 - e^{-2a(t_i - t_{i-1})}\big).$$

Notice that the times t_i appearing in (1.30) are quite arbitrarily spaced. It makes the CAR(1) process useful for modeling irregularly spaced data.

If the observations are regularly spaced, say $t_i = i$, $i = 1, \ldots, n$, then the joint density g is exactly the same as the joint density of observations of the discrete-time Gaussian AR(1) process

$$X_n = e^{-a} X_{n-1} + Z_n,$$

where $\{Z_n\}$ is a white noise process with mean 0 and variance $b^2(1 - e^{-2a})/(2a)$. This shows that the "embedded" discrete-time process $\{Y(i), i = 1, 2, \ldots\}$ of the CAR(1) process is a discrete-time AR(1) with coefficient e^{-a}.

Example 1.6. The CARMA(2,1) process with parameters a_1, a_2, b_0, b_1 is a stationary solution of the stochastic differential equation

$$D^2Y(t) + a_1 DY(t) + a_2 Y(t) = (b_0 + b_1 D) DW(t), \quad t \geq 0. \tag{1.31}$$

In order for a causal stationary solution to exist it is necessary that the roots of the equation

$$\lambda^2 + a_1 \lambda + a_2 = 0 \tag{1.32}$$

have negative real parts. For a minimum phase solution we also require that $b_0 \geq 0$ and $b_1 > 0$. In the case when (1.32) has two distinct complex conjugate roots

$$\lambda_1 = \alpha + i\beta \text{ and } \lambda_2 = \alpha - i\beta, \quad \alpha < 0, \beta > 0 \tag{1.33}$$

it follows from (1.26) that the autocovariance function of $\{Y(t)\}$ is

$$\gamma_Y(h) = \gamma_Y(0)e^{\alpha|h|}\left[\cos(\beta h) + \sin(\beta|h|)\frac{\alpha(b_1^2 a_2 - b)^2)}{\beta(b_1^2 a_2 + b_0^2)}\right], \tag{1.34}$$

where

$$\gamma_Y(0) = \frac{b_1^2 a_2 + b_0^2}{2a_1 a_2}. \tag{1.35}$$

Note that if $\lambda = \alpha + i\beta$ is any complex number with non-zero imaginary part and if $\omega = a + ib$ is any complex number, then

$$\kappa(h) = \omega e^{\lambda|h|} + \bar{\omega}e^{\bar{\lambda}|h|}, \quad -\infty < h < \infty \tag{1.36}$$

is the autocovariance function of a CARMA(2,1) or CARMA(2,0) process if and only if

$$\alpha < 0, \tag{1.37}$$

$$a > 0 \tag{1.38}$$

and

$$|\beta| \leq a|\alpha|/|b|. \tag{1.39}$$

Condition (1.38) expresses the obvious requirement that $\kappa(0) > 0$ and a straightforward calculation shows that (1.39) is then necessary and sufficeient for the Fourier transform

$$f(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ih\omega}\kappa(h)dh$$

to be non-negative for all $\omega \in (-\infty, \infty)$.

1.5 GARCH processes

In early econometric models, the variance of the daily percent change in asset price was assumed to be a constant. However, many econometric time-series models exhibit periods of unusually large volatility followed by a periods of relative tranquility. In such circumstances, the assumption of constant variance (homoscedasticity) is inappropriate. Engle (1982) introduced the *ARCH(p) process* $\{X_n\}$ as

$$X_n = \sqrt{v_n}\,\varepsilon_n \tag{1.40}$$

where $\{\varepsilon_n\}$ is an i.i.d. $N(0,1)$ sequence and v_n is the (positive) function of $\{X_k, k < n\}$, defined by

$$v_n = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{n-i}^2, \tag{1.41}$$

with $\alpha_0 > 0$ and $\alpha_i \geq 0$, $i = 1, \ldots, p$. The name ARCH signifies autoregressive conditional heteroscedasticity. v_n is the conditional variance of X_n given $\{X_k, k < n\}$.

The ARCH(p, q) process given by (1.40) and (1.41) has been extended by Bollerslev (1986). The generalized ARCH(p, q) model, called GARCH(p, q), is the process in which the variance equation (1.41) is replaced by

$$v_n = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{n-i}^2 + \sum_{j=1}^{q} \beta_j v_{n-j}, \tag{1.42}$$

with $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, \ldots, p$ and $\beta_j \geq 0$, $j = 1, \ldots, q$.

The simplest example from the class of GARCH models is the ARCH(1) process. The recursions (1.40) and (1.41) give

$$\begin{aligned}
X_n^2 &= \alpha_0 \varepsilon_n^2 + \alpha_1 X_{n-1}^2 \varepsilon_n^2 \\
&= \alpha_0 \varepsilon_n^2 + \alpha_1 \alpha_0 \varepsilon_n^2 \varepsilon_{n-1}^2 + \alpha_1^2 X_{n-2}^2 \varepsilon_n^2 \varepsilon_{n-1}^2 \\
&= \cdots \\
&= \alpha_0 \sum_{j=0}^{k} \alpha_1^j \varepsilon_n^2 \varepsilon_{n-1}^2 \cdots \varepsilon_{n-j}^2 + \alpha_1^{k+1} X_{n-k-1}^2 \varepsilon_n^2 \varepsilon_{n-1}^2 \cdots \varepsilon_{n-k}^2.
\end{aligned}$$

If $|\alpha_1| < 1$ then the last term converges to 0 with probability one as $k \to \infty$. The first term converges with probability one by Proposition 3.1.1 of Brockwell and Davis (1991), and hence

$$X_n^2 = \alpha_0 \sum_{j=0}^{\infty} \alpha_1^j \varepsilon_n^2 \varepsilon_{n-1}^2 \cdots \varepsilon_{n-j}^2. \tag{1.43}$$

From (1.43) we immediately find that

$$E[X_t^2] = \alpha_0/(1-\alpha_1). \tag{1.44}$$

Since

$$X_n = \varepsilon_n \sqrt{\alpha_0\Big(1+\sum_{j=1}^{\infty}\alpha_1^j\varepsilon_{n-1}^2\cdots\varepsilon_{n-j}^2\Big)}. \tag{1.45}$$

it is clear that $\{X_n\}$ is strictly stationary and hence, since $E[X_n^2] < \infty$, also (weakly) stationary. Also it is straight-forward to find that

$$E[X_n] = E(E[X_n|\varepsilon_k, k<n]) = 0, \tag{1.46}$$

$$E[X_{n+h}X_n] = E(E[X_{n+h}X_n|\varepsilon_k, k<n+h]) = 0, \quad h>0. \tag{1.47}$$

Thus the ARCH(1) process with $|\alpha_1| < 1$ is strictly stationary white noise. However, it is not an i.i.d. sequence, since from (1.40) and (1.41),

$$E[X_n^2|X_{n-1}] = (\alpha_0+\alpha_1 X_{n-1}^2)E[\varepsilon_n^2|X_{n-1}] = \alpha_0+\alpha_1 Z_{n-1}^2 \neq EX_n^2.$$

This also shows that $\{X_n\}$ is not Gaussian, since strictly stationary Gaussian white noise is necessarily i.i.d.. From (1.43) it is easy to show that $E[X_n^4]$ is finite if and only if $3\alpha_1^2 < 1$. If $E[X_n^4] < \infty$, the squared process $Y_n = X_n^2$ has the same autocovariance function as the AR(1) process $W_n = \alpha_1 W_{n-1} + \varepsilon_n$, a result that extends also to ARCH(p) processes. It can be shown that for every α_1 in the interval $(0,1)$, $E[Z_n^{2k}] = \infty$ for some integer k. This indicates the "heavy-tailed" nature of the marginal distribution of X_n.

The ARCH(p) process is conditionally Gaussian, in the sense that for given values of $\{X_k, k=n-1,\ldots,n-p\}$, X_n is Gaussian with known distribution. This makes it easy to write down the likelihood of $Z_{p+1},\ldots,Z_n$ conditional on $\{Z_1,\ldots,Z_p\}$ and hence, by numerical maximization, to compute conditional maximum likelihood estimates for the parameters. For example, the conditional likelihood of observations $\{z_2,\ldots,z_n\}$ of the ARCH(1) process, given $Z_1 = z_1$, is

$$L = \prod_{t=2}^{n}\frac{1}{\sqrt{2\pi(\alpha_0+\alpha_1 z_{t-1}^2)}}\exp\Big\{-\frac{z_t^2}{2(\alpha_0+\alpha_1 z_{t-1}^2)}\Big\}.$$

2 COGARCH processes

2.1 The COGARCH$(1,1)$ process

The construction of the COGARCH$(1,1)$ process due to Klüppelberg et al. (2004) starts from the defining equations of the discrete time GARCH$(1,1)$ process $(\xi_n)_{n\in\mathbb{N}_0}$,

$$\xi_n = \varepsilon_n\sigma_n, \quad \sigma_n^2 = \alpha_0 + \alpha_1\xi_{n-1}^2 + \beta_1\sigma_{n-1}^2, \quad n \in \mathbb{N}_0, \tag{2.1}$$

where α_0, α_1, and β_1 are all strictly positive and $(\varepsilon_n)_{n\in\mathbb{N}_0}$ is a sequence of iid (independent and identically distributed) random variables with mean zero and variance 1. The recursions (2.1) can be solved to give

$$\begin{aligned}
\sigma_n^2 &= \alpha_0 \sum_{i=0}^{n-1}\prod_{j=i+1}^{n-1}(\beta_1+\alpha_1\varepsilon_j^2) + \sigma_0^2\prod_{j=0}^{n-1}(\beta_1+\alpha_1\varepsilon_j^2)\\
&= \left(\sigma_0^2 + \alpha_0\int_0^n \exp\Big[-\sum_{j=0}^{\lfloor s\rfloor}\log(\beta_1+\alpha_1\varepsilon_j^2)\Big]\,ds\right)\exp\Big[\sum_{j=0}^{n-1}\log(\beta_1+\alpha_1\varepsilon_j^2)\Big],
\end{aligned}$$

where $\lfloor s\rfloor$ denotes the integer part of $s \in \mathbb{R}$. The COGARCH$(1,1)$ equations are then obtained by replacing the driving noise sequence $(\varepsilon_n)_{n\in\mathbb{N}_0}$ by the jumps $(\Delta L_t = L_t - L_{t-})_{t\geq 0}$ of a Lévy process. More precisely, observing that

$$\sum_{j=0}^{n-1}\log(\beta_1+\alpha_1\varepsilon_j^2) = n\log\beta_1 + \sum_{j=0}^{n-1}\log(1+(\alpha_1/\beta_1)\varepsilon_j^2)$$

for $\beta_1 > 0$, and writing η for $-\log\beta_1$, ω_0 for α_0 and ω_1 for α_1, leads to the equations

$$\begin{aligned}
dG_t &= \sigma_t\,dL_t, \quad t>0, \quad G_0 = 0, &(2.2)\\
\sigma_t^2 &= \left(\sigma_0^2 + \omega_0\int_0^t e^{X_s}\,ds\right)e^{-X_{t-}}, \quad t\geq 0, &(2.3)
\end{aligned}$$

where the auxiliary process $(X_t)_{t\geq 0}$ is defined as

$$X_t := \eta t - \sum_{0<s\leq t}\log\left(1+\omega_1 e^{\eta}(\Delta L_s)^2\right). \tag{2.4}$$

Here, $\omega_0 > 0$, $\omega_1 \geq 0$, $\eta > 0$ and σ_0^2 is independent of $(L_t)_{t\geq 0}$. The COGARCH$(1,1)$ process is the solution G of these equations and, under specified conditions on the coefficients and the distribution of σ_0^2, the volatility process σ^2 is strictly stationary and G has stationary increments.

As shown in Proposition 3.2 of Klüppelberg et al. (2004), the process $(\sigma_t^2)_{t\geq 0}$ satisfies the stochastic differential equation

$$d\sigma_{t+}^2 = \omega_0 dt + \sigma_t^2 e^{X_{t-}} d(e^{-X_t}), \quad t > 0,$$

and

$$\sigma_t^2 = \omega_0 t - \eta \int_0^t \sigma_s^2 ds + \omega_1 e^{\eta} \sum_{0<s<t} \sigma_s^2 (\Delta L_s)^2 + \sigma_0^2, \quad t \geq 0.$$

The COGARCH$(1,1)$ process with stationary volatility has been shown to have many of the features of the discrete time GARCH$(1,1)$ process. As shown in Klüppelberg et al. (2004, 2006), the COGARCH$(1,1)$ process has uncorrelated increments, while the autocorrelation functions of the volatility σ^2 and of the squared increments of G decay exponentially. Further, the COGARCH$(1,1)$ process has heavy tails and volatility clusters at high levels (see Klüppelberg et al. (2006) and Fasen et al. (2006)). While the volatility clustering can be also achieved in the stochastic volatility model of Barndorff-Nielsen and Shephard if the driving Lévy process has regularly varying tails (see Fasen et al. (2006) or Fasen (2004)), this is impossible for the GARCH diffusion (1.1). For an overview of extremes of stochastic volatility models, see Fasen et al. (2006). Also, observe that many of the features of the COGARCH$(1,1)$ process can be obtained in a more general setting, as in Lindner and Maller (2005).

2.2 The COGARCH(p, q) equations

Let $(\varepsilon_n)_{n\in\mathbb{N}_0}$ be an i.i.d. sequence of random variables. Let $p, q \geq 0$. Then the GARCH(p,q) process $(\xi_n)_{n\in\mathbb{N}_0}$ is defined by the equations,

$$\begin{aligned} \xi_n &= \sigma_n \varepsilon_n, \\ \sigma_n^2 &= \alpha_0 + \alpha_1 \xi_{n-1}^2 + \ldots + \alpha_p \xi_{n-p}^2 + \beta_1 \sigma_{n-1}^2 + \ldots + \beta_q \sigma_{n-q}^2, \quad n \geq s, \end{aligned} \tag{2.5}$$

where $s := \max(p, q)$, $\sigma_0^2, \ldots, \sigma_{s-1}^2$ are i.i.d. and independent of the i.i.d. sequence $(\varepsilon_n)_{n \geq s}$, and $\xi_n = G_{n+1} - G_n$ represents the increment at time n of the log asset price process $(G_n)_{n \in \mathbb{N}_0}$. Note that the continuous-time GARCH process will be a model for $(G_t)_{t \geq 0}$ and not for its increments as in discrete-time.

Equation (2.5) shows that the volatility process $(\sigma_n^2)_{n \in \mathbb{N}_0}$ can be viewed as a "self-exciting" ARMA$(q, p-1)$ process driven by the noise sequence $(\sigma_{n-1}^2 \varepsilon_{n-1}^2)_{n \in \mathbb{N}}$. Motivated by this observation, we will define a continuous time GARCH model for the log asset price process $(G_t)_{t \geq 0}$ of order (p, q) by

$$dG_t = \sigma_t \, dL_t, \quad t > 0, \quad G_0 = 0,$$

where $(\sigma_t^2)_{t \geq 0}$ is a CARMA$(q, p-1)$ process driven by a suitable replacement for the discrete time driving noise sequence $(\sigma_{n-1}^2 \varepsilon_{n-1}^2)_{n \in \mathbb{N}}$.

The state-space representation of a Lévy-driven CARMA$(q, p-1)$ process $(\psi_t)_{t \geq 0}$ with driving Lévy process L, location parameter c, moving average coefficients $\alpha_1, \ldots,$ α_p, autoregressive coefficients $\beta_1, \ldots, \beta_q$ and $q \geq p$ is (see Brockwell (2001)),

$$\begin{aligned} \psi_t &= c + \mathbf{a}' \zeta_t, \\ d\zeta_t &= \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 1 \\ -\beta_q & -\beta_{q-1} & -\beta_{q-2} & \cdots & -\beta_1 \end{bmatrix} \zeta_t dt + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} dL_t, \end{aligned}$$

where $\mathbf{a}' = [\alpha_1, \ldots, \alpha_q]$, $\alpha_j := 0$ for $j > p$, and the coefficient matrix in the last equation is $-\beta_1$ if $q = 1$. (The CARMA$(q, p-1)$ process, $(\psi_t)_{t \geq 0}$, is a strictly stationary solution of these equations, which exists under conditions found in Brockwell (2001).) In order to obtain a continuous-time analog of the equation (2.5) we suppose that the volatility process $(\sigma_t^2)_{t \geq 0}$ has the state-space representation of a CARMA$(q, p-1)$ process in which the driving Lévy process (L_t) is replaced by a continuous-time analog of the driving process $(\sigma_{n-1}^2 \varepsilon_{n-1}^2)_{n \in \mathbb{N}}$ in (2.5).

The increments of the driving process in continuous time should correspond to the increments of the discrete-time process,

$$R_n^{(d)} := \sum_{i=0}^{n-1} \xi_i^2 = \sum_{i=0}^{n-1} \sigma_i^2 \varepsilon_i^2.$$

We therefore replace the innovations ε_n by the jumps ΔL_t of a Lévy process $(L_t)_{t\geq 0}$ to obtain the continuous-time analogue,

$$R_t := \sum_{0<s\leq t} \sigma_{s-}^2 (\Delta L_s)^2, \quad t > 0.$$

If L has no Gaussian part (i.e. $\tau_L^2 = 0$ in (1.14)), we recognise R as the quadratic covariation of G, i.e.

$$R_t = \sum_{0<s\leq t} \sigma_{s-}^2 (\Delta L_s)^2 = \int_0^t \sigma_{s-}^2 d[L,L]_s = [G,G]_t.$$

If L has a Gaussian part, then $\sum_{0<s\leq t} (\Delta L_s)^2 = [L,L]^{(d)}$, the *discrete part of the quadratic covariation*, and we have in general

$$R_t = \int_0^t \sigma_{s-}^2 d[L,L]_s^{(d)}, \quad \text{i.e.} \quad dR_t = \sigma_{t-}^2 \, d[L,L]_t^{(d)}.$$

The COGARCH(p,q) equations will now be obtained by specifying that the volatility process $V(=\sigma^2)$ should satisfy continuous-time ARMA equations driven by the process R defined above. Provided V is non-negative almost surely (conditions for which are given in Section 2.5), we can define a process G by the equations $G_0 = 0$ and $dG_t = \sqrt{V_t} dL_t$. Under conditions ensuring that V is also strictly stationary, we refer to G as a COGARCH(p,q) process. As we shall see, when $p = q = 1$, the solution of the COGARCH equations coincides with that of the COGARCH$(1,1)$ equations (1.3)–(1.5) of Klüppelberg et al. (2004). (The parameters $\beta_1, \ldots, \beta_q$ and $\alpha_1, \ldots, \alpha_p$ in the following definition should not be confused with the parameters denoted by the same symbols in the defining equation (2.5) of the discrete-time GARCH process.)

Definition 2.1. (The COGARCH(p,q) equations) If p and q are integers such that $q \geq p \geq 1$, $\alpha_0 > 0$, $\alpha_1, \ldots, \alpha_p \in \mathbb{R}$, $\beta_1, \ldots, \beta_q \in \mathbb{R}$, $\alpha_p \neq 0$, $\beta_q \neq 0$, and

$\alpha_{p+1} = \ldots = \alpha_q = 0$, we define the $(q \times q)$–matrix B and the vectors $\mathbf{a}$ and $\mathbf{e}$ by

$$B = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ -\beta_q & -\beta_{q-1} & -\beta_{q-2} & \ldots & -\beta_1 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{q-1} \\ \alpha_q \end{bmatrix}, \quad \mathbf{e} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

with $B := -\beta_1$ if $q = 1$. Then if $L = (L_t)_{t \geq 0}$ is a Lévy process with non-trivial Lévy measure, we define the (left-continuous) *volatility process* $V = (V_t)_{t \geq 0}$ with parameters B, $\mathbf{a}$, α_0 and driving Lévy process L by

$$V_t = \alpha_0 + \mathbf{a}'\mathbf{Y}_{t-}, \quad t > 0, \quad V_0 = \alpha_0 + \mathbf{a}'\mathbf{Y}_0,$$

where the *state process* $\mathbf{Y} = (\mathbf{Y}_t)_{t \geq 0}$ is the unique càdlàg solution of the stochastic differential equation

$$d\mathbf{Y}_t = B\mathbf{Y}_{t-}\,dt + \mathbf{e}(\alpha_0 + \mathbf{a}'\mathbf{Y}_{t-})\,d[L, L]_t^{(d)}, \quad t > 0, \tag{2.6}$$

with initial value $\mathbf{Y}_0$, independent of the driving Lévy process $(L_t)_{t \geq 0}$. If the process $(V_t)_{t \geq 0}$ is strictly stationary and non-negative almost surely, we say that $G = (G_t)_{t \geq 0}$, given by

$$dG_t = \sqrt{V_t}\,dL_t, \quad t > 0, \quad G_0 = 0,$$

is a *COGARCH(p, q) process* with parameters B, $\mathbf{a}$, α_0 and driving Lévy process L .
□

That there is in fact a unique solution of (2.6) for any starting random vector $\mathbf{Y}_0$ follows from standard theorems on stochastic differential equations (e.g. Protter (2004), Chapter V, Theorem 7). The stochastic integrals are interpreted with respect to the filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$, which is defined to be the smallest right-continuous filtration such that $\mathcal{F}_0$ contains all the P-null sets of $\mathcal{F}$, $(L_t)_{t \geq 0}$ is adapted and $\mathbf{Y}_0$ is $\mathcal{F}_0$ measurable.

Without restrictions on α_0, $\mathbf{a}$ and B, the process V is not necessarily non-negative, in which case G is undefined. Conditions which ensure that V is non-negative will be discussed in Section 2.5. In particular, it will be shown that if $\mathbf{a}'e^{Bt}\mathbf{e} \geq 0$ for all $t \geq 0$ and $\mathbf{Y}_0$ is such that V is strictly stationary, then V is non-negative with probability one. Even if V takes negative values however, the process is of some interest in its own right and many of our results for V are valid without the non-negativity restriction.

Conditions for stationarity of V are discussed in Section 2.3.

We next show that if $p = q = 1$, the solution of the COGARCH equations in Definition 2.1 coincides with the solution of the COGARCH$(1,1)$ equations of Klüppelberg et al. (2004).

Theorem 2.2. *Suppose that $p = q = 1$, and that α_0, α_1 and β are all strictly positive. Then the processes $(G_t)_{t\geq 0}$ and $(V_t)_{t\geq 0}$ of Definition 2.1 are respectively the processes $(G_t)_{t\geq 0}$ and $(\sigma_t^2)_{t\geq 0}$ defined by (2.2) - (2.4), with parameters $\omega_0 = \alpha_0\beta_1$, $\omega_1 = \alpha_1 e^{-\beta_1}$ and $\eta = \beta_1$.*

Proof. From $d\mathbf{Y}_t = -\beta_1 \mathbf{Y}_t\, dt + V_t\, d[L,L]_t^{(d)}$ and $V_{t+} = \alpha_0 + \alpha_1 \mathbf{Y}_t$ follows that

$$dV_{t+} = \alpha_1\, d\mathbf{Y}_t = -\alpha_1\beta_1 \frac{V_t - \alpha_0}{\alpha_1}\, dt + \alpha_1 V_t\, d[L,L]_t^{(d)},$$

and hence that

$$V_{t+} = \alpha_0\beta_1 t - \beta_1 \int_0^t V_s\, ds + \alpha_1 \sum_{0<s\leq t} V_s(\Delta L_s)^2 + V_0.$$

But this equation is also satisfied by the volatility process $(\sigma_t^2)_{t\geq 0}$ of (2.3) when $\omega_0 = \alpha_0\beta_1$, $\eta = \beta_1$ and $\omega_1 = \alpha_1 e^{-\beta_1}$, as shown in Proposition 3.2 of Klüppelberg et al. (2004), and uniqueness of the solution gives the claim. $\square$

We conclude this section with a few definitions and some notation which will be used throughout the paper.

Definition 2.3. Let $\mathbf{a}$ and B be as in Definition 2.1. Then the *characteristic polynomials* associated with $\mathbf{a}$ and B are given by

$$\begin{aligned} a(z) &:= \alpha_1 + \alpha_2 z + \ldots + \alpha_p z^{p-1}, \quad z \in \mathbb{C}, \\ b(z) &:= z^q + \beta_1 z^{q-1} + \ldots + \beta_q, \quad z \in \mathbb{C}. \end{aligned}$$

The eigenvalues of the matrix B (which are exactly the zeroes of b) will be denoted by $\lambda_1, \ldots, \lambda_q$ and assumed to be ordered in such a way that

$$\Re\lambda_q \leq \Re\lambda_{q-1} \leq \ldots \leq \Re\lambda_1$$

(where $\Re\lambda_i$ denotes the real part of λ_i). Further, define

$$\lambda := \lambda(B) := \Re\lambda_1.$$

For the rest of the paper, convergence in probability will be denoted by "$\xrightarrow{P}$", uniform convergence on compacts in probability by "$\xrightarrow{ucp}$", and equality in distribution by "$\stackrel{d}{=}$". For $x \in \mathbb{R}$ we shall write $\log^+(x)$ for $\log(\max\{1, x\})$. The transpose of a column vector $\mathbf{c} \in \mathbb{C}^q$ will be denoted by $\mathbf{c}'$. If $\|\cdot\|$ is a vector norm in $\mathbb{C}^q$, then the natural matrix norm of the $(q \times q)$-matrix C is defined as $\|C\| = \sup_{\mathbf{c}\in\mathbb{C}^q\setminus\{0\}} \frac{\|C\mathbf{c}\|}{\|\mathbf{c}\|}$. Correspondingly, for $r \in [1, \infty]$ we denote by $\|\cdot\|_r$ both the vector L^r-norm and the associated natural matrix norm. Recall that the natural matrix norms of the L^1, L^2 and L^∞ vector norms are the column-sum norm, the spectral norm and the row-sum norm, respectively.

The $(q \times q)$–identity matrix will be denoted by I_q or simply I, and the canonical vector $(0, \ldots, 0, 1, 0, \ldots, 0)'$, with i^{th} component equal to 1, by $\mathbf{e}_i$. For $\mathbf{e}_q$ we simply write $\mathbf{e}$. By $\operatorname{diag}(\lambda_1, \ldots, \lambda_q)$ we mean the diagonal $(q \times q)$-matrix with these entries on the diagonal. The Kronecker product of two $(q \times q)$–matrices A and B will be denoted by $A \otimes B$, and by $\operatorname{vec}(A)$ we denote the column vector in $\mathbb{C}^{q^2}$ which arises from A by stacking the columns of A in a vector (starting with the first column). For the properties of the Kronecker product we refer to Lütkepohl (1996).

2.3 Stationarity conditions

In this section we consider conditions under which the volatility process $(V_t)_{t\geq 0}$ specified in Definition 2.1 is strictly stationary. The parameters B, $\mathbf{a}$ and α_0, and the state process $(\mathbf{Y}_t)_{t\geq 0}$ are as specified in Definition 2.1. The condition (2.8) established in Theorem 2.4 below is necessary and sufficient for stationarity in the special case $p = q = 1$. For larger values of p and q it is sufficient only, but not unduly restrictive since there is a rich class of models satisfying the condition. Without serious loss of generality we shall assume that the matrix B can be diagonalised. Since the only eigenvectors corresponding to the eigenvalue λ_i are constant multiples of $[1, \lambda_i, \lambda_i^2, \ldots, \lambda_i^{q-1}]'$, this is equivalent to the assumption that the eigenvalues of B are distinct. Let S be a matrix such that $S^{-1}BS$ is a diagonal matrix, for example,

$$S = \begin{bmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_q \\ \vdots & \cdots & \vdots \\ \lambda_1^{q-1} & \cdots & \lambda_q^{q-1} \end{bmatrix}. \tag{2.7}$$

(For this particular choice, $S^{-1}BS = \operatorname{diag}(\lambda_1, \ldots, \lambda_q)$.)

Theorem 2.4. *Let $(\mathbf{Y}_t)_{t\geq 0}$ be the state process of the COGARCH(p, q) process with parameters B, $\mathbf{a}$ and α_0. Suppose that all the eigenvalues of B are distinct. Let L be a Lévy process with non-trivial Lévy measure ν_L, and suppose there is some $r \in [1, \infty]$ such that*

$$\int_{\mathbb{R}} \log(1 + \|S^{-1}\mathbf{e}\mathbf{a}'S\|_r\, y^2)\, d\nu_L(y) < -\lambda = -\lambda(B), \tag{2.8}$$

for some matrix S such that $S^{-1}BS$ is diagonal. Then $\mathbf{Y}_t$ converges in distribution to a finite random variable $\mathbf{Y}_\infty$, as $t \to \infty$. It follows that if $\mathbf{Y}_0 \stackrel{d}{=} \mathbf{Y}_\infty$, then $(\mathbf{Y}_t)_{t\geq 0}$ and $(V_t)_{t\geq 0}$ are strictly stationary.

Remark 2.5. *(a)* If $(V_t)_{t\geq 0}$ is the volatility of a COGARCH(1, 1) process with parameters $B = -\beta_1 < 0$, $\alpha_0 > 0$ and $\alpha_1 > 0$, then $\|S^{-1}\mathbf{e}\mathbf{a}'S\|_r = \alpha_1$ and, as already indicated, the condition (2.8) is necessary and sufficient for the existence of a strictly stationary COGARCH$(1,1)$ volatility process. (See Klüppelberg et al. (2004), Theorem 3.1.)

(*b*) For general $q \geq 2$, the quantity $\|S^{-1}\mathbf{e}\mathbf{a}S\|_r$ depends on the specific choice of S and on r. Observe that it is sufficient to find *some* S and *some* r such that (2.8) holds. $\square$

The proof of Theorem 2.4 will make heavy use of the general theory of multivariate random recurrence equations, as discussed by Bougerol and Picard (1992), Kesten (1973), and Brandt (1986) (in the one–dimensional case). The proof is given after the proof of Theorem 2.8, since equation (2.15) will be needed in the proof of Theorems 2.4 and 2.6.

The COGARCH state vector satisfies such a multivariate random recurrence equation, as indicated in the following theorem.

Theorem 2.6. *Let $(\mathbf{Y}_t)_{t\geq 0}$ be the state process of the COGARCH(p, q) process with parameters B, $\mathbf{a}$ and α_0, and driving Lévy process L. Then there exists a family $(J_{s,t}, \mathbf{K}_{s,t})_{0\leq s\leq t}$ of random $(q\times q)$–matrices $J_{s,t}$ and random vectors $\mathbf{K}_{s,t}$ in $\mathbb{R}^q$ such that*

$$\mathbf{Y}_t = J_{s,t}\mathbf{Y}_s + \mathbf{K}_{s,t}, \quad 0 \leq s \leq t. \tag{2.9}$$

Further, the distribution of $(J_{s,t}, \mathbf{K}_{s,t})$ depends only on $t-s$, $(J_{s_1,t_1}, \mathbf{K}_{s_1,t_1})$ and $(J_{s_2,t_2}, \mathbf{K}_{s_2,t_2})$ are independent for $0 \leq s_1 \leq t_1 \leq s_2 \leq t_2$, and for $0 \leq s \leq u \leq t$,

$$J_{s,t} = J_{u,t}\, J_{s,u}. \tag{2.10}$$

If additionally the conditions of Theorem 2.4 hold, then the distribution of the vector $\mathbf{Y}_\infty$ is for any $h > 0$ the unique solution of the random fixed point equation,

$$\mathbf{Y}_\infty \stackrel{d}{=} J_{0,h}\mathbf{Y}_\infty + \mathbf{K}_{0,h}, \tag{2.11}$$

with $\mathbf{Y}_\infty$ independent of $(J_{0,h}, \mathbf{K}_{0,h})$ on the right hand side of (2.11).

Remark 2.7. (*a*) The stationarity condition (2.8) is easy to check. However, as the proofs of Theorems 2.4 and 2.6 show, a weaker stationarity condition is the existence of a vector norm $\|\cdot\|$ and $t_0 > 0$ such that J_{0,t_0} and $\mathbf{K}_{0,t_0}$ satisfy the conditions

$$E \log \|J_{0,t_0}\| < 0 \quad \text{and} \quad E \log^+ \|\mathbf{K}_{0,t_0}\| < \infty. \tag{2.12}$$

By (2.10), $E \log \|J_{0,t_0}\| < 0$ is equivalent to the requirement that there is a strictly positive value of t_1 such that the Lyapunov exponent of the iid sequence $(J_{t_1 n, t_1(n+1)})_{n \in \mathbb{N}_0}$, i.e.

$$\lim_{n \to \infty} \frac{1}{n} E \left(\log \|J_{t_1(n-1), t_1 n} \cdots J_{0,t_1}\| \right) = \inf_{n \in \mathbb{N}} \left(\frac{1}{n} E \left(\log \|J_{t_1(n-1), t_1 n} \cdots J_{0,t_1}\| \right) \right),$$

(which is independent of the specific norm) is strictly negative. As shown by Bougerol and Picard (1992), provided $E \log^+ \|J_{0,t_1}\| < \infty$, $E \log^+ \|\mathbf{K}_{0,t_1}\| < \infty$ and a certain irreducibility condition holds, then strict negativity of the Lyapunov exponent is not only sufficient but also necessary for the existence of stationary solutions of such random recurrence equations.

(*b*) The conditions of Theorem 2.4 imply the conditions (2.12) with the matrix norm defined as the natural norm $\|A\|_{B,r} = \|S^{-1} A S\|_r$, corresponding to the vector norm,

$$\|\mathbf{c}\|_{B,r} := \|S^{-1} \mathbf{c}\|_r, \quad \mathbf{c} \in \mathbb{C}^q. \tag{2.13}$$

Observe, however, that the conditions of Theorem 2.4 are in general not necessary for stationarity. For example, using methods similar to those in the proofs of Theorems 2.4 and 2.6, it can be shown that for any vector norm $\|\cdot\|$, and for $t \geq 0$,

$$\|J_{0,t}\| \leq \|e^{Bt}\| + e^{\|B\| t} \exp\left(\sum_{0 < s \leq t} \log \left(1 + (\Delta L_s)^2 \|\mathbf{e}\mathbf{a}'\|\right) \right) \|\mathbf{e}\mathbf{a}'\| \sum_{0 < s \leq t} (\Delta L_s)^2.$$

Now if $\lambda(B) < 0$, then $\|e^{Bt}\| \to 0$ as $t \to \infty$, and (2.12) can be satisfied without assuming that all the eigenvalues of B are distinct, but choosing $\|\mathbf{a}\|$ sufficiently small and imposing certain integrability conditions on L. We shall not pursue this argument here as the conditions of Theorem 2.4 will be sufficient for our purposes.
□

The matrices $J_{s,t}$ and the vector $\mathbf{K}_{s,t}$ of Theorem 2.6 will be constructed explicitly when L is compound-Poisson, and in the general case will be obtained as the limit of the corresponding quantities for compound-Poisson-driven processes. In the compound-Poisson case we shall show that the stationary state vector satisfies a distributional fixed point equation which is much easier to handle than (2.11). Also, we compare the stationary distribution of $\mathbf{Y}_\infty$ with the stationary distribution of the state vector when sampled at the jump times of the Lévy process. This is the content of the next theorem.

Theorem 2.8. *(a) Let $(\mathbf{Y}_t)_{t\geq 0}$ be the state process of a COGARCH(p, q) process with parameters B, $\mathbf{a}$ and α_0. Suppose that the Lévy measure ν_L of the driving Lévy process L is finite and write the compound Poisson process $[L, L]^{(d)}$ in the form*

$$[L, L]_t^{(d)} = \sum_{0<s\leq t} (\Delta L_s)^2 = \sum_{i=1}^{N(t)} Z_i,$$

where $N(t)$ is the number of jumps of L in the time interval $(0, t]$ and Z_i is the square of the i^{th} jump size. Let T_1 denote the time at which the first jump occurs and let T_j, $j = 2, 3, \ldots$, be the time intervals between the $(j-1)^{\text{th}}$ and j^{th} jumps. Further, let (T_0, Z_0) be independent of $(T_i, Z_i)_{i\in\mathbb{N}}$ with the same distribution as (T_1, Z_1). For $i \in \mathbb{N}_0$, let

$$\begin{aligned} C_i &= (I + Z_i \mathbf{e}\mathbf{a}')e^{BT_i}, \\ \mathbf{D}_i &= \alpha_0 Z_i \mathbf{e}, \end{aligned}$$

and $\Gamma_n = \sum_{i=1}^n T_i$ (where $\Gamma_0 := 0$). Then the discrete time process $(\mathbf{Y}_{\Gamma_n})_{n\in\mathbb{N}_0}$ satisfies the random recurrence equation

$$\mathbf{Y}_{\Gamma_{n+1}} = C_{n+1}\mathbf{Y}_{\Gamma_n} + \mathbf{D}_{n+1}, \quad n \in \mathbb{N}_0. \tag{2.14}$$

Further, for any $t > 0$,

$$\begin{aligned}
\mathbf{Y}_t &= e^{B(t-\Gamma_{N(t)})}\Big[\mathbf{1}_{\{N(t)\neq 0\}}\mathbf{D}_{N(t)} + \sum_{i=0}^{N(t)-2} C_{N(t)}\cdots C_{N(t)-i}\mathbf{D}_{N(t)-i-1} \\
&\quad + C_{N(t)}\cdots C_1\mathbf{Y}_0\Big] \\
&\stackrel{d}{=} e^{B(t-\Gamma_{N(t)})}\Big[\mathbf{1}_{\{N(t)\neq 0\}}\,\mathbf{D}_1 + \sum_{i=1}^{N(t)-1} C_1\cdots C_i\mathbf{D}_{i+1} + C_1\cdots C_{N(t)}\mathbf{Y}_0\Big].
\end{aligned} \tag{2.15}$$

(*b*) *Assume additionally that the conditions of Theorem 2.4 are satisfied. Then the infinite sum $\sum_{i=0}^{\infty} C_1\cdots C_i\mathbf{D}_{i+1}$ converges almost surely absolutely to a random vector $\widehat{\mathbf{Y}}$, which has the stationary distribution of the sequence $(\mathbf{Y}_{\Gamma_n})_{n\in\mathbb{N}_0}$. The stationary state vector $\mathbf{Y}_\infty$ satisfies*

$$\mathbf{Y}_\infty \stackrel{d}{=} e^{BT}\widehat{\mathbf{Y}}, \tag{2.16}$$

where T is independent of $(T_i, Z_i)_{i\in\mathbb{N}_0}$ and has the distribution of T_1. Further, $\mathbf{Y}_\infty$ is the unique solution in distribution of the distributional fixed point equation

$$\mathbf{Y}_\infty \stackrel{d}{=} Q\mathbf{Y}_\infty + \mathbf{R}, \tag{2.17}$$

where $\mathbf{Y}_\infty$ is independent of $(Q, \mathbf{R})$ and

$$\begin{aligned}
Q &:= e^{BT_0}(I + Z_0\mathbf{e}\mathbf{a}'), \\
\mathbf{R} &:= \alpha_0\, Z_0\, e^{BT_0}\mathbf{e}.
\end{aligned}$$

The fixed point equation (2.17) will play a crucial role in the determination of the covariance matrix of $\mathbf{Y}_\infty$, studied in the next section.

Proof of Theorem 2.8. (*a*) It follows from (2.6) that $\mathbf{Y}_t$ satisfies $d\mathbf{Y}_t = B\mathbf{Y}_t\,dt$ for $t \in [\Gamma_n, \Gamma_{n+1})$, so that

$$\mathbf{Y}_t = e^{B(t-\Gamma_n)}\mathbf{Y}_{\Gamma_n}, \quad t \in [\Gamma_n, \Gamma_{n+1}), \quad n \in \mathbb{N}_0. \tag{2.18}$$

At time Γ_{n+1} a jump of size $\mathbf{e}(\alpha_0 + \mathbf{a}'\mathbf{Y}_{\Gamma_{n+1}-})Z_{n+1}$ occurs, so that

$$\begin{aligned} \mathbf{Y}_{\Gamma_{n+1}} &= \mathbf{Y}_{\Gamma_{n+1}-} + \mathbf{e}(\alpha_0 + \mathbf{a}'\mathbf{Y}_{\Gamma_{n+1}-})Z_{n+1} \\ &= (I + Z_{n+1}\mathbf{e}\mathbf{a}')\mathbf{Y}_{\Gamma_{n+1}-} + \alpha_0 Z_{n+1}\mathbf{e} \\ &= C_{n+1}\mathbf{Y}_{\Gamma_n} + \mathbf{D}_{n+1}, \quad n \in \mathbb{N}_0, \end{aligned}$$

which is (2.14). Solving this recursion gives

$$\mathbf{Y}_{\Gamma_n} = \mathbf{D}_n + \sum_{i=0}^{n-2} C_n \cdots C_{n-i}\mathbf{D}_{n-i-1} + C_n \cdots C_1 \mathbf{Y}_0, \quad n \in \mathbb{N},$$

and the first equality in (2.15) follows from this and $\mathbf{Y}_t = e^{B(t-\Gamma_{N(t)})}\mathbf{Y}_{\Gamma_{N(t)}}$. The second equality in (2.15) is a consequence of the fact that the infinite random element $(N(t), \Gamma_{N(t)}, C_{N(t)}, \mathbf{D}_{N(t)}, \ldots, C_1, \mathbf{D}_1, 0, 0, \ldots)$ has the same distribution as $(N(t), \Gamma_{N(t)}, C_1, \mathbf{D}_1, \ldots, C_{N(t)}, \mathbf{D}_{N(t)}, 0, 0, \ldots)$; indeed, for any $n \in \mathbb{N}_0$ and $c \geq 0$ the random vectors $(C_1, \mathbf{D}_1), \ldots, (C_n, \mathbf{D}_n)$ are iid and depend on the restriction $\{N(t) = n, \ \Gamma_{N(t)} \geq c\}$ only in terms of $\sum_{i=1}^n T_i$ and T_{n+1}, but not on the T_i, $i = 1, \ldots, n$, individually.

(*b*) Let S be such that $S^{-1}BS =: \Lambda$ is diagonal and define the vector norm $\|\mathbf{c}\|_{B,r} = \|S^{-1}\mathbf{c}\|_r$ as in equation (2.13), so that the associated natural matrix norm is $\|A\|_{B,r} = \|S^{-1}AS\|_r$. Then we have for $t \geq 0$,

$$\left\|e^{Bt}\right\|_{B,r} = \left\|S\, e^{\Lambda t}\, S^{-1}\right\|_{B,r} = \left\|e^{\Lambda t}\right\|_r = e^{\lambda t}. \tag{2.19}$$

This gives $\|C_1\|_{B,r} \leq (1 + Z_1\|\mathbf{e}\mathbf{a}'\|_{B,r})e^{\lambda T_1}$ and $\|\mathbf{D}_1\|_{B,r} = \alpha_0\|\mathbf{e}\|_{B,r}\, Z_1$, so that, using $\nu_{[L,L]}([x,\infty)) = \nu_L(\{y \in \mathbb{R} : |y| \geq \sqrt{x}\}$ for $x \geq 0$,

$$\begin{aligned} E \log \|C_1\|_{B,r} &\leq \lambda\, E(T_1) + E \log(1 + Z_1\|\mathbf{e}\mathbf{a}'\|_{B,r}) \\ &= \frac{\lambda}{\nu_L(\mathbb{R})} + \frac{1}{\nu_L(\mathbb{R})} \int_{(0,\infty)} \log(1 + \|\mathbf{e}\mathbf{a}'\|_{B,r}\, y^2)\, d\nu_L(y) < 0 \end{aligned}$$

by (2.8) and

$$E \log^+(Z_1) = \frac{1}{\nu_L(\mathbb{R})} \int_{\mathbb{R}} \log^+(y^2)\, d\nu_L(y) < \infty.$$

From the general theory of random recurrence equations this implies the almost sure absolute convergence of $\sum_{i=0}^{\infty} C_1 \cdots C_i \mathbf{D}_{i+1}$ to $\widehat{\mathbf{Y}}$ which has the stationary distribution of $(\mathbf{Y}_{\Gamma_n})_{n\in\mathbb{N}}$, see e.g. Bougerol and Picard (1992).

To prove (2.16), for $m \in \mathbb{N}$ let

$$\widehat{\mathbf{Y}}_m := \sum_{i=0}^{m-1} C_1 \cdots C_i \mathbf{D}_{i+1} + C_1 \cdots C_m \mathbf{Y}_0,$$

and

$$\mathbf{Y}_{t,m} := e^{B(t-\Gamma_{N(t)})} \widehat{\mathbf{Y}}_m, \quad t \geq 0.$$

Since the random variable $(t - \Gamma_{N(t)})$ is asymptotically independent of $T_1, Z_1, \ldots, T_m, Z_m$ (for $t \to \infty$, m fixed), it follows that $e^{B(t-\Gamma_{N(t)})}$ is asymptotically independent of $\widehat{\mathbf{Y}}_m$, and hence $\mathbf{Y}_{t,m}$ converges in distribution to $e^{BT}\widehat{\mathbf{Y}}_m$, as $t \to \infty$, where T is exponentially distributed with parameter $\nu_L(\mathbb{R})$ (e.g. Taylor and Karlin (1994), Section 7.4.4) and independent of $T_1, Z_1, \ldots, T_m, Z_m$ and hence can be chosen to be independent of $(T_i)_{i\in\mathbb{N}}$, $(Y_i)_{i\in\mathbb{N}}$ (as in the statement of the theorem). Moreover, $e^{BT}\widehat{\mathbf{Y}}_m$ converges almost surely, hence in distribution to $e^{BT}\widehat{\mathbf{Y}}$, as $m \to \infty$. Denote by $\widetilde{\mathbf{Y}}_t$ the expression in the lower line of (2.15). Then (2.16), and in particular the existence of the limit variable $\mathbf{Y}_\infty$ in the compound Poisson case, follow from (2.15) and a variant of Slutsky's Theorem (e.g. Brockwell and Davis (1991), Proposition 6.3.9), provided

$$\lim_{m\to\infty} \limsup_{t\to\infty} P(\|\widetilde{\mathbf{Y}}_t - \mathbf{Y}_{t,m}\|_{B,r} > \varepsilon) = 0, \quad \forall\, \varepsilon > 0. \tag{2.20}$$

Since $\|e^{B(t-\Gamma_{N(t)})}\|_{B,r} \leq 1$, and $\mathbf{1}_{\{N(t)\neq 0\}}\,\mathbf{D}_1 + \sum_{i=1}^{N(t)-1} C_1 \cdots C_i \mathbf{D}_{i+1} + C_1 \cdots C_{N(t)} \mathbf{Y}_0 - \widehat{\mathbf{Y}}_m$ converges almost surely, hence in probability as $t \to \infty$ to $\sum_{i=m}^{\infty} C_1 \cdots C_i \mathbf{D}_{i+1} - C_1 \cdots C_m \mathbf{Y}_0$, which itself converges almost surely to 0 as $m \to \infty$, (2.20) is true and (2.16) follows. That $\mathbf{Y}_\infty$ satisfies (2.17) is clear from (2.16), and that it is the unique solution follows from $E \log \|Q\|_{B,r} < 0$ and $E \log^+ \|\mathbf{R}\|_{B,r} < \infty$. $\square$

The proof of Theorem 2.8 (b) already showed the existence of the limit variable $\mathbf{Y}_\infty$ for the case of a driving compound Poisson process. Nevertheless, this existence will be reestablished in the proof of Theorems 2.4 and 2.6 for the general case, making use of Theorem 2.8 (a) only. We shall use an approximation argument and introduce the following notation:

Definition 2.9. Let L be a Lévy process. Then for any $\varepsilon > 0$, the *$\sqrt{\varepsilon}$-cut Lévy process* $(L_t^{(\varepsilon)})_{t\geq 0}$ is defined by

$$L_t^{(\varepsilon)} := \sum_{0<s\leq t, |\Delta L_s|\geq\sqrt{\varepsilon}} \Delta L_s, \quad t \geq 0.$$

If $(\mathbf{Y}_t)_{t\geq 0}$ is a state process of a COGARCH(p,q) process driven by L, then the COGARCH(p,q) process with the same parameters and starting vector but driving Lévy process $(L_t^{(\varepsilon)})_{t\geq 0}$ will be denoted by $(\mathbf{Y}_t^{(\varepsilon)})_{t\geq 0}$.

The quadratic covariation of $L^{(\varepsilon)}$ is given by

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

In particular, the corresponding COGARCH volatility results in being driven by a compound Poisson process. With this notation, we have the following lemma:

Lemma 2.10. *Let $(\mathbf{Y}_t)_{t\geq 0}$ be the state process of a COGARCH(p,q) process. Then $\mathbf{Y}_t^{(\varepsilon)}$ converges in ucp to $\mathbf{Y}_t$, as $\varepsilon \to 0$.*

Proof of Lemma 2.10. This is an easy consequence of perturbation results in stochastic differential equations: recalling the definition of prelocal convergence in $\underline{\underline{H}}^p$, $1 \leq p < \infty$, as in Protter (2004), page 260, it is easy to see that $[L^{(\varepsilon)}, L^{(\varepsilon)}]$ converges prelocally to $[L, L]^{(d)}$ in $\underline{\underline{H}}^p$, $1 \leq p < \infty$, as $\varepsilon \to 0$ (for example, with stopping times $T_k = k$). The claim then follows from Theorems 14 and 15 of Chapter V in Protter (2004). $\square$

Proof of Theorems 2.4 and 2.6. We shall first concentrate on (2.9) and (2.10) and then prove Theorem 2.4 and the rest of Theorem 2.6 simultaneously. Let $\varepsilon > 0$, and assume the representation

$$[L^{(\varepsilon)}, L^{(\varepsilon)}]_t = \sum_{i=1}^{N_\varepsilon(t)} Z_i^{(\varepsilon)},$$

where $L^{(\varepsilon)}$ is the $\sqrt{\varepsilon}$–cut Lévy process of Definition 2.9. Define $C_i^{(\varepsilon)}$ and $\mathbf{D}_i^{(\varepsilon)}$ similarly as in Theorem 2.8. Further, let

$$J_{0,t}^{(\varepsilon)} := e^{B(t-\Gamma^{(\varepsilon)}_{N_\varepsilon(t)})} C^{(\varepsilon)}_{N_\varepsilon(t)} \cdots C_1^{(\varepsilon)},$$

$$\mathbf{K}_{0,t}^{(\varepsilon)} := e^{B(t-\Gamma^{(\varepsilon)}_{N_\varepsilon(t)})} \Big[\mathbf{1}_{\{N_\varepsilon(t)\neq 0\}} \mathbf{D}^{(\varepsilon)}_{N_\varepsilon(t)} + \sum_{i=0}^{N_\varepsilon(t)-2} C^{(\varepsilon)}_{N_\varepsilon(t)} \cdots C^{(\varepsilon)}_{N_\varepsilon(t)-i} \mathbf{D}^{(\varepsilon)}_{N_\varepsilon(t)-i-1} \Big].$$

Then, by Theorem 2.8 (a),

$$\mathbf{Y}_t^{(\varepsilon)} = J_{0,t}^{(\varepsilon)} \mathbf{Y}_0 + \mathbf{K}_{0,t}^{(\varepsilon)}. \tag{2.21}$$

From the previous Lemma we know that $\mathbf{Y}_t^{(\varepsilon)}$ converges in ucp to $\mathbf{Y}_t$ as $\varepsilon \to 0$. Since this is true for any starting value $\mathbf{Y}_0$, it holds in particular for $\mathbf{Y}_0 = 0$, and from (2.21) follows that $\mathbf{K}_{0,t}^{(\varepsilon)}$ converges in ucp to some $\mathbf{K}_{0,t}$, as $\varepsilon \to 0$. Hence, again from (2.21) it follows that for arbitrary $\mathbf{Y}_0$,

$$J_{0,t}^{(\varepsilon)} \mathbf{Y}_0 = \mathbf{Y}_t^{(\varepsilon)} - \mathbf{K}_{0,t}^{(\varepsilon)} \stackrel{ucp}{\to} \mathbf{Y}_t - \mathbf{K}_{0,t}, \quad \text{as} \quad \varepsilon \to 0.$$

Since this holds for arbitrary $\mathbf{Y}_0$, we conclude that $J_{0,t}^{(\varepsilon)}$ converges in ucp to some $J_{0,t}$ as $\varepsilon \to 0$. From (2.21) then follows

$$\mathbf{Y}_t = J_{0,t} \mathbf{Y}_0 + \mathbf{K}_{0,t}.$$

By starting at an arbitrary time s instead of at time 0, we obtain (2.9). For example, $J_{s,t}^{(\varepsilon)}$ is given by

$$J_{s,t}^{(\varepsilon)} = e^{B(t-\Gamma^{(\varepsilon)}_{N_\varepsilon(t)})} C_{N_\varepsilon(t)} \ldots C_{N_\varepsilon(s)+2} (I + Z_{N_\varepsilon(s)+1} \mathbf{e}\mathbf{a}') e^{B(\Gamma^{(\varepsilon)}_{N_\varepsilon(s)+1} - s)}, 0 \le s \le t,$$

giving (2.10). The independence and stationarity assertions on $(J_{s,t}, \mathbf{K}_{s,t})$ are clear, since $J_{s,t}$ and $\mathbf{K}_{s,t}$ are constructed only from the segment $(L_u)_{s<u\le t}$ of the Lévy process L.

Now assume that all eigenvalues of B are distinct and that (2.8) holds. Applying (2.19) to $J_{0,t}^{(\varepsilon)}$ gives

$$\begin{aligned}
\|J_{0,t}^{(\varepsilon)}\|_{B,r} &\le \left\|e^{B(t-\Gamma^{(\varepsilon)}_{N_\varepsilon(t)})}\right\|_{B,r} \left\|C^{(\varepsilon)}_{N_\varepsilon(t)}\right\|_{B,r} \cdots \left\|C_1^{(\varepsilon)}\right\|_{B,r} \\
&\le e^{\lambda(t-\Gamma^{(\varepsilon)}_{N_\varepsilon(t)})} \prod_{i=1}^{N_\varepsilon(t)} \left((1+Z_i^{(\varepsilon)}\,\|\mathbf{ea}'\|_{B,r})e^{\lambda(\Gamma_i^{(\varepsilon)}-\Gamma_{i-1}^{(\varepsilon)})}\right) \\
&= e^{\lambda t}\exp\left(\sum_{i=1}^{N_\varepsilon(t)} \log(1+Z_i^{(\varepsilon)}\,\|S^{-1}\mathbf{ea}'S\|_r)\right) && (2.22) \\
&\le e^{\lambda t}\exp\left(\sum_{0<s\le t} \log(1+(\Delta L_s)^2\|S^{-1}\mathbf{ea}'S\|_r)\right). && (2.23)
\end{aligned}$$

Since $\|J_{0,t}\|_{B,r} \le \limsup_{\varepsilon\to 0}\|J_{0,t}^{(\varepsilon)}\|_{B,r}$, we conclude that

$$\log\|J_{0,t}\|_{B,r} \le \lambda t + \sum_{0<s\le t}\log(1+(\Delta L_s)^2\|S^{-1}\mathbf{ea}'S\|_r), \tag{2.24}$$

giving

$$E\log\|J_{0,t}\|_{B,r} \le t\left(\lambda + \int_{\mathbb{R}}\log(1+\|S\mathbf{ea}'S^{-1}\|y^2)\,d\nu_L(y)\right) < 0$$

by (2.8) (see e.g. Protter (2004), Chapter I, Theorems 36 and 38). This is the left hand inequality of (2.12). To show that $E\log^+\|\mathbf{K}_{0,t}\|_{B,r} < \infty$, observe that

$$\begin{aligned}
&\|\mathbf{K}_{0,t}^{(\varepsilon)}\|_{B,r} \\
&\le e^{\lambda(t-\Gamma^{(\varepsilon)}_{N_\varepsilon(t)})}\mathbf{1}_{\{N_\varepsilon(t)\neq 0\}}\,\alpha_0\|\mathbf{e}\|_{B,r}\,Z^{(\varepsilon)}_{N_\varepsilon(t)} \\
&\quad +\alpha_0\|\mathbf{e}\|_{B,r}\sum_{i=0}^{N_\varepsilon(t)-2} e^{\lambda(t-\Gamma^{(\varepsilon)}_{N_\varepsilon(t)-i-1})}\left(1+Z^{(\varepsilon)}_{N_\varepsilon(t)}\|\mathbf{ea}'\|_{B,r}\right)\cdots \\
&\quad \cdots\left(1+Z^{(\varepsilon)}_{N_\varepsilon(t)-i}\|\mathbf{ea}'\|_{B,r}\right)\,Z^{(\varepsilon)}_{N_\varepsilon(t)-i-1} && (2.25) \\
&\le \alpha_0\|\mathbf{e}\|_{B,r}\,\mathbf{1}_{\{N_\varepsilon(t)\neq 0\}}Z^{(\varepsilon)}_{N_\varepsilon(t)} \\
&\quad +\alpha_0\,\|\mathbf{e}\|_{B,r}\sum_{i=0}^{N_\varepsilon(t)-2}\exp\Big[\sum_{0<s\le t}\log(1+(\Delta L_s)^2\|\mathbf{ea}'\|_{B,r})\Big]Z^{(\varepsilon)}_{N_\varepsilon(t)-i-1} \\
&\le \alpha_0\|S^{-1}\mathbf{e}\|_r\exp\Big[\sum_{0<s\le t}\log(1+(\Delta L_s)^2\|S^{-1}\mathbf{ea}'S\|_r)\Big]\sum_{0<s\le t}(\Delta L_s)^2. && (2.26)
\end{aligned}$$

From this follows that

$$\log\|\mathbf{K}_{0,t}\|_{B,r} \le \log(\alpha_0\|S^{-1}\mathbf{e}\|_r) + \sum_{0<s\le t}\log(1+(\Delta L_s)^2\|S^{-1}\mathbf{ea}'S\|_r) + \log\,[L,L]_t^{(d)}.$$

The expectation of the second summand is finite as shown above, and $E(\log[L,L]_t^{(d)}) < \infty$ since $\int_{(1,\infty)} \log x \, d\nu_{[L,L]}(x) = \int_{\mathbb{R}\setminus[-1,1]} \log x^2 \, d\nu_L(x) < \infty$, showing the right hand inequality of (2.12).

Let $(J_n, \mathbf{K}_n)_{n\in\mathbb{N}}$ be an iid sequence with distribution $(J_{0,1}, \mathbf{K}_{0,1})$, independent of L and $\mathbf{Y}_0$. Let $\gamma \in [0,1)$ and $n \in \mathbb{N}$. Then it follows from (2.9) that

$$\begin{aligned}
\mathbf{Y}_{n+\gamma} &= \mathbf{K}_{n+\gamma-1,n+\gamma} + \sum_{i=0}^{n-2} J_{n+\gamma-1,n+\gamma} \cdots J_{n+\gamma-i-1,n+\gamma-i} \mathbf{K}_{n+\gamma-i-2,n+\gamma-i-1} \\
&\quad + J_{n+\gamma-1,n+\gamma} \cdots J_{\gamma,\gamma+1} \mathbf{Y}_\gamma \\
&\overset{d}{=} \mathbf{K}_1 + \sum_{i=1}^{n-1} J_1 \cdots J_i \mathbf{K}_{i+1} + J_1 \cdots J_n \mathbf{Y}_\gamma \\
&=: \mathbf{G}_n + H_n \mathbf{Y}_\gamma, \quad \text{say.}
\end{aligned}$$

Since $E \log \|J_1\|_{B,r} < 0$ and $E \log^+ \|\mathbf{K}_1\|_{B,r} < \infty$, it follows from the general theory of random recurrence equations (e.g. Bougerol and Picard (1992)) that H_n converges almost surely to 0 as $n \to \infty$ and that $\mathbf{G}_n$ converges almost surely absolutely to some random vector $\mathbf{G}$, as $n \to \infty$. Since $\mathbf{Y}$ has càdlàg paths, it follows that $\sup_{\gamma\in[0,1)} \|\mathbf{Y}_\gamma\|_{B,r}$ is almost surely finite. Hence

$$\lim_{n\to\infty} \sup_{\gamma\in[0,1)} \|H_n \mathbf{Y}_\gamma\|_{B,r} = 0 \quad a.s.,$$

and it follows that $\mathbf{Y}_t$ converges in distribution to $\mathbf{Y}_\infty := \mathbf{G}$ as $t \to \infty$. That $\mathbf{Y}_\infty$ satisfies (2.11) and is the unique solution is clear by the theory of random recurrence equations. Equations (2.11) and (2.9) then imply that if $\mathbf{Y}_0 \overset{d}{=} \mathbf{Y}_\infty$, then $\mathbf{Y}_t \overset{d}{=} \mathbf{Y}_\infty$ for all $t > 0$, showing strict stationarity of $(\mathbf{Y}_t)_{t\geq 0}$ since it is a Markov process. □

2.4 Second order properties of the volatility process

In this section $(\mathbf{Y}_t)_{t\geq 0}$ denotes the state process defined by (2.6), with parameters B, $\mathbf{a}$ and α_0 and driving Lévy process L with Lévy measure ν_L. The aim of this section is to study the autocorrelation function of the volatility process $(V_t)_{t\geq 0}$. We shall write

$$\mu := \int_{\mathbb{R}} y^2 \, d\nu_L(y) \quad \text{and} \quad \rho := \int_{\mathbb{R}} y^4 \, d\nu_L(y),$$

and, if $\mu < \infty$ (i.e. $EL_1^2 < \infty$),

$$\widetilde{B} := B + \mu \mathbf{e}\mathbf{a}'. \tag{2.27}$$

Observe that $\widetilde{B}$ has the same form as B, but with last row given by $(-\beta_q + \mu\alpha_1, \ldots, -\beta_1 + \mu\alpha_q)$. We first give sufficient conditions for the moments of $\mathbf{Y}_t$ to exist.

Proposition 2.11. *Suppose that the eigenvalues of B are distinct, $\lambda = \lambda(B) < 0$, $\|\cdot\|$ is any vector norm on $\mathbb{C}^q$ and $k \in \mathbb{N}$. Then the following results hold.*

(a) If $E|L_1|^{2k} < \infty$ and $E\|\mathbf{Y}_0\|^k < \infty$,

$$E\|\mathbf{Y}_t\|^k < \infty \quad \forall\, t \geq 0.$$

(b) If $E|L_1|^{2k} < \infty$, $r \in [1, \infty]$, S is a matrix such that $S^{-1}BS$ is diagonal and

$$\int_{\mathbb{R}} \left((1 + \|S^{-1}\mathbf{e}\mathbf{a}'S\|_r\, y^2)^k - 1\right)\, d\nu_L(y) < -\lambda k,$$

then S and r satisfy (2.8) and $E\|\mathbf{Y}_\infty\|^k < \infty$. In particular, $E(\mathbf{Y}_\infty)$ exists if

$$EL_1^2 < \infty \quad \text{and} \quad \|S^{-1}\mathbf{e}\mathbf{a}'S\|_r\, \mu < -\lambda, \tag{2.28}$$

and the covariance matrix, $\operatorname{cov}(\mathbf{Y}_\infty)$, exists if

$$EL_1^4 < \infty \quad \text{and} \quad \|S^{-1}\mathbf{e}\mathbf{a}'S\|_r^2\, \rho < 2(-\lambda - \|S^{-1}\mathbf{e}\mathbf{a}'S\|_r \mu). \tag{2.29}$$

Further, (2.29) implies (2.28), and (2.28) implies that all the eigenvalues of $\widetilde{B}$ have strictly negative real parts, in particular that $\widetilde{B}$ is invertible and $\beta_q \neq \alpha_1\mu$.

In order to prove Proposition 2.11, we will show that the state process $(\mathbf{Y}_t)_{t\geq 0}$ can be majorised by the state process of a COGARCH$(1,1)$ process, for which we can apply the moment conditions of Klüppelberg et al. (2004). We further show that under the conditions of Theorem 2.4, the stationary distribution $\mathbf{Y}_\infty$ can be approximated by stationary distributions of compound Poisson driven COGARCH processes, and that there is a majorant for this approximation. This will allow to restrict attention to compound Poisson driven processes when calculating autocorrelations, the general case following from Lebesgue's dominated convergence theorem. This is the content of the next lemma:

Lemma 2.12. *Let $(\mathbf{Y}_t)_{t\geq 0}$ be the state process of a COGARCH(p, q) process with parameters B,* $\mathbf{a}$ *and $\alpha_0 > 0$ such that all eigenvalues of B are distinct and that $\lambda = \lambda(B) < 0$. Let $r \in [1, \infty]$, S such that $S^{-1}BS$ is diagonal, and denote by $\|\cdot\|_{B,r}$ the vector norm defined in* (2.13). *Further, denote by $(\overline{\mathbf{Y}}_t)_{t\geq 0}$ the state process of a COGARCH(1, 1) process with (1×1)-matrix λ, vector $\|\mathbf{e}\mathbf{a}'\|_{B,r} \in \mathbb{R}^1$, scaling parameter $\alpha_0\|\mathbf{e}\|_{B,r} > 0$ and initial state vector $\overline{\mathbf{Y}}_0 := \|\mathbf{Y}_0\|_{B,r}$. Then*

$$\|\mathbf{Y}_t\|_{B,r} \leq \overline{\mathbf{Y}}_t, \quad t \geq 0. \tag{2.30}$$

If (2.8) *is satisfied for this r, then there exist versions of $\mathbf{Y}_\infty$ and $\overline{\mathbf{Y}}_\infty$ such that*

$$\|\mathbf{Y}_\infty\|_{B,r} \leq \overline{\mathbf{Y}}_\infty. \tag{2.31}$$

Further, if $(\mathbf{Y}_t^{(\varepsilon)})_{t\geq 0}$ is the process defined in Definition 2.9 for $\varepsilon > 0$, then versions of $\mathbf{Y}_\infty^{(\varepsilon)}$ can be chosen such that $\|\mathbf{Y}_\infty^{(\varepsilon)}\|_{B,r} \leq \overline{\mathbf{Y}}_\infty$ for all $\varepsilon > 0$ and $\mathbf{Y}_\infty^{(\varepsilon)} \xrightarrow{P} \mathbf{Y}_\infty$, as $\varepsilon \to 0$.

Proof of Lemma 2.12. We use the notations and setup of the proof of Theorems 2.4 and 2.6. Let $\varepsilon > 0$ and define a COGARCH(1, 1) state process $\overline{\mathbf{Y}}^{(\varepsilon)}$ similarly as above (with respect to $\mathbf{Y}^{(\varepsilon)}$). Let $\overline{J}_{0,t}^{(\varepsilon)}$ and $\overline{\mathbf{K}}_{0,t}^{(\varepsilon)}$ be defined similarly as $J_{0,t}^{(\varepsilon)}$ and $\mathbf{K}_{0,t}^{(\varepsilon)}$ (with respect to $\overline{\mathbf{Y}}^{(\varepsilon)}$). Then it is easy to see that $\overline{J}_{0,t}^{(\varepsilon)}$ and $\overline{\mathbf{K}}_{0,t}^{(\varepsilon)}$ are the right hand sides of (2.22) and (2.25), respectively. In particular, $\|J_{0,t}^{(\varepsilon)}\|_{B,r} \leq \overline{J}_{0,t}^{(\varepsilon)}$ and $\|\mathbf{K}_{0,t}^{(\varepsilon)}\|_{B,r} \leq \overline{\mathbf{K}}_{0,t}^{(\varepsilon)}$, and since $\overline{J}_{0,t}^{(\varepsilon)}$ and $\overline{\mathbf{K}}_{0,t}^{(\varepsilon)}$ converge in ucp as $\varepsilon \to 0$ to some $\overline{J}_{0,t}$ and $\overline{\mathbf{K}}_{0,t}$ such that

$$\overline{\mathbf{Y}}_t = \overline{J}_{0,t}\overline{\mathbf{Y}}_0 + \overline{\mathbf{K}}_{0,t},$$

it follows that $\|\mathbf{Y}_t\|_{B,r} \leq \overline{\mathbf{Y}}_t$ for fixed $t \geq 0$, giving (2.30).

Similar quantities such as $\overline{J}_{s,t}^{(\varepsilon)}$ and $\overline{J}_{s,t}$ can be defined when going from time s to time t, and similar results hold. Let $\overline{V}_t^{(\varepsilon)} := \alpha_0\|\mathbf{e}\|_{B,r} + \|\mathbf{e}\mathbf{a}'\|_{B,r}\,\overline{\mathbf{Y}}_{t-}^{(\varepsilon)}$ be the COGARCH(1, 1) volatility corresponding to $\overline{\mathbf{Y}}^{(\varepsilon)}$. Define

$$\begin{aligned} X_t &:= -\lambda t - \sum_{0<s\leq t} \log(1 + (\Delta L_s)^2\|\mathbf{e}\mathbf{a}'\|_{B,r}), \\ X_t^{(\varepsilon)} &:= -\lambda t - \sum_{0<s\leq t, (\Delta L_s)^2 \geq \varepsilon} \log(1 + (\Delta L_s)^2\|\mathbf{e}\mathbf{a}'\|_{B,r}). \end{aligned}$$

Then it follows from Theorem 2.2 and (2.3), that

$$\overline{V}_{t+}^{(\varepsilon)} = \left(\overline{V}_0 - \alpha_0 \|\mathbf{e}\|_{B,r} \lambda \int_0^t e^{X_s^{(\varepsilon)}}\, ds\right) e^{-X_t^{(\varepsilon)}}.$$

Thus we have $\overline{J}_{0,t}^{(\varepsilon)} = e^{-X_t^{(\varepsilon)}}$ and obtain another formula for $\overline{\mathbf{K}}_{0,t}^{(\varepsilon)}$, namely

$$\overline{\mathbf{K}}_{0,t}^{(\varepsilon)} = \|\mathbf{e}\mathbf{a}'\|_{B,r}^{-1} \left[\alpha_0 \|\mathbf{e}\|_{B,r}\, e^{-X_t^{(\varepsilon)}} - \alpha_0 \|\mathbf{e}\|_{B,r} \lambda \int_0^t e^{-(X_t^{(\varepsilon)} - X_s^{(\varepsilon)})}\, ds - \alpha_0 \|\mathbf{e}\|_{B,r}\right].$$

From this it can be seen that $\overline{J}_{0,t}^{(\varepsilon)}$ and $\overline{\mathbf{K}}_{0,t}^{(\varepsilon)}$ are bounded by $\overline{J}_{0,t} = e^{-X_t}$ and

$$\overline{\mathbf{K}}_{0,t} = \|\mathbf{e}\mathbf{a}'\|_{B,r}^{-1} \alpha_0 \|\mathbf{e}\|_{B,r} \left[e^{-X_t} - \lambda \int_0^t e^{-(X_t - X_s)}\, ds - 1\right],$$

respectively. Now define the versions

$$\begin{aligned}
\overline{\mathbf{Y}}_\infty &:= \sum_{i=0}^{\infty} \overline{J}_{0,1} \cdots \overline{J}_{i-1,i} \overline{\mathbf{K}}_{i,i+1}, \\
\mathbf{Y}_\infty^{(\varepsilon)} &:= \sum_{i=0}^{\infty} J_{0,1}^{(\varepsilon)} \cdots J_{i-1,i}^{(\varepsilon)} \mathbf{K}_{i,i+1}^{(\varepsilon)}, \\
\mathbf{Y}_\infty &:= \sum_{i=0}^{\infty} J_{0,1} \cdots J_{i-1,i} \mathbf{K}_{i,i+1}.
\end{aligned}$$

In the proof of Theorems 2.4 and 2.6 we have seen that (2.8) implies that the sum defining $\overline{\mathbf{Y}}_\infty$ converges almost surely. This then gives the claim, since

$$\|J_{i-1,i}\|_{B,r},\ \|J_{i-1,i}^{(\varepsilon)}\|_{B,r} \le \overline{J}_{i-1,i}, \quad \|\mathbf{K}_{i,i+1}\|_{B,r},\ \|\mathbf{K}_{i,i+1}^{(\varepsilon)}\|_{B,r} \le \overline{\mathbf{K}}_{i,i+1},$$

and $J_{i-1,i}^{(\varepsilon)}$ and $\mathbf{K}_{i,i+1}^{(\varepsilon)}$ converge in probability to $J_{i-1,i}$ and $\mathbf{K}_{i,i+1}$ as $\varepsilon \to 0$, respectively.

□

Proof of Proposition 2.11. All assertions apart from the implication "(2.28) $\Longrightarrow \lambda(\widetilde{B}) < 0$" follow immediately from Lemma 2.12 (observing that the existence of $E\|Y_t\|^k$ is independent of the specific matrix norm) and the corresponding properties of the COGARCH(1,1) process, see Section 4 in Klüppelberg et al. (2004). That (2.28) implies $\lambda(\widetilde{B}) < 0$ is a consequence of the Bauer-Fike perturbation result on eigenvalues, stating that for every eigenvalue $\widetilde{\lambda}_j$ of $\widetilde{B}$ we have

$$\min_{i=1,\ldots,q} |\lambda_i - \widetilde{\lambda}_j| \le \|S^{-1}(\widetilde{B} - B)S\|_r = \mu\, \|S^{-1}\mathbf{e}\mathbf{a}'S\|_r,$$

see e.g. Theorem 7.2.2 and its proof in Golub and van Loan (1989). □

Next, we determine the autocovariance function of the (not necessarily stationary) volatility process of Definition 2.1.

Theorem 2.13. *Let $(V_t)_{t\geq 0}$ be the volatility process specified in Definition 2.1, with state process $(\mathbf{Y}_t)_{t\geq 0}$ and parameters B, $\mathbf{a}$ and α_0. Suppose that $EL_1^4 < \infty$ and that $E\|\mathbf{Y}_t\|^2 < \infty \; \forall \, t \geq 0$ (as is the case for example if the conditions of Proposition 2.11 are satisfied). Then, with $\widetilde{B}$ defined as in (2.27),*

$$\operatorname{cov}(V_{t+h}, V_t) = \mathbf{a}' e^{\widetilde{B}h} \operatorname{cov}(\mathbf{Y}_t)\, \mathbf{a}, \quad t, h \geq 0. \tag{2.32}$$

Proof of Theorem 2.13. Since for fixed t, almost surely $V_t = V_{t+} = \alpha_0 + \mathbf{a}'\mathbf{Y}_t$, we obtain

$$\operatorname{cov}(V_{t+h}, V_t) = \mathbf{a}' \operatorname{cov}(\mathbf{Y}_{t+h}, \mathbf{Y}_t)\, \mathbf{a}. \tag{2.33}$$

For the ease of notation, we will assume that $t = 0$. Let $J_h := J_{0,h}$ and $\mathbf{K}_h := \mathbf{K}_{0,h}$ as constructed in the proof of Theorem 2.6. Then, using that $\|e^{Bt}\| \leq e^{\|B\|t}$ for any vector norm $\|\cdot\|$, it follows as in the proof of (2.23) that

$$E\|J_h\| \leq e^{\|B\|t}\, E\left\{ \exp\left(\sum_{0<s\leq h} \log(1 + (\Delta L_s)^2 \|\mathbf{e}\mathbf{a}'\|) \right) \right\} < \infty \tag{2.34}$$

by Klüppelberg et al. (2004), Lemma 4.1 (a). Using that $\mathbf{Y}_h = J_h\mathbf{Y}_0 + \mathbf{K}_h$, we conclude that $E\|\mathbf{K}_h\| < \infty$ and that

$$\begin{aligned} E(\mathbf{Y}_h\mathbf{Y}_0') &= E(E(\mathbf{Y}_h\mathbf{Y}_0' | J_h, \mathbf{K}_h)) \\ &= E\left(J_h\, E(\mathbf{Y}_0\mathbf{Y}_0') + \mathbf{K}_h\, E(\mathbf{Y}_0')\right) \\ &= E(J_h)\, E(\mathbf{Y}_0\mathbf{Y}_0') + E(\mathbf{K}_h)\, E(\mathbf{Y}_0'). \end{aligned}$$

On the other hand,

$$E(\mathbf{Y}_h)\, E(\mathbf{Y}_0') = E(J_h)\, E(\mathbf{Y}_0)\, E(\mathbf{Y}_0') + E(\mathbf{K}_h)\, E(\mathbf{Y}_0'),$$

so that $\operatorname{cov}(\mathbf{Y}_h, \mathbf{Y}_0) = E(J_h)\operatorname{cov}(\mathbf{Y}_0)$, and (2.32) will follow from (2.33) once we have shown that

$$E(J_t) = e^{\widetilde{B}t}, \quad t \geq 0. \tag{2.35}$$

To do that, it suffices to assume that $[L, L]_t$ is a compound Poisson process. The general case then follows from the fact that $J_t^{(\varepsilon)}$ as defined in the proof of Theorem 2.4 converges to J_t in L^1 as $\varepsilon \to 0$, since it converges stochastically and since there is an integrable majorant by (2.34) and its proof. So suppose that $[L, L]_t = \sum_{i=1}^{N(t)} Z_i$ is compound Poisson with intensity $c > 0$ and let $C_i = (I + Z_i \mathbf{e}\mathbf{a}')\, e^{B(\Gamma_i - \Gamma_{i-1})}$. Then, for $0 \leq s, t$, it follows from (2.10) and the independence of $J_{0,s}$ and $J_{s,s+t}$ that

$$E(J_{s+t}) = E(J_s)E(J_t).$$

It is easy to see that $E(J_t)$ is a continuous function in $t \in [0, \infty)$. Further, $E(J_0) = I$, and we conclude that $(E(J_t))_{t \geq 0}$ is a semigroup. We shall show that its generator A_J satisfies

$$A_J := \lim_{t \to 0} \frac{1}{t}(E(J_t) - I) = B + \int_{\mathbb{R}} y^2 \, d\nu_L(y)\, \mathbf{e}\mathbf{a}' = \widetilde{B}. \tag{2.36}$$

This then implies (2.35), since $E(J_t) = e^{tA_J}$, see e.g. Goldstein (1985), Proposition 2.5. To show (2.36), write

$$J_t = e^{Bt}\, \mathbf{1}_{\{N(t)=0\}} + e^{B(t-\Gamma_1)} C_1\, \mathbf{1}_{\{N(t)=1\}} + e^{B(t-\Gamma_{N(t)})} C_{N(t)} \cdots C_1\, \mathbf{1}_{\{N(t) \geq 2\}}. \tag{2.37}$$

We have $P(N(t) = k) = e^{-ct}(ct)^k/(k!)$ since $N(t)$ is Poisson distributed with parameter ct. Then by (2.34),

$$
\begin{aligned}
& E\left(e^{B(t-\Gamma_{N(t)})} C_{N(t)} \cdots C_1 \mathbf{1}_{\{N(t)\geq 2\}}\right) \\
\leq \;& e^{\|B\|t} E\left(\exp\left(\sum_{i=1}^{N(t)} \log(1+Z_i\|\mathbf{ea}'\|)\right) \mathbf{1}_{\{N(t)\geq 2\}}\right) \\
= \;& e^{\|B\|t} E\left(\exp\left(\sum_{i=1}^{N(t)} \log(1+Z_i\|\mathbf{ea}'\|)\right) \Big| N(t)\geq 2\right) P(N(t)\geq 2) \\
\leq \;& e^{\|B\|t} E\left(\exp\left(\sum_{i=1}^{N(t)+2} \log(1+Z_i\|\mathbf{ea}'\|)\right)\right) P(N(t)\geq 2) \\
= \;& e^{\|B\|t} E\left((1+Z_1\|\mathbf{ea}'\|)(1+Z_2\|\mathbf{ea}'\|)\right. \\
& \times E\left(\exp\left(\sum_{0<s\leq t} \log(1+(\Delta L_s)^2\|\mathbf{ea}'\|)\right)\right) P(N(t)\geq 2) \\
= \;& o(t) \quad \text{as} \quad t\to 0, \qquad (2.38)
\end{aligned}
$$

since $P(N(t)\geq 2) = o(t)$ as $t\to 0$. Further, since Γ_1 is uniformly distributed on $(0,t)$, conditional that $N(t)=1$, it follows that

$$
\begin{aligned}
& E\left(e^{B(t-\Gamma_1)} C_1 \mathbf{1}_{\{N(t)=1\}}\right) \\
= \;& E\left(e^{B(t-\Gamma_1)}(I+Z_1\mathbf{ea}')e^{B\Gamma_1} \Big| N(t)=1\right) P(N(t)=1) \\
= \;& \int_0^t e^{B(t-s)}\,(I+E(Z_1)\mathbf{ea}')e^{Bs}\,\frac{ds}{t}\,e^{-ct}\,ct.
\end{aligned}
$$

Since $\sup_{0\leq s\leq t}\|e^{Bs}-I\|$ converges to 0 as $t\to 0$, we conclude that

$$\lim_{t\to 0}\frac{1}{t}E\left(e^{B(t-\Gamma_1)}C_1\mathbf{1}_{\{N(t)=1\}}\right) = (I+E(Z_1)\mathbf{ea}')c.$$

Now (2.37) and (2.38) give (2.36), since

$$
\begin{aligned}
\lim_{t\to 0}\frac{E(J_t)-I}{t} &= \lim_{t\to 0}\frac{e^{Bt}e^{-ct}-I}{t} + c(I+E(Z_1)\mathbf{ea}') \\
&= -cI + B + c(I+E(Z_1)\mathbf{ea}') = \widetilde{B}.
\end{aligned}
$$

□

Since we are primarily interested in the stationary volatility process, we need to evaluate $\mathrm{cov}\,(\mathbf{Y}_\infty)$. But first we need an expression for $E(\mathbf{Y}_\infty)$.

Lemma 2.14. *Suppose that all the eigenvalues of B are distinct and that* (2.28) *holds. Then*

$$E(\mathbf{Y}_\infty) \quad = \quad -\alpha_0\mu\widetilde{B}^{-1}\mathbf{e} = \frac{\alpha_0\mu}{\beta_q - \alpha_1\mu}\mathbf{e}_1. \tag{2.39}$$

We need the following lemma:

Lemma 2.15. *Let T be exponentially distributed with parameter c, and suppose that $\lambda(B) < 0$. Let*

$$M := E(e^{BT} \otimes e^{BT}).$$

Then

$$E(e^{BT}) \quad = \quad \left(I - c^{-1}B\right)^{-1}, \tag{2.40}$$

$$M^{-1} \quad = \quad I_{q^2} - \left(I \otimes (c^{-1}B)\right) - \left((c^{-1}B) \otimes I\right). \tag{2.41}$$

Further, $(I \otimes B) + (B \otimes I)$ is invertible, and for any real $(q \times q)$-matrix U the unique solution of $((I \otimes B) + (B \otimes I))\,\mathbf{x} = \mathrm{vec}\,(U)$ is given by

$$\mathbf{x} = \mathrm{vec}\left(-\int_0^\infty e^{Bt}\,U\,e^{B't}\,dt\right). \tag{2.42}$$

Here, we denote by I the $(q \times q)$-identity matrix, and by I_{q^2} the $(q^2 \times q^2)$-identity matrix.

Proof. Equations (2.40) and (2.41) follow by simple calculations and a diagonalisation argument, while invertibility of $(I \otimes B) + (B \otimes I)$ and (2.42) are consequences of Lyapunov's theorem for the solution of Lyapunov equations, see e.g. Section 9.3 in Godunov (1998). □

Proof of Lemma 2.14. Suppose first that the Lévy measure of L is finite and let Q and $\mathbf{R}$ be as in Theorem 2.8 (*b*) (writing (T, Z) instead of (T_0, Z_0)). Then by Lemma 2.15,

$$E(Q) \quad = \quad \left(I - c^{-1}B\right)^{-1}(I + E(Z)\mathbf{e}\mathbf{a}'),$$

$$E(\mathbf{R}) \quad = \quad \alpha_0 E(Z)\left(I - c^{-1}B\right)^{-1}\mathbf{e},$$

so that (2.17) gives

$$(I - E(Q))E(\mathbf{Y}_\infty) = E(\mathbf{R}).$$

Further,

$$\left(I - c^{-1}B\right)\left(I - E(Q)\right) = \left[\left(I - c^{-1}B\right) - I - E(Z)\mathbf{e}\mathbf{a}'\right] = -\frac{1}{c}(B + \mu\mathbf{e}\mathbf{a}'),$$

giving

$$E(\mathbf{Y}_\infty) = -c(B + \mu\mathbf{e}\mathbf{a}')^{-1}\left(I - c^{-1}B\right)E(\mathbf{R}) = -\alpha_0\mu(B + \mu\mathbf{e}\mathbf{a}')^{-1}\mathbf{e}.$$

Denoting $\mathbf{u} = (u_1, \ldots, u_q)' := (B + \mu\mathbf{e}\mathbf{a}')^{-1}\mathbf{e}$, it is easy to see that $u_2 = \ldots = u_q = 0$ and $u_1 = 1/(\alpha_1\mu - \beta_q)$. In the case when ν_L is infinite the result follows from Lemma 2.12, using that $\overline{Y}_\infty$ is an integrable majorant by (2.28). $\square$

The following theorem contains the main results of this section. It demonstrates that the autocorrelation function of the stationary COGARCH volatility process is the same as that of a continuous-time ARMA process. This reflects the corresponding discrete-time result that the autocorrelation function of a GARCH volatility process is the same as that of a discrete-time ARMA process.

Theorem 2.16. *Suppose that the eigenvalues of the matrix B are distinct, $\lambda(B) < 0$ and (2.29) holds. Then the matrix $(I \otimes \widetilde{B}) + (\widetilde{B} \otimes I) + \rho((\mathbf{e}\mathbf{a}') \otimes (\mathbf{e}\mathbf{a}'))$ is invertible, and the covariance matrix of $\mathbf{Y}_\infty$ is the unique solution of*

$$\begin{aligned}\left[(I \otimes \widetilde{B}) + (\widetilde{B} \otimes I) + \rho((\mathbf{e}\mathbf{a}') \otimes (\mathbf{e}\mathbf{a}'))\right] &\operatorname{vec}(\operatorname{cov}(\mathbf{Y}_\infty)) \\ &= \frac{-\alpha_0^2\beta_q^2\rho}{(\beta_q - \mu\alpha_1)^2}\operatorname{vec}(\mathbf{e}\mathbf{e}').\end{aligned} \tag{2.43}$$

Let $(\psi_t)_{t\geq 0}$ be a stationary CARMA$(q, p-1)$ process (as defined in Section 2) with location parameter 0, moving average coefficients $\alpha_1, \ldots, \alpha_p$, autoregressive coefficients $\beta_1 - \mu\alpha_q, \beta_2 - \mu\alpha_{q-1}, \ldots, \beta_q - \alpha_1\mu$, driving Lévy process $\widetilde{L}$ and corresponding state process $(\zeta_t)_{t\geq 0}$. Suppose that $E(\widetilde{L}_1)^2 < \infty$, $E(\widetilde{L}_1) = \mu$ and $\operatorname{var}(\widetilde{L}_1) = \rho$ and define

$$m := \rho\int_0^\infty \mathbf{a}'e^{\widetilde{B}t}\mathbf{e}\mathbf{e}'e^{\widetilde{B}'t}\mathbf{a}\,dt = \operatorname{var}(\psi_t).$$

Then $0 \leq m < 1$, and

$$
\begin{aligned}
\operatorname{cov}(\mathbf{Y}_\infty) &= \frac{\alpha_0^2\beta_q^2}{(\beta_q-\mu\alpha_1)^2(1-m)} \operatorname{cov}(\zeta_\infty) \\
&= \frac{\alpha_0^2\beta_q^2\rho}{(\beta_q-\mu\alpha_1)^2(1-m)} \int_0^\infty e^{\widetilde{B}t}\mathbf{e}\mathbf{e}'e^{\widetilde{B}'t}\,dt, && (2.44)\\
\operatorname{var}(V_\infty) &= \frac{\alpha_0^2\beta_q^2}{(\beta_q-\mu\alpha_1)^2}\,\frac{m}{1-m}, && (2.45)\\
E(V_\infty) &= \frac{\alpha_0\beta_q}{\beta_q-\mu\alpha_1}, && (2.46)\\
E(\psi_\infty) &= \frac{\alpha_1\mu}{\beta_q-\mu\alpha_1}. && (2.47)
\end{aligned}
$$

If $(V_t)_{t\geq 0}$ is the stationary COGARCH volatility process, then

$$
\operatorname{cov}(V_{t+h},V_t) = \frac{\alpha_0^2\beta_q^2}{(\beta_q-\mu\alpha_1)^2(1-m)}\operatorname{cov}(\psi_{t+h},\psi_t), \quad t,h\geq 0, \tag{2.48}
$$

showing, in particular, that V has the same autocorrelation function as ψ. If the eigenvalues $\widetilde{\lambda}_1,\ldots,\widetilde{\lambda}_q$ of $\widetilde{B}$ are also distinct, and $a(z)$ and $\widetilde{b}(z)$ are the characteristic polynomials associated with $\mathbf{a}$ and $\widetilde{B}$, then

$$
\operatorname{cov}(V_{t+h},V_t) = \frac{\alpha_0^2\beta_q^2\rho}{(\beta_q-\mu\alpha_1)^2(1-m)}\sum_{j=1}^q \frac{a(\widetilde{\lambda}_j)\,a(-\widetilde{\lambda}_j)}{\widetilde{b}'(\widetilde{\lambda}_j)\,\widetilde{b}(-\widetilde{\lambda}_j)}\,e^{\widetilde{\lambda}_j h}, \quad t,h\geq 0, \tag{2.49}
$$

where $\widetilde{b}'$ denotes the derivative of $\widetilde{b}$.

Proof of Theorem 2.16. By Lemma 2.12 and the dominated convergence theorem, it is sufficient to assume that $[L,L]$ is a compound Poisson process. Hence, let Q and $\mathbf{R}$ be as in Theorem 2.8, writing (T,Z) instead of (T_0,Z_0), where T is exponentially distributed with parameter $c>0$. Then

$$
E(\mathbf{Y}_\infty\mathbf{Y}'_\infty) - E(Q\mathbf{Y}_\infty\mathbf{Y}'_\infty Q') = E(Q\mathbf{Y}_\infty\mathbf{R}') + E(\mathbf{R}\mathbf{Y}'_\infty Q') + E(\mathbf{R}\mathbf{R}') \tag{2.50}
$$

by (2.17), and all these expectations exist by (2.29). Now

$$
\begin{aligned}
E(Q\mathbf{Y}_\infty\mathbf{Y}'_\infty Q') &= E\left(E[Q\mathbf{Y}_\infty\mathbf{Y}'_\infty Q'|Q]\right)\\
&= E\left(E[Q\,E(\mathbf{Y}_\infty\mathbf{Y}'_\infty)\,Q'|T]\right)\\
&= E\left(e^{BT}\,E\left[(I+Z\mathbf{e}\mathbf{a}')E(\mathbf{Y}_\infty\mathbf{Y}'_\infty)\,(I+Z\mathbf{a}\mathbf{e}')\right]e^{B'T}\right).
\end{aligned}
$$

Using that $\mathrm{vec}\,(A_1A_2A_3) = (A_3' \otimes A_1)\mathrm{vec}\,(A_2)$ for matrices A_1, A_2 and A_3 it follows with M as in Lemma 2.15 that

$$\begin{aligned}
&\mathrm{vec}\,(E(Q\mathbf{Y}_\infty\mathbf{Y}_\infty' Q')) \\
&\quad = M\,\mathrm{vec}\,(E\,((I + Z\mathbf{ea}')E(\mathbf{Y}_\infty\mathbf{Y}_\infty')(I + Z\mathbf{ae}'))) \\
&\quad = M\,(E\,((I + Z\mathbf{ea}') \otimes (I + Z\mathbf{ea}')))\,\mathrm{vec}\,(E(\mathbf{Y}_\infty\mathbf{Y}_\infty')) \\
&\quad = M\big(I_{q^2} + E(Z)((\mathbf{ea}') \otimes I) + E(Z)(I \otimes (\mathbf{ea}')) + E(Z^2)((\mathbf{ea}') \otimes (\mathbf{ea}'))\big) \\
&\qquad \times \mathrm{vec}\,(E(\mathbf{Y}_\infty\mathbf{Y}_\infty')).
\end{aligned}$$

Similar expressions can be obtained for $\mathrm{vec}\,(E(Q\mathbf{Y}_\infty\mathbf{R}'))$, $\mathrm{vec}\,(E(\mathbf{R}\mathbf{Y}_\infty' Q'))$ and $\mathrm{vec}\,(E(\mathbf{R}\mathbf{R}'))$ and we obtain from (2.50) that

$$\begin{aligned}
&\big[I_{q^2} - M\big(I_{q^2} + E(Z)((\mathbf{ea}') \otimes I) + E(Z)(I \otimes (\mathbf{ea}')) + E(Z^2)((\mathbf{ea}') \otimes (\mathbf{ea}'))\big)\big] \\
&\qquad\qquad \times \mathrm{vec}\,(E(\mathbf{Y}_\infty\mathbf{Y}_\infty')) \\
&= M\,\mathrm{vec}\,\big[\alpha_0^2E(Z^2)\mathbf{ee}' + \alpha_0(E(Z)I + E(Z^2)\mathbf{ea}')E(\mathbf{Y}_\infty)\mathbf{e}' + \alpha_0\mathbf{e}E(\mathbf{Y}_\infty') \\
&\qquad\qquad \times (E(Z)I + E(Z^2)\mathbf{ae}')\big]
\end{aligned}$$

Multiplying this equation by cM^{-1}, using (2.41), (2.39) as well as $\mu = c\,E(Z)$ and $\rho = c\,E(Z^2)$, we obtain

$$\begin{aligned}
&-\Big[(I \otimes (B + \mu\mathbf{ea}')) + ((B + \mu\mathbf{ea}') \otimes I) + \rho((\mathbf{ea}') \otimes (\mathbf{ea}'))\Big]\,\mathrm{vec}\,(E(\mathbf{Y}_\infty\mathbf{Y}_\infty')) \\
&= \mathrm{vec}\,\Big[\alpha_0^2\rho\mathbf{ee}' - \alpha_0^2(\mu I + \rho\mathbf{ea}')\mu(B + \mu\mathbf{ea}')^{-1}\mathbf{ee}' - \alpha_0^2\mathbf{ee}'(B' + \mu\mathbf{ae}')^{-1}\mu \\
&\qquad\qquad \times (\mu I + \rho\mathbf{ae}')\Big].
\end{aligned}$$

Adding to this

$$\begin{aligned}
&\Big[(I \otimes \widetilde{B}) + (\widetilde{B} \otimes I) + \rho((\mathbf{ea}') \otimes (\mathbf{ea}'))\Big]\,\mathrm{vec}\,(E(\mathbf{Y}_\infty)E(\mathbf{Y}_\infty')) \\
&\quad = \mathrm{vec}\,\Big[\widetilde{B}\,E(\mathbf{Y}_\infty)E(\mathbf{Y}_\infty') + E(\mathbf{Y}_\infty)E(\mathbf{Y}_\infty')\,\widetilde{B}' + \rho\mathbf{ea}'E(\mathbf{Y}_\infty)E(\mathbf{Y}_\infty')\mathbf{ae}'\Big] \\
&\quad = \alpha_0^2\,\mathrm{vec}\,\Big[\mu^2\mathbf{ee}'(\widetilde{B}')^{-1} + \mu^2\widetilde{B}^{-1}\mathbf{ee}' + \rho\mu^2\mathbf{ea}'\widetilde{B}^{-1}\mathbf{ee}'(\widetilde{B}')^{-1}\mathbf{ae}'\Big]
\end{aligned}$$

on both sides results in

$$
\begin{aligned}
&-\left[(I \otimes \widetilde{B}) + (\widetilde{B} \otimes I) + \rho((\mathbf{e}\mathbf{a}') \otimes (\mathbf{e}\mathbf{a}'))\right] \operatorname{vec}(\operatorname{cov}(\mathbf{Y}_0)) \\
&= \alpha_0^2 \rho \left[1 - \mu \left(\mathbf{a}' \widetilde{B}^{-1} \mathbf{e}\right)\right]^2 \operatorname{vec}(\mathbf{e}\mathbf{e}') = \frac{\alpha_0^2 \beta_q^2 \rho}{(\beta_q - \mu\alpha_1)^2} \operatorname{vec}(\mathbf{e}\mathbf{e}'),
\end{aligned}
$$

which is (2.43), where we used (2.39) in the last equation.

Now let $A := (I \otimes \widetilde{B}) + (\widetilde{B} \otimes I)$ and $\mathbf{x} := \operatorname{vec}(\operatorname{cov}(\mathbf{Y}_\infty))$. By Proposition 2.11 and Lemma 2.15, A is invertible. Observe that the matrix $\rho((\mathbf{e}\mathbf{a}') \otimes (\mathbf{e}\mathbf{a}'))$ has non-zero entries only in the last row. Denote this row by $\mathbf{c}'$. Further, set $\gamma := \rho\alpha_0^2\beta_q^2(\mu\alpha_1 - \beta_q)^{-2}$. Then (2.43) can be written as

$$
A\mathbf{x} + (\mathbf{c}'\mathbf{x})\mathbf{e}_{q^2} = -\gamma\, \mathbf{e}_{q^2}.
$$

We know already that a solution to this equation exists. Suppose there are two of them and, call them $\mathbf{x}_1$ and $\mathbf{x}_2$. Then $A\mathbf{x}_1 = -(\gamma + \mathbf{c}'\mathbf{x}_1)\mathbf{e}_{q^2}$ and $A\mathbf{x}_2 = -(\gamma + \mathbf{c}'\mathbf{x}_2)\mathbf{e}_{q^2}$. Denoting the unique solution of $A\mathbf{y} = -n\,\mathbf{e}_{q^2}$ by $\mathbf{y}(n)$, $n \in \mathbb{R}$, it follows that $\mathbf{x}_1 = \mathbf{y}(\gamma + \mathbf{c}'\mathbf{x}_1)$ and $\mathbf{x}_2 = \mathbf{y}(\gamma + \mathbf{c}'\mathbf{x}_2)$. Since $\mathbf{x}_1 \neq \mathbf{0} \neq \mathbf{x}_2$, this implies that $\gamma + \mathbf{c}'\mathbf{x}_1 \neq \mathbf{0} \neq \gamma + \mathbf{c}'\mathbf{x}_2$, and using the linearity of the solution $\mathbf{y}(n)$ in n it follows that there is $\kappa \neq 0$ such that $\mathbf{x}_2 = \kappa\mathbf{x}_1$. Thus we have $A\mathbf{x}_1 = -(\gamma + \mathbf{c}'\mathbf{x}_1)\mathbf{e}_{q^2}$ and $\kappa A\mathbf{x}_1 = -(\gamma + \kappa\mathbf{c}'\mathbf{x}_1)\mathbf{e}_{q^2}$, and this is only possible if $\kappa = 1$, so $\mathbf{x}_1 = \mathbf{x}_2$. So the solution of (2.43) is unique, implying that the matrix $A + \rho((\mathbf{e}\mathbf{a}') \otimes (\mathbf{e}\mathbf{a}'))$ is invertible.

By (2.42), the solution $y(n)$ of $Ay = -ne_{q^2}$ is given by

$$
y(n) = \operatorname{vec}\left(n \int_0^\infty e^{\widetilde{B}t}\mathbf{e}\mathbf{e}'e^{\widetilde{B}'t}\, dt\right). \tag{2.51}
$$

This gives

$$
\operatorname{cov}(\mathbf{Y}_\infty) = (\gamma + \mathbf{c}' \operatorname{vec}(\operatorname{cov}(\mathbf{Y}_\infty))) \int_0^\infty e^{\widetilde{B}t}\mathbf{e}\mathbf{e}'e^{\widetilde{B}'t}\, dt.
$$

Since both $\operatorname{cov}(\mathbf{Y}_\infty)$ and $\int_0^\infty e^{\widetilde{B}t}\mathbf{e}\mathbf{e}'e^{\widetilde{B}'t}\, dt$ are positive semidefinite, it follows that $\gamma + \mathbf{c}' \operatorname{vec}(\operatorname{cov}(\mathbf{Y}_\infty)) > 0$. By Brockwell (2001), the stationary CARMA state vector ζ_∞ has covariance matrix

$$
\operatorname{cov}(\zeta_\infty) = \rho \int_0^\infty e^{\widetilde{B}t}\mathbf{e}\mathbf{e}'e^{\widetilde{B}'t}\, dt,
$$

so that there is $u > 0$ such that

$$\operatorname{cov}(\mathbf{Y}_\infty) = u \operatorname{cov}(\zeta_\infty). \tag{2.52}$$

Inserting (2.52) in (2.43) and using (2.51) shows

$$-u\rho \operatorname{vec}(\mathbf{e}\mathbf{e}') + u\rho^2 \operatorname{vec}\left(\mathbf{e}\mathbf{a}' \int_0^\infty e^{\widetilde{B}t}\mathbf{e}\mathbf{e}' e^{\widetilde{B}'t}\, dt\, \mathbf{a}\mathbf{e}'\right) = \frac{-\alpha_0^2\beta_q^2\rho}{(\beta_q - \mu\alpha_1)^2} \operatorname{vec}(\mathbf{e}\mathbf{e}'),$$

so that

$$-u(1-m)\operatorname{vec}(\mathbf{e}\mathbf{e}') = \frac{-\alpha_0^2\beta_q^2}{(\beta_q - \mu\alpha_1)^2}\operatorname{vec}(\mathbf{e}\mathbf{e}').$$

Since $u > 0$ and $\alpha_0, \beta_q \neq 0$, it follows that $0 \leq m < 1$ and that

$$u = \frac{\alpha_0^2\beta_q^2}{(\beta_q - \mu\alpha_1)^2(1-m)},$$

giving (2.44). This implies (2.45) using $V_\infty = \alpha_0 + \mathbf{a}'\mathbf{Y}_\infty$, and (2.46) follows from (2.39). Finally,

$$E(\psi_\infty) = \mathbf{a}'\, E\int_0^\infty e^{\widetilde{B}t}\mathbf{e}\, d\widetilde{L}_t = \mu \int_0^\infty \mathbf{a}' e^{\widetilde{B}t}\mathbf{e}\, dt = -\mu\mathbf{a}'\widetilde{B}^{-1}\mathbf{e},$$

giving (2.47), and (2.48) and (2.49) are direct consequences of (2.32), (2.44) and the autocovariance function of a CARMA process (see Brockwell (2001)). □

2.5 Positivity conditions for the volatility

In order for the definition of the COGARCH price process $dG_t = \sqrt{V_t}\, dt$ to make sense it is necessary that V_t be non-negative for all $t \geq 0$. The following Theorem gives necessary and sufficient conditions for this to occur with probability 1.

Theorem 2.17. *(a) Let $(\mathbf{Y}_t)_{t\geq 0}$ be the state vector of a COGARCH(p,q) volatility process $(V_t)_{t\geq 0}$ with parameters B, $\mathbf{a}$ and $\alpha_0 > 0$. Let $\gamma \geq -\alpha_0$ be a real constant. Suppose that the following two conditions hold:*

$$\mathbf{a}'e^{Bt}\mathbf{e} \;\geq\; 0 \quad \forall\, t \geq 0, \tag{2.53}$$

$$\mathbf{a}'e^{Bt}\mathbf{Y}_0 \;\geq\; \gamma \quad a.s. \quad \forall\, t \geq 0. \tag{2.54}$$

Then for any driving Lévy process, with probability one,

$$V_t \geq \alpha_0 + \gamma \geq 0 \quad \forall\, t \geq 0. \tag{2.55}$$

Conversely, if either (2.54) *fails, or* (2.54) *holds with* $\gamma > -\alpha_0$ *and* (2.53) *fails, then there exists a driving compound Poisson process* L *and* $t_0 \geq 0$ *such that* $P(V_{t_0} < 0) > 0$.

(*b*) *Suppose that all the eigenvalues of* B *are distinct and that* (2.8) *and* (2.53) *both hold. Then with probability one the stationary COGARCH(p,q) volatility process* $(V_t)_{t\geq 0}$ *satisfies*

$$V_t \geq \alpha_0 > 0 \quad \forall\, t \geq 0.$$

Proof of Theorem 2.17. (*a*) Suppose that (2.53) and (2.54) both hold. By Lemma 2.10, it suffices to show (2.55) for the case that $[L, L] = \sum_{i=1}^{N(t)} Z_i$ is a compound Poisson process, with jump times $(\Gamma_n)_{n\in\mathbb{N}}$. Then it follows easily by induction from (2.6) and (2.18) that

$$\mathbf{Y}_t = e^{Bt}\mathbf{Y}_0 + \sum_{i=1}^{N(t)} e^{B(t-\Gamma_i)}\mathbf{e}V_{\Gamma_i}Z_i, \quad t \geq 0. \tag{2.56}$$

In view of the proof of (*b*) below, let $s \geq 0$. Then

$$\begin{aligned} \mathbf{a}'e^{Bs}\mathbf{Y}_t &= \mathbf{a}'e^{B(s+t)}\mathbf{Y}_0 + \sum_{i=1}^{N(t)} \mathbf{a}'e^{B(s+t-\Gamma_i)}\,\mathbf{e}\,V_{\Gamma_i}Z_i && (2.57)\\ &\geq \gamma + \sum_{i=1}^{N(t)} \mathbf{a}'e^{B(s+t-\Gamma_i)}\,\mathbf{e}\,V_{\Gamma_i}Z_i. && (2.58) \end{aligned}$$

Setting $s = 0$, it follows that $V_t = \alpha_0 + \mathbf{a}'\mathbf{Y}_{t-} \geq \alpha_0 + \gamma$ for $t \in [0, \Gamma_1]$, hence also $V_{\Gamma_1+} \geq \alpha_0 + \gamma \geq 0$ by (2.53) and (2.58), and an induction argument shows that $V_t \geq \alpha_0 + \gamma$ for all $t \geq 0$, i.e. (2.55) holds.

For the converse, suppose first that (2.54) fails. Then, using the continuity of the function $t \mapsto e^{Bt}$, it follows that there is $(t_1, t_2) \subset (0, \infty)$ such that $P(\alpha_0 + \mathbf{a}'e^{Bt}\mathbf{Y}_0 < 0\ \forall\, t \in (t_1, t_2)) > 0$, and since $P(\Gamma_1 > t_2) > 0$ we get the claim from (2.57). So suppose that (2.54) holds with $\gamma > -\alpha_0$, but (2.53) fails. Suppose that the support of the Lévy measure of the compound Poisson process $[L, L]$ (and hence the support of the jump distribution Z_1) is unbounded. Let $(t_3, t_4) \subset (0, \infty)$ be an interval such that $\mathbf{a}'e^{Bt}\mathbf{e} \leq -c_1 < 0$ for all $t \in (t_3, t_4)$ for some $c_1 < 0$. Let $t_5 > t_4$. By (2.54) we have $P(V_{\Gamma_1} \geq \alpha_0 + \gamma) = 1$, so that it is easy to see that the set

$$A := \{\Gamma_1 < t_5 < \Gamma_2,\ t_5 - \Gamma_1 \in (t_3, t_4),\ V_{\Gamma_1} \geq \alpha_0 + \gamma\}$$

has positive probability. On A, we have by (2.57)

$$V_{t_5} = \alpha_0 + \mathbf{a}'e^{Bt_5}\mathbf{Y}_0 + \mathbf{a}'e^{B(t_5-\Gamma_1)}\,\mathbf{e}\,V_{\Gamma_1}Z_1.$$

Now $\mathbf{a}'e^{B(t_5-\Gamma_1)}\mathbf{e} \leq -c_1$, and by choosing Z_1 (which is independent of Γ_1, Γ_2 and $\mathbf{Y}_0$) large enough we obtain $P(V_{t_5} < 0) > 0$.

(b) In view of (a) it remains to show that $\mathbf{Y}_\infty$ satisfies (2.54). For the proof of this, it suffices by Lemma 2.12 to assume that $[L, L]$ is compound Poisson. Let $(\widetilde{\mathbf{Y}}_t)_{t\geq 0}$ be a state process with $\widetilde{\mathbf{Y}}_0 = 0$. Then (2.54) holds for $\widetilde{\mathbf{Y}}_0$ with $\gamma = 0$, and it follows from (2.58), (2.53) and (2.55) that $\mathbf{a}'e^{Bs}\widetilde{\mathbf{Y}}_t \geq 0$ for all $s, t \geq 0$. Since $\widetilde{\mathbf{Y}}_t$ converges in distribution to $\mathbf{Y}_\infty$ as $t \to \infty$, (2.54) follows with $\gamma = 0$. $\square$

For the stationary COGARCH volatility process or for the process with $\mathbf{Y}_0 = \mathbf{0}$, the condition (2.53) alone is sufficient for almost sure non-negativity. The expression $\mathbf{a}'e^{Bt}\mathbf{e}$ is in fact the kernel of a CARMA process with autoregressive coefficients $b_1, \ldots, b_q$ and moving average coefficients $a_1, \ldots, a_q$. Results pertaining to non-negativity of a CARMA kernel have been recently obtained by Tsai and Chan (2004). We state their results in the next theorem in the context of COGARCH rather than CARMA processes. Statement (e) below has also been obtained by Todorov and Tauchen (2004). Recall that a function ϕ on $(0, \infty)$ is called *completely monotone* if it possesses derivatives of all orders and satisfies $(-1)^n\frac{d^n\phi}{dt^n}(t) \geq 0$ for all $t > 0$ and all $n \in \mathbb{N}_0$.

Theorem 2.18. *Let B and* $\mathbf{a}$ *be the parameters of a COGARCH(p, q) process. If $\lambda(B) < 0$ and $\alpha_1 > 0$ we have the following results.*

(a) For the COGARCH(p,q) process, equation (2.53) *holds if and only the ratio of the characteristic polynomials* $a(\cdot)/b(\cdot)$ *is completely monotone on* $(0,\infty)$.

(b) A sufficient condition for (2.53) *to hold for the COGARCH($1,q$) process is that either*

(i) all eigenvalues of B are real and negative, or

(ii) if $(\lambda_{i_1}, \lambda_{i_1+1}), \ldots, (\lambda_{i_r}, \lambda_{i_r+1})$ *is a partition of the set of all pairs of complex conjugate eigenvalues of B (counted with multiplicity), then there exists an injective mapping* $u : \{1, \ldots, r\} \to \{1, \ldots, n\}$ *such that* $\lambda_{u(j)}$ *is a real eigenvalue of B satisfying* $\lambda_{u(j)} \geq \Re(\lambda_{i_j})$.

(c) A necessary condition for (2.53) *to hold for the COGARCH($1,q$) process is that there exists a real eigenvalue of B not smaller than the real part of all other eigenvalues of B.*

(d) Suppose $2 \leq p \leq q$, *that all eigenvalues of B are negative and ordered as in Definition 2.3, and that the roots* γ_j *of* $a(z) = 0$ *are negative and ordered such that* $\gamma_{p-1} \leq \ldots \leq \gamma_1 < 0$. *Then a sufficient condition for* (2.53) *to hold for the COGARCH(p,q) process is that*

$$\sum_{i=1}^{k} \gamma_i \leq \sum_{i=1}^{k} \lambda_i \quad \forall\, k \in \{1, \ldots, p-1\}.$$

(e) A necessary and sufficient condition for (2.53) *in the COGARCH($2,2$) case is that both eigenvalues of B are real, that* $\alpha_2 \geq 0$ *and that* $\alpha_1 \geq -\alpha_2 \lambda(B)$.

Although characterisation (a) may be difficult to check in general, it gives a method of constructing further pairs ($\mathbf{a}$, B) for which (2.53) holds, since the product of two completely monotone functions is again completely monotone.

2.6 The autocorrelation of the squared increments

In Section 2.4 we investigated the behaviour of the autocorrelation function of the volatility process. Since one of the striking features of observed financial time series is that the returns have negligible correlation while the squared returns are significantly correlated, we now turn to the second-order properties of the increments of the COGARCH process itself. We therefore assume that V is strictly stationary and non-negative and define, for $r > 0$,

$$G_t^{(r)} := G_{t+r} - G_t = \int_{(t,t+r]} \sqrt{V_s}\, dL_s, \quad t \geq 0.$$

It is easy to see that $(G_t^{(r)})_{t\geq 0}$ is a stationary process. Let μ and $\widetilde{B}$ be defined as in Section 2.4. We then have the following theorem.

Theorem 2.19. *Let B, $\mathbf{a}$ and α_0 be the parameters of a COGARCH(p,q) process whose driving Lévy process has no Gaussian part and for which $EL_1 = 0$. Suppose that the eigenvalues of B are distinct, that (2.28) and (2.53) hold and that V is the stationary volatility process. Then for any $t \geq 0$ and $h \geq r > 0$,*

$$E(G_t^{(r)}) = 0, \tag{2.59}$$

$$E((G_t^{(r)})^2) = \frac{\alpha_0 \beta_q r}{\beta_q - \mu \alpha_1} E(L_1^2), \tag{2.60}$$

$$\operatorname{cov}(G_t^{(r)}, G_{t+h}^{(r)}) = 0. \tag{2.61}$$

If in addition (2.29) holds then

$$\operatorname{cov}((G_t^{(r)})^2, (G_{t+h}^{(r)})^2) = E(L_1^2)\, \mathbf{a}' e^{\widetilde{B}h} \widetilde{B}^{-1}(I - e^{-\widetilde{B}r}) \operatorname{cov}(\mathbf{Y}_r, G_r^2), \quad h \geq r, \tag{2.62}$$

$$\operatorname{var}((G_t^{(r)})^2) = 6E(L_1^2)\, \mathbf{a}'\mathbf{K}_r + 2(rE(L_1^2)E(V_\infty))^2 + rE(L_1^4)E(V_\infty^2), \tag{2.63}$$

where

$$\operatorname{cov}(\mathbf{Y}_r, G_r^2) = [(I - e^{\widetilde{B}r})\operatorname{cov}(\mathbf{Y}_\infty) - \widetilde{B}^{-1}(e^{\widetilde{B}r} - I)\operatorname{cov}(\mathbf{Y}_\infty)B']\mathbf{e}$$

and

$$\mathbf{K}_r := \left[(rI - \widetilde{B}^{-1}(e^{\widetilde{B}r} - I))\operatorname{cov}(\mathbf{Y}_\infty) - \widetilde{B}^{-1}(\widetilde{B}^{-1}(e^{\widetilde{B}r} - I) - rI)\operatorname{cov}(\mathbf{Y}_\infty)B'\right]\mathbf{e}.$$

The autocovariance function (2.62), like that of the CARMA process with parameters $\widetilde{B}$ and $\mathbf{a}$, is a linear combination of terms of the form $e^{\tilde{\lambda}_j h}$, $j = 1, \ldots, q$, where $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_q$ are the eigenvalues of $\widetilde{B}$.

Proof of Theorem 2.19. We mimic the proof of Proposition 5.1 of Klüppelberg et al. (2004), i.e. in the COGARCH(1, 1) case. Observe that (2.59) and (2.61) follow immediately, since $(L_t)_{t\geq 0}$ is a zero-mean martingale. Further, $(G_t)_{t\geq 0}$ is a square integrable martingale, and using the compensation formula (e.g. Bertoin (1996), page 7), we have

$$EG_r^2 = E\int_0^r V_s\, d[L,L]_s = E\sum_{0<s\leq r} V_s(\Delta L_s)^2 = E(L_1)^2 rE(V_\infty),$$

and (2.60) follows from (2.46). Before showing (2.62), we verify that $EG_t^4 < \infty$ if (2.29) is satisfied: it follows from the Burkholder-Davis-Gundy inequality (see e.g. Protter (2004), page 222) that $EG_t^4 < \infty$ if $E[G,G]_t^2 < \infty$. Let $\overline{V}_t = \alpha_0\|\mathbf{e}\|_{B,r} + \|\mathbf{e}\mathbf{a}'\|_{B,r}\overline{\mathbf{Y}}_{t-}$ the volatility of the COGARCH(1, 1) process constructed in Lemma 2.12, and let $\overline{G}_t = \int_0^t \sqrt{\overline{V}_t}\, dL_t$ the corresponding GOGARCH(1, 1) price process. Then it follows from Lemma 2.12 that there is $C_1 > 0$ such that

$$0 \leq V_s = \alpha_0 + \mathbf{a}'\mathbf{Y}_{s-} \leq \alpha_0 + C_1\overline{\mathbf{Y}}_{s-} = \alpha_0 + C_1\frac{\overline{V}_s - \alpha_0\|\mathbf{e}\|_{B,r}}{\|\mathbf{e}\mathbf{a}'\|_{B,r}}.$$

Then

$$\begin{aligned}
[G,G]_t &= \int_0^t V_s\, d[L,L]_s \\
&\leq \frac{C_1}{\|\mathbf{e}\mathbf{a}'\|_{B,r}}\int_0^t \overline{V}_s\, d[L,L]_s + \left(\alpha_0 - \frac{C_1\alpha_0\|\mathbf{e}\|_{B,r}}{\|\mathbf{e}\mathbf{a}'\|_{B,r}}\right)[L,L]_t \\
&= \frac{C_1}{\|\mathbf{e}\mathbf{a}'\|_{B,r}}[\overline{G},\overline{G}]_t + \left(\alpha_0 - \frac{C_1\alpha_0\|\mathbf{e}\|_{B,r}}{\|\mathbf{e}\mathbf{a}'\|_{B,r}}\right)[L,L]_t,
\end{aligned}$$

so that again by the Burkholder-Davis-Gundy inequality and Doob's maximal inequality, finiteness of $E\overline{G}_t^4$ implies finiteness of $E[\overline{G},\overline{G}]_t^2$ and hence of EG_t^4. That $E\overline{G}_t^4 < \infty$ was already used in Klüppelberg et al. (2004).

Denote by E_r the conditional expectation with respect to the σ-algebra $\mathcal{F}_r$. Using partial integration, we have

$$\begin{aligned}
(G_h^{(r)})^2 &= 2\int_{h+}^{h+r} G_{s-}\, dG_s + [G,G]_{h+}^{h+r} \\
&= 2\int_h^{h+r} G_{s-}\sqrt{V_s}\, dL_s + \sum_{h<s\leq h+r} V_s(\Delta L_s)^2.
\end{aligned}$$

Since the increments of L on the interval $(h, h+r]$ are independent of $\mathcal{F}_r$ and since L has expectation 0, it follows that

$$E_r \int_{h+}^{h+r} G_{s-}\sqrt{V_s}\, dL_s = 0.$$

Recall that $\mathbf{Y}_s = J_{r,s}\mathbf{Y}_r + \mathbf{K}_{r,s}$ by (2.9). Hence we also have $\mathbf{Y}_{s-} = J_{r,s-}\mathbf{Y}_r + \mathbf{K}_{r,s-}$, so that by the compensation formula,

$$\begin{aligned}
E_r(G_h^{(r)})^2 &= E_r \sum_{h<s\le h+r} (\alpha_0 + \mathbf{a}'\mathbf{Y}_{s-})(\Delta L_s)^2 \\
&= E_r \sum_{h<s\le h+r} (\alpha_0 + \mathbf{a}' J_{r,s-}\mathbf{Y}_r + \mathbf{a}'\mathbf{K}_{r,s-})(\Delta L_s)^2 \\
&= E(L_1^2)\alpha_0 r + E(L_1^2)\mathbf{a}' \int_{h+}^{h+r} (EJ_{r,s-})\mathbf{Y}_r\, ds \\
&\quad + E(L_1^2)\mathbf{a}' \int_{h+}^{h+r} (E\mathbf{K}_{r,s-})\, ds \\
&= E(L_1^2) \int_h^{h+r} E_r(V_s)\, ds.
\end{aligned} \tag{2.64}$$

Since $\mathbf{Y}_\infty \stackrel{d}{=} J_{r,s}\mathbf{Y}_\infty + \mathbf{K}_{r,s}$ by (2.11), with $\mathbf{Y}_\infty$ independent on the right hand side, and $EJ_{r,s} = e^{\widetilde{B}(s-r)}$ by the proof of Theorem 2.13, it follows from (2.39) that

$$E\mathbf{K}_{r,s} = (I - e^{\widetilde{B}(s-r)})\frac{\alpha_0\mu}{\beta_q - \alpha_1\mu}\mathbf{e}_1.$$

Hence

$$\begin{aligned}
E_r(V_s) &= \alpha_0 + \mathbf{a}' e^{\widetilde{B}(s-r)}\mathbf{Y}_r + \mathbf{a}'\frac{\alpha_0\mu}{\beta_q - \alpha_1\mu}(I - e^{\widetilde{B}(r-s)})\mathbf{e}_1 \\
&= \frac{\alpha_0\beta_q}{\beta_q - \alpha_1\mu} + \mathbf{a}' e^{\widetilde{B}(s-r)}\left(\mathbf{Y}_r - \frac{\alpha_0\mu}{\beta_q - \alpha_1\mu}\mathbf{e}_1\right).
\end{aligned} \tag{2.65}$$

Combining $\int_h^{h+r} e^{\widetilde{B}(s-r)}\, ds = e^{\widetilde{B}h}\widetilde{B}^{-1}(I - e^{-\widetilde{B}r})$ with (2.64), (2.65) and (2.39) gives

$$E_r(G_h^{(r)})^2 = E(L_1^2)\left(\frac{\alpha_0 r\beta_q}{\beta_q - \alpha_1\mu} + \mathbf{a}' e^{\widetilde{B}h}\widetilde{B}^{-1}(I - e^{-\widetilde{B}r})\,(\mathbf{Y}_r - E\mathbf{Y}_r)\right),$$

and we conclude with (2.60) that

$$\begin{aligned}
&E((G_0^{(r)})^2(G_h^{(r)})^2) = E(E_r((G_h^{(r)})^2 G_r^2)) \\
&= E(L_1^2)\, E\left(\frac{\alpha_0 r\beta_q}{\beta_q - \alpha_1\mu}G_r^2 + \mathbf{a}' e^{\widetilde{B}h}\widetilde{B}^{-1}(I - e^{-\widetilde{B}r})(\mathbf{Y}_r - E\mathbf{Y}_r)G_r^2\right) \\
&= (E(G_r^2))^2 + E(L_1^2)\mathbf{a}' e^{\widetilde{B}h}\widetilde{B}^{-1}(I - e^{-\widetilde{B}r})\left[E(\mathbf{Y}_r G_r^2) - (E\mathbf{Y}_r)(EG_r^2)\right],
\end{aligned}$$

showing (2.62). To calculate $\operatorname{cov}(\mathbf{Y}_r, G_r^2)$, integrate by parts to get

$$G_r^2 = 2\int_0^r G_{s-}\,dG_s + [G,G]_r = 2\int_0^r G_{s-}\sqrt{V_s}\,dL_s + \int_0^r V_s\,d[L,L]_s,$$

therefore

$$\begin{aligned}\operatorname{cov}(\mathbf{Y}_r, G_r^2) &= \operatorname{cov}(\mathbf{Y}_r, 2\int_0^r G_{s-}\sqrt{V_s}\,dL_s + \int_0^r V_s\,d[L,L]_s) \\ &= 2\operatorname{cov}(\mathbf{Y}_r, \int_0^r G_{s-}\sqrt{V_s}\,dL_s) + \operatorname{cov}(\mathbf{Y}_r, \int_0^r V_s\,d[L,L]_s).\end{aligned}$$

To calculate the first term, let $I_r := \int_0^r G_{s-}\sqrt{V_s}\,dL_s$. We already have that $E(I_t) = 0$ and that $\int_0^r G_{s-}V_s\sqrt{V_s}\,dL_s = 0$. Integrating by parts and substituting $dV_{r+} = \mathbf{a}'B\mathbf{Y}_r\,dr + \alpha_q V_r\,d[L,L]_r^{(d)}$,

$$\begin{aligned}I_r V_{r+} &= \int_0^r I_{s-}\,dV_{r+} + \int_0^r V_s\,dI_s + [V_+, I]_r \\ &= \mathbf{a}'B\int_0^r I_{s-}\mathbf{Y}_s\,ds + \alpha_q\int_0^r I_{s-}V_s\,d[L,L]_s^{(d)} + \int_0^r G_{s-}V_s\sqrt{V_s}\,dL_s \\ &\quad + \left[\alpha_0 + \mathbf{a}'B\int_0^{\cdot}\mathbf{Y}_s\,ds + \alpha_q\int_0^{\cdot} V_s\,d[L,L]_s^{(d)}, \int_0^{\cdot} G_{s-}\sqrt{V_s}\,dL_s\right]_r.\end{aligned}$$

Taking expectation,

$$\begin{aligned}E(I_r V_{r+}) &= \mathbf{a}'B\int_0^r E(I_{s-}\mathbf{Y}_s)\,ds + \alpha_q E(L_1^2)\int_0^r E(I_{s-}V_s)\,ds + 0 \\ &\quad + \alpha_q E\int_0^r V_s\sqrt{V_s}G_{s-}\,d\sum_{0<u\le s}(\Delta L_u)^3 \\ &= \mathbf{a}'B\int_0^r E(I_{s-}\mathbf{Y}_s)\,ds + \alpha_q E(L_1^2)\mathbf{a}'\int_0^r E(I_{s-}\mathbf{Y}_{s-})\,ds.\end{aligned}$$

where we have used $\int_{\mathbb{R}} x^3\nu_L(dx) = 0$ and $E(I_r V_{r+}) = \mathbf{a}'E(I_r\mathbf{Y}_r)$. $I_s\mathbf{Y}_s = I_{s-}\mathbf{Y}_s = I_{s-}\mathbf{Y}_{s-}$ almost surely for fixed s, so we have

$$\mathbf{a}'E(I_s\mathbf{Y}_s) = \mathbf{a}'(B + \alpha_q E(L_1^2)I)\int_0^r E(I_s\mathbf{Y}_s)\,ds.$$

The equality holds for any vector $\mathbf{a}$, hence

$$E(I_r\mathbf{Y}_r) = (B + \alpha_q E(L_1^2)I)\int_0^r E(I_s\mathbf{Y}_s)\,ds.$$

Solving the integral equation and using $I_0 = 0$, we get $E(I_r \mathbf{Y}_r) = 0$ for all $r \geq 0$. So, the first term of $\text{cov}(\mathbf{Y}_r, G_r^2)$ is equal to 0.

To calculate the second term of the covariance, start with

$$\begin{aligned} d\mathbf{Y}_r &= B\mathbf{Y}_{r-}\,dr + \mathbf{e}V_r\,d[L,L]_r, \\ \mathbf{e}V_r\,d[L,L]_r &= d\mathbf{Y}_r - B\mathbf{Y}_{r-}\,dr, \\ V_r\,d[L,L]_r &= (d\mathbf{Y}_r - B\mathbf{Y}_{r-}\,dr)'\mathbf{e}, \end{aligned}$$

therefore

$$\begin{aligned} \int_0^r V_s\,d[L,L]_s &= \int_0^r (d\mathbf{Y}_s - B\mathbf{Y}_{s-}\,ds)'\mathbf{e} \\ &= (\mathbf{Y}_r - \mathbf{Y}_0 - \int_0^r B\mathbf{Y}_{s-}\,ds)'\mathbf{e} \\ &= (\mathbf{Y}_r' - \mathbf{Y}_0' - \int_0^r \mathbf{Y}_{s-}'\,ds\,B')\mathbf{e}. \end{aligned}$$

Thus, by the stationarity of $\mathbf{Y}$ and using $\text{cov}(\mathbf{Y}_r, \mathbf{Y}_0) = e^{\widetilde{B}r}\text{cov}(\mathbf{Y}_0)$,

$$\begin{aligned} \text{cov}(\mathbf{Y}_r, \int_0^r V_s\,d[L,L]_s) &= (\text{cov}(\mathbf{Y}_r) - \text{cov}(\mathbf{Y}_r, \mathbf{Y}_0) - \text{cov}(\mathbf{Y}_r, \int_0^r \mathbf{Y}_{s-}\,ds)B')\mathbf{e} \\ &= (\text{cov}(\mathbf{Y}_0) - e^{\widetilde{B}r}\text{cov}(\mathbf{Y}_0) - \int_0^r e^{\widetilde{B}(r-s)}\text{cov}(\mathbf{Y}_{s-})\,ds\,B')\mathbf{e} \\ &= ((I - e^{\widetilde{B}r})\text{cov}(\mathbf{Y}_0) - \widetilde{B}^{-1}(e^{\widetilde{B}r} - I)\text{cov}(\mathbf{Y}_0)B')\mathbf{e}. \end{aligned}$$

Finally, to calculate $\text{var}((G_t^{(r)})^2)$, integrate by parts and get

$$\begin{aligned} G_r^4 &= 2\int_0^r G_{s-}^2\,dG_s^2 + [G^2, G^2]_r \\ &= 2\int_0^r G_{s-}^2\big(2G_{s-}\sqrt{V_s}\,dL_s + V_s\,d[L,L]_s\big) \\ &\quad + \left[2\int_0^{\cdot} G_{s-}\sqrt{V_s}\,dL_s + \int_0^{\cdot} V_s\,d[L,L]_s, 2\int_0^{\cdot} G_{s-}\sqrt{V_s}\,dL_s + \int_0^{\cdot} V_s\,d[L,L]_s\right]_r \\ &= 4\int_0^r G_{s-}^3\sqrt{V_s}\,dL_s + 2\int_0^r G_{s-}^2 V_s\,d[L,L]_s + 4\int_0^r G_{s-}^2 V_s\,d[L,L]_s \\ &\quad + 4\int_0^r G_{s-}V_s\sqrt{V_s}\,d\big[[L,L],L\big]_s + \int_0^r V_s^2\,d\big[[L,L],[L,L]\big]_s, \end{aligned}$$

Taking expectation and using $E(L_1) = E(L_1^3) = 0$,

$$\begin{aligned}
E(G_r^4) &= 6E(L_1^2)\int_0^r E(G_{s-}^2 V_s)\,ds + E(L_1^4)\int_0^r E(V_s^2)\,ds \\
&= 6E(L_1^2)\int_0^r \left[\mathrm{cov}(G_{s-}^2, V_s) + E(G_{s-}^2)E(V_s)\right] ds + rE(L_1^4)E(V_\infty^2) \\
&= 6E(L_1^2)\int_0^r \left[\mathbf{a}'\mathrm{cov}(G_{s-}^2, \mathbf{Y}_{s-}) + sE(L_1^2)(E(V_\infty))^2\right] ds + rE(L_1^4)E(V_\infty^2) \\
&= 6E(L_1^2)\,\mathbf{a}'\int_0^r \mathrm{cov}(G_{s-}^2, \mathbf{Y}_{s-})\,ds + 3r^2(E(L_1^2)E(V_\infty))^2 + rE(L_1^4)E(V_\infty^2).
\end{aligned}$$

The integral in the first term is

$$\begin{aligned}
\mathbf{K}_r &:= \int_0^r \mathrm{cov}(G_{s-}^2, \mathbf{Y}_{s-})\,ds = \int_0^r \mathrm{cov}(G_s^2, \mathbf{Y}_s)\,ds \\
&= \int_0^r \left[(I - e^{\widetilde{B}s})\mathrm{cov}(\mathbf{Y}_\infty) - \widetilde{B}^{-1}(e^{\widetilde{B}s} - I)\mathrm{cov}(\mathbf{Y}_\infty)B'\right]\mathbf{e}\,ds \\
&= \left[(rI - \widetilde{B}^{-1}(e^{\widetilde{B}r} - I))\mathrm{cov}(\mathbf{Y}_\infty) - \widetilde{B}^{-1}(\widetilde{B}^{-1}(e^{\widetilde{B}r} - I) - rI)\mathrm{cov}(\mathbf{Y}_\infty)B'\right]\mathbf{e}.
\end{aligned}$$

So we have the variance of $(G_t^{(r)})^2$ as follows,

$$\begin{aligned}
\mathrm{var}((G_t^{(r)})^2) &= E((G_0^{(r)})^4) - (E((G_0^{(r)})^2))^2 = E(G_r^4) - (E(G_r^2))^2 \\
&= 6E(L_1^2)\,\mathbf{a}'\mathbf{K}_r + 2(rE(L_1^2)E(V_\infty))^2 + rE(L_1^4)E(V_\infty^2). \quad \square
\end{aligned}$$

The autocovariance function (2.62) and the variance (2.63) can also be written down, using (2.27) and (2.44)–(2.46), as the second degree polynomials in terms of parameter $\mathbf{a}$:

$$\gamma(0) := \mathrm{var}((G_t^{(r)})^2) = \frac{\alpha_0^2\beta_q^2}{(1-m)(\beta_q - \mu\alpha_1)^2}(\mathbf{a}'P_0\,\mathbf{a} + \mathbf{a}'\mathbf{Q}_0 + R), \tag{2.66}$$

$$\gamma(h) := \mathrm{cov}((G_t^{(r)})^2, (G_{t+h}^{(r)})^2) = \frac{\alpha_0^2\beta_q^2}{(1-m)(\beta_q - \mu\alpha_1)^2}(\mathbf{a}'P_h\,\mathbf{a} + \mathbf{a}'\mathbf{Q}_h), \quad h \geq r. \tag{2.67}$$

The autocorrelation function is therefore written as follows:

$$\rho(h) := \mathrm{corr}((G_t^{(r)})^2, (G_{t+h}^{(r)})^2) = \frac{\mathbf{a}'P_h\,\mathbf{a} + \mathbf{a}'\mathbf{Q}_h}{\mathbf{a}'P_0\,\mathbf{a} + \mathbf{a}'\mathbf{Q}_0 + R_0}, \quad h \geq r, \tag{2.68}$$

where

$$\begin{aligned}
P_0 &= 2\mu^2\big[3\widetilde{B}^{-1}(\widetilde{B}^{-1}(e^{\widetilde{B}r}-I)-rI)-I\big]\mathrm{cov}(\zeta_\infty),\\
\mathbf{Q}_0 &= 6\mu\big[(rI-\widetilde{B}^{-1}(e^{\widetilde{B}r}-I))\mathrm{cov}(\zeta_\infty)\\
&\quad -\widetilde{B}^{-1}(\widetilde{B}^{-1}(e^{\widetilde{B}r}-I)-rI)\mathrm{cov}(\zeta_\infty)\widetilde{B}'\big]\mathbf{e},\\
R_0 &= 2r^2\mu^2+\rho
\end{aligned}$$

and

$$\begin{aligned}
P_h &= \mu^2 e^{\widetilde{B}h}\widetilde{B}^{-1}(I-e^{-\widetilde{B}r})\widetilde{B}^{-1}(e^{\widetilde{B}r}-I)\mathrm{cov}(\zeta_\infty),\\
\mathbf{Q}_h &= \mu e^{\widetilde{B}h}\widetilde{B}^{-1}(I-e^{-\widetilde{B}r})\big[(I-e^{\widetilde{B}r})\mathrm{cov}(\zeta_\infty)\\
&\quad -\widetilde{B}^{-1}(e^{\widetilde{B}r}-I)\mathrm{cov}(\zeta_\infty)\widetilde{B}'\big]\mathbf{e}.
\end{aligned}$$

Notice that the autocorrelation function is free of α_0 and β_q, and is a function of only $\tilde{\beta}_1,\ldots,\tilde{\beta}_q$, therefore enables us to estimate the parameters $\tilde{\beta}_1,\ldots,\tilde{\beta}_q$ using the sample autocorrelations. The matrices P_0, P_h, $\mathbf{Q}_0$ and $\mathbf{Q}_h$ are easily calculated since $\mathrm{cov}(\zeta_\infty)$ is straightforward to find, for example, when $q=2$,

$$\mathrm{cov}(\zeta_\infty)=\frac{\rho}{2\tilde{\beta}_1\tilde{\beta}_2}\begin{pmatrix}1 & 0\\ 0 & \tilde{\beta}_2\end{pmatrix},$$

and when $q=3$,

$$\mathrm{cov}(\zeta_\infty)=\frac{\rho}{2(\tilde{\beta}_1\tilde{\beta}_2-\tilde{\beta}_3)}\begin{pmatrix}\frac{\tilde{\beta}_1}{\tilde{\beta}_3} & 0 & -1\\ 0 & 1 & 0\\ -1 & 0 & \tilde{\beta}_2\end{pmatrix},$$

etc.

2.7 An Example

In this section we illustrate the properties established above using the COGARCH(1,3) process driven by a compound Poisson process with jump-rate 2 and normally distributed jumps with mean zero and variance 0.74. The COGARCH coefficients are

$\alpha_0 = \alpha_1 = 1, \beta_1 = 1.2, \beta_2 = .48 + \pi^2$, and $\beta_3 = .064 + .4\pi^2$, from which we find that the eigenvalues of B are $-.4$, $-.4 + \pi i$ and $-.4 - \pi i$. With S defined as in (2.7), $||S^{-1}\mathbf{e}\mathbf{a}'S||_2 = 0.21493$ and it is easy to check from this that the conditions (2.28) and (2.29) are satisfied. Condition (b)(ii) of Theorem 2.18 also implies that the volatility process is non-negative.

The eigenvalues of the matrix $\widetilde{B} = B + \mu\mathbf{e}\mathbf{a}'$ are $-.25038$, $-.47481 + 3.14426i$ and $-.47481 - 3.14426i$. From (2.49) we conclude that the autocorrelation of the volatility in this case is a linear combination of $\exp(-.25038t)$ and a damped sinusoid with period approximately equal to 2 and damping factor $\exp(-.47481t)$.

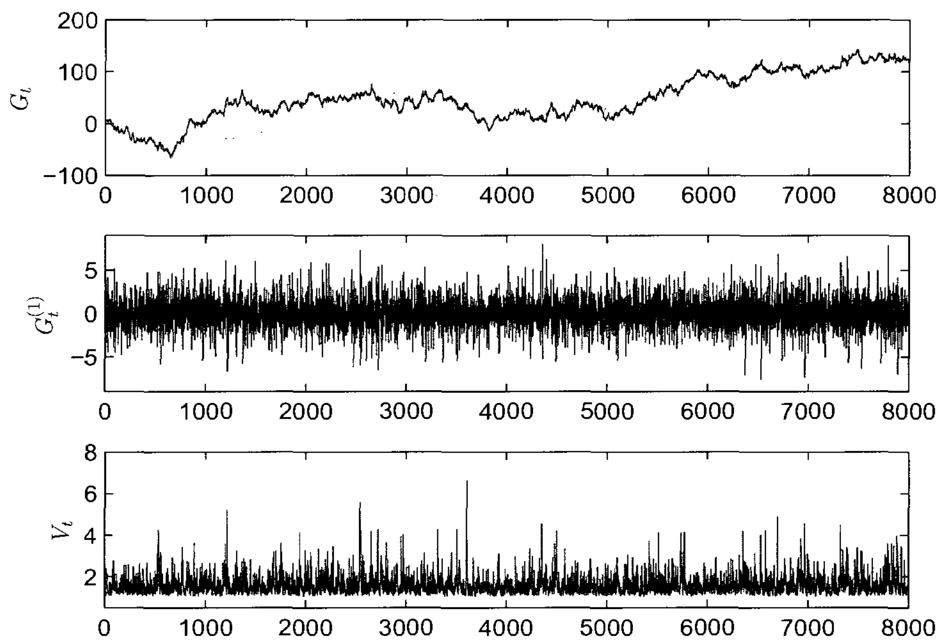

Figure 3: The simulated compound-Poisson driven COGARCH(1,3) process with jump-rate 2, normally distributed jumps with mean zero and variance 0.74 and coefficients $\alpha_0 = \alpha_1 = 1$, $\beta_1 = 1.2$, $\beta_2 = .48 + \pi^2$ and $\beta_3 = .064 + .4\pi^2$. The graphs show the process (G_t) sampled at integer times (top),the corresponding increments $(G_t^{(1)} = G_{t+1} - G_t)$ (centre), and the corresponding volatility sequence $(V_t = \sigma_t^2)$ (bottom).

The top graph in Figure 3 shows the values at integer times $101, \ldots, 8100$ of a simulated series (G_t) with the parameters specified above, $\mathbf{Y}_0 = (1,1,1)'$ and $G(0) = 0$. The second graph shows the differenced series $(G_{t+1} - G_t)_{t=100,\ldots,8099}$ and the last graph shows the volatility $(\sigma_t^2)_{t=101,\ldots,8100}$.

As is the case for a discrete-time GARCH process, the increments $(G_{t+1} - G_t)$ exhibit no significant correlation, but the squared increments $((G_{t+1} - G_t)^2)$ have highly significant correlations as shown in the second graph of Figure 4. The first graph in Figure 4 shows the sample autocorrelation function of the volatility process at integer lags. This too is highly significant for large lags, reflecting the long-memory property frequently observed in financial time series. As expected from the remarks in the first paragraph above, it has the form of an exponentially decaying term plus a small damped sinusoidal term with period approximately equal to two.

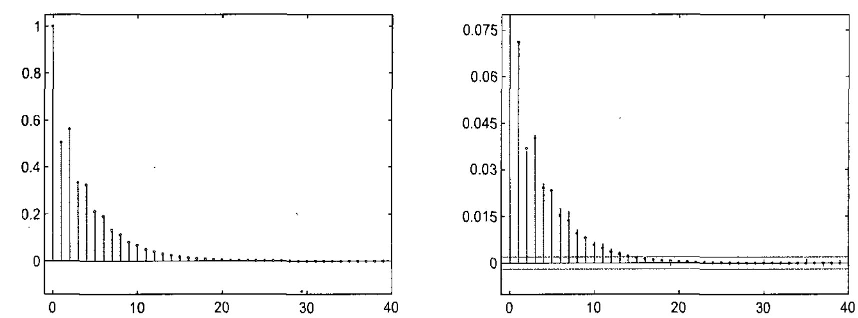

Figure 4: The sample autocorrelation functions of the volatilities (V_t) (left) and of the squared COGARCH increments $((G_{t+1} - G_t)^2)$ (right) of a realisation of length 1000000 of the COGARCH process with parameters as specified in Figure 3. The red dots indicate the theoretical autocorrelations.

3 Parameter estimation

As mentioned earlier, the ACF $\rho(h)$, $h \geq 1$ of the process $((G_t^{(r)})^2)_{t\geq 0}$ is a linear combination of $e^{\tilde{\lambda}_j h}$, $j = 1, \dots, q$, where $\tilde{\lambda}_1, \dots, \tilde{\lambda}_q$ are the eigenvalues of $\widetilde{B}$ (see (2.62)-(2.63)). It is therefore the ACF of an ARMA(q,q) process, namely

$$\rho(h) = c_1\xi_1^{-h} + c_2\xi_2^{-h} + \cdots + c_q\xi_q^{-h}, \quad h = 1, 2, \dots \tag{3.1}$$

where $\xi_j = e^{-\tilde{\lambda}_j}$, $j = 1, \dots, q$ are the autoregressive roots of the ARMA(q,q) process, and $c_1, \dots, c_q$ are constants. Fitting an ARMA(q,q) process that gives the same ACF as that of the squared increment process $((G_t^{(r)})^2)_{t\geq 0}$ is the main idea of the estimation procedure proposed in this chapter.

3.1 Preliminary estimation

The autoregressive coefficients $\boldsymbol{\phi} = (\phi_1, \dots, \phi_q)'$ can be initially estimated by using the sample autocorrelations of the squared increment process $((G_t^{(r)})^2)_{t\geq 0}$. Multiplying the identity $\phi(\xi_j) = 0$ by $c_j\xi_j^{-i}$ and summing over $j = 1, \dots, q$, we find that

$$\rho(q+i) - \phi_1\rho(q+i-1) - \phi_2\rho(q+i-2) - \cdots - \phi_q\rho(i) = 0, \quad i = 1, 2, \dots \tag{3.2}$$

Writing (3.2) for $i = 1, \dots, q$, we obtain the following system of linear equations of variables $\boldsymbol{\phi}$.

$$\mathbf{R}\boldsymbol{\phi} = \boldsymbol{\rho}_q, \tag{3.3}$$

where $\boldsymbol{\rho}_q = (\rho(q+1), \dots, \rho(2q))'$ and $\mathbf{R}$ is defined as

$$\mathbf{R} := \begin{pmatrix} \rho(q) & \rho(q-1) & \cdots & \rho(1) \\ \rho(q+1) & \rho(q) & \cdots & \rho(2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(2q-1) & \rho(2q-2) & \cdots & \rho(q) \end{pmatrix}. \tag{3.4}$$

In order the equations (3.3) have unique solution, the matrix $\mathbf{R}$ must be non-singular. The determinant of $\mathbf{R}$ is calculated in the next lemma.

Lemma 3.1.

$$\det \mathbf{R} = \prod_{i=1}^{q} c_i \xi_i^{-1} \prod_{1 \le i < j \le q} (\xi_i^{-1} - \xi_j^{-1})^2 \tag{3.5}$$

Proof. Define a column vector

$$\xi_i^{-(j)} := [\xi_i^{-j}, \ldots, \xi_i^{-q-j+1}]', \quad i, j = 1, \ldots, q.$$

By the multilinearity property and using Vandermonde determinant (see Rao and Rao (1998)),

$$\begin{aligned}
\det \mathbf{R} &= \det[\sum_{i=1}^{q} c_i \xi_i^{-(1)}, \ldots, \sum_{i=1}^{q} c_i \xi_i^{-(q)}] \\
&= \prod_{i=1}^{q} c_i \sum_{\pi \in \Pi} \det[\xi_{\pi(1)}^{-(1)}, \ldots, \xi_{\pi(q)}^{-(q)}] \\
&= \prod_{i=1}^{q} c_i \sum_{\pi \in \Pi} \xi_{\pi(1)}^{-1} \cdots \xi_{\pi(q)}^{-q} \det[\xi_{\pi(1)}^{-(0)}, \ldots, \xi_{\pi(q)}^{-(0)}] \\
&= \prod_{i=1}^{q} c_i \xi_i^{-1} \sum_{\pi \in \Pi} \xi_{\pi(1)}^{0} \cdots \xi_{\pi(q)}^{-q+1} \operatorname{sign}(\pi) \det[\xi_1^{-(0)}, \ldots, \xi_q^{-(0)}] \\
&= \prod_{i=1}^{q} c_i \xi_i^{-1} \sum_{\pi \in \Pi} \operatorname{sign}(\pi) \xi_{\pi(1)}^{0} \cdots \xi_{\pi(q)}^{-q+1} \prod_{1 \le i < j \le q} (\xi_i^{-1} - \xi_j^{-1}) \\
&= \prod_{i=1}^{q} c_i \xi_i^{-1} \prod_{1 \le i < j \le q} (\xi_i^{-1} - \xi_j^{-1}) \sum_{\pi \in \Pi} \operatorname{sign}(\pi) \varphi_{\pi(1)}^{0} \cdots \xi_{\pi(q)}^{-q+1} \\
&= \prod_{i=1}^{q} c_i \xi_i^{-1} \prod_{1 \le i < j \le q} (\xi_i^{-1} - \xi_j^{-1})^2,
\end{aligned}$$

where $\operatorname{sign}(\pi)$ is 1 if π is an even permutation, -1 if π is an odd permutation. □

Since $\xi_1, \ldots, \xi_q$ are assumed to be distinct, from (3.5), $\det \mathbf{R} \neq 0$, therefore the equations (3.3) have a unique solution $\boldsymbol{\phi}$.

Replacing the autocorrelations $\rho(1), \ldots, \rho(2q)$ by the corresponding sample autocorrelations $\hat{\rho}(1), \ldots, \hat{\rho}(2q)$, we obtain the preliminary estimators $\hat{\boldsymbol{\phi}}$ of $\boldsymbol{\phi}$ which can be used as starting point for the least-squares estimation in Section 3.2. The preliminary estimator is not very efficient because it uses the first $2q$ lags of the autocorrelations and the determinant (3.5) is very close to zero therefore a small variation in the sample autocorrelations would result in a significant change in the determinant (3.5), hence in the estimators of ϕ.

The preliminary estimator of the autoregressive coefficients ϕ of $((G_t^{(1)})^2)_{t\geq 0}$ is strongly consistent in the case of a compound Poisson driven COGARCH(2,2) process. Let the estimator of $\phi = (\phi_1, \ldots, \phi_q)'$ based on n observations be denoted as $\hat{\phi}_n$.

Theorem 3.2. *For the compound Poisson driven COGARCH(2,2) process $\hat{\phi}_n$ converges to ϕ almost surely, as $n \to \infty$.*

Proof. The squared increment process $((G_t^{(1)})^2)_{t\geq 0}$ is mixing, therefore ergodic. Hence, by the theorem IV.2.2 of Hannan (1970) the empirical moments and the empirical autocovariances converge almost surely to their theoretical counterparts:

$$\hat{m} \xrightarrow{\text{a.s.}} E(G_t^{(1)})^2,$$

$$\hat{\gamma}(h) \xrightarrow{\text{a.s.}} \gamma(h), \quad h = 0, 1, \ldots,$$

as $N \to \infty$. The preliminary estimators of ϕ are continuous functions of the sample autocovariances and the sample moment of the observed data: $\hat{m}, \hat{\gamma}(0), \ldots, \hat{\gamma}(2q)$, hence we conclude that

$$\hat{\phi}_n \xrightarrow{\text{a.s.}} \phi \text{ as } n \to \infty,$$

i.e. the preliminary estimators are strongly consistent. □

3.2 Least-squares estimators

The squared increment process $((G_t^{(r)})^2)_{t\geq 0}$ has an autocovariance structure of an ARMA(q,q) process. Thus, fitting an ARMA(q,q) process to the squared increment process would enable us to estimate the parameters of the COGARCH process. The noise involved in the ARMA equation in Chapter 1 is assumed to be an i.i.d. sequence, or a martingale difference sequence. However, the noise sequence in the squared increment process $((G_t^{(r)})^2)_{t\geq 0}$ does not satisfy such strong assumptions.

However, under strong mixing and moment conditions, the least squares estimators (LSE) of ARMA representations in which the noise is the linear innovation process, so called *weak ARMA models*, are strongly consistent and asymptotically normal (see Francq and Zakoïan (1998)). In the case of the compound Poisson driven COGARCH(2,2) process the squared increment process $((G_t^{(r)})^2)_{t\geq 0}$ satisfies these conditions, as will be shown in Section 3.3, therefore we can find estimators of the ARMA parameters of $((G_t^{(r)})^2)_{t\geq 0}$ which lead in turn to strongly consistent and asymptotically normal estimators for the COGARCH parameters.

We now introduce the notation of Francq and Zakoïan (1998) below. Let $(X_t)_{t\geq 0}$ be a second-order stationary process satisfying

$$X_t = \sum_{i=1}^{q} \phi_i X_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t, \tag{3.6}$$

where $\{\varepsilon_t\}$ is a sequence of uncorrelated random variables with zero mean and common variance σ^2, and the polynomials $\phi(z) = 1 - \phi_1 z - \ldots - \phi_q z^q$ and $\theta(z) = 1 + \theta_1 z + \ldots + \theta_q z^q$ have all their zeros outside the unit disk and have no zero in common. Let $\boldsymbol{\beta}_0 = (\phi_1, \ldots, \phi_q, \theta_1, \ldots, \theta_q)'$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{2q})'$, and denote by Θ the parameter space

$$\Theta := \{\boldsymbol{\beta} \in \mathbb{R}^{2q} : \sum_{i=1}^{q} \beta_i z^i \text{ and } \sum_{j=q+1}^{2q} \beta_j z^j \text{ have all their zeros outside the unit disk}\}.$$

For all $\boldsymbol{\beta} \in \Theta$, let $\{\varepsilon_t(\boldsymbol{\beta})\}$ be the second-order stationary process (See Brockwell and Davis (1991), Chapter 3 for the existence and uniqueness of such a process) defined as the solution of

$$\varepsilon_t(\boldsymbol{\beta}) = X_t - \sum_{i=1}^{q} \beta_i X_{t-i} - \sum_{j=1}^{q} \beta_{q+j} \varepsilon_{t-j}(\boldsymbol{\beta}), \quad t \in \mathbb{Z}. \tag{3.7}$$

If $\beta_i = \phi_i$, $i = 1, \ldots, q$ and $\beta_j = \theta_{j-q}$, $j = q+1, \ldots, 2q$ then $\{\varepsilon_t(\boldsymbol{\beta})\}$ is the linear innovation of $\{X_t\}$, i.e.

$$\varepsilon_t = X_t - E(X_t | H_X(t-1)) \tag{3.8}$$

where $H_X(t-1)$ is the Hilbert space generated by $(X_s, s < t)$.

Given a realization of length n, $X_1, \ldots, X_n$, $\varepsilon_t(\boldsymbol{\beta})$ can be approximated, for $0 < t \leq n$, by $e_t(\boldsymbol{\beta})$ defined recursively by

$$e_t(\boldsymbol{\beta}) = X_t - \sum_{i=1}^{q} \beta_i X_{t-i} - \sum_{j=1}^{q} \beta_{q+j} e_{t-j}(\boldsymbol{\beta}), \tag{3.9}$$

where the starting values are $e_0(\boldsymbol{\beta}) = \cdots = e_{-q+1}(\boldsymbol{\beta}) = 0$ and $X_0 = \cdots = X_{-q+1} = 0$. Let δ be a strictly positive constant chosen so that the true parameter $\boldsymbol{\beta}_0$ belongs to the compact set

$$\Theta_\delta := \{\boldsymbol{\beta} \in \mathbb{R}^{2q}; \text{the zeros of } \sum_{i=1}^{q} \beta_i z^i \text{ and } \sum_{j=q+1}^{2q} \beta_j z^j \text{ have moduli } \geq 1 + \delta\}.$$

The random variable $\hat{\boldsymbol{\beta}}_n$ is called the least-squares estimator if it satisfies, almost surely,

$$\hat{\boldsymbol{\beta}}_n = \arg\min_{\boldsymbol{\beta} \in \Theta_\delta} \frac{1}{n} \sum_{t=1}^{n} e_t^2(\boldsymbol{\beta}) \tag{3.10}$$

For any $\boldsymbol{\beta} \in \Theta_\delta$, let

$$O_n(\hat{\boldsymbol{\beta}}_n) = \frac{1}{n} \sum_{t=1}^{n} \varepsilon_t^2(\boldsymbol{\beta})$$

and $\frac{\partial}{\partial \boldsymbol{\beta}} O_n(\boldsymbol{\beta}) = (\frac{\partial}{\partial \beta_1} O_n(\boldsymbol{\beta}), \ldots, \frac{\partial}{\partial \beta_{p+q}} O_n(\boldsymbol{\beta}))'$. Consider the following matrices:

$$I(\boldsymbol{\beta}) = \lim_{n \to \infty} \mathrm{Var}\Big(\sqrt{n} \frac{\partial}{\partial \boldsymbol{\beta}} O_n(\boldsymbol{\beta})\Big)$$

and

$$J(\boldsymbol{\beta}) = \lim_{n \to \infty} \mathrm{Var}\Big(\frac{\partial^2}{\partial \beta_i \partial \beta_j} O_n(\boldsymbol{\beta})\Big),$$

where $[A(i,j)]$ denotes the matrix A with elements $A(i,j)$.

The following is the main result that shows the strong consistency and the asymptotic normality of the least-squares estimators.

Theorem 3.3 (Francq and Zakoïan (1998)). *Let $(X_t)_{t \in \mathbb{Z}}$ be a strictly stationary ergodic process satisfying* (3.6). *Let $(\hat{\boldsymbol{\beta}}_n)$ be a sequence of least-squares estimators defined by* (3.10). *Suppose $\boldsymbol{\beta}_0 \in \Theta_\delta$. Then*

$$\hat{\boldsymbol{\beta}}_n \to \boldsymbol{\beta}_0 \text{ a.s. as } n \to \infty.$$

If in addition, $(X_t)_{t\in\mathbb{Z}}$ *satisfies* $E|X_t|^{4+2\nu} < \infty$ *and strongly mixing with the mixing rate such that* $\sum_{k=0}^{\infty} \alpha_k^{\nu/(2+\nu)} < \infty$ *for some* $\nu > 0$. *Then*

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_0) \sim N(0, \Sigma)$$

as $n \to \infty$, *where* $\Sigma := J^{-1}(\boldsymbol{\beta}_0) I(\boldsymbol{\beta}_0) J^{-1}(\boldsymbol{\beta}_0)$.

3.3 Strong consistency and asymptotic normality of the least-squares estimators

In this section we establish the conditions of Fracq and Zakoian for strong consistency and asymptotic normailty of the least-squares estimators of the COGARCH(2,2) parameters when the driving Lévy process is compound Poisson. We conjecture that the conditions hold more generally, but the results of this section cover the examples considered later in the chapter.

We'll show that the state process (and therefore the volatility process) of the COGARCH(2,2) process is strongly mixing with geometric rate when the driving Lévy process is compound Poisson. If the volatility process is strictly stationary and strongly mixing, then the squared increment process is also strongly mixing (see Haug et. al. (2007)). The mixing property guarantees strong consistency of the estimators proposed in Chapter 3 and the strong mixing property guarantees asymptotic normality of the estimators.

A stationary Markov chain $(\mathbf{Y}_n)$ is said to be *strongly mixing with geometric rate* when there exist constants C and $a \in (0, 1)$ such that

$$\sup_{f,g} |\mathrm{cov}\,(f(\mathbf{Y}_0), g(\mathbf{Y}_k))| =: \alpha_k \leq C a^k,$$

where the sup is taken over all measurable functions f and g with $|f| \leq 1$ and $|g| \leq 1$. The function α_k is called the *mixing rate function* of for the Markov chain $(\mathbf{Y}_n)$ and is equal to

$$\alpha_k = \sup_{f,g} |\mathrm{cov}\,(f(\mathbf{Y}_0), g(\mathbf{Y}_k))| = \sup_{A\in\sigma(\mathbf{Y}_0,),\, B\in\sigma(\mathbf{Y}_k)} |P(A \cap B) - P(A)P(B)|,$$

where the last equality follows from Doukhan (1994).

Before showing the mixing conditions, we first recall the following definitions. A Markov chain $\{\mathbf{Y}_n\}$ with state space $E \subset \mathbb{R}^q$ is said to be *μ-irreducible* for some measure μ on $(E, \mathcal{E})$ ($\mathcal{E}$ is the Borel σ-field on E), if

$$\sum_{n>0} p^n(\mathbf{y}, C) > 0 \quad \text{for all } \mathbf{y} \in E, \text{ whenever } \mu(C) > 0.$$

Here $p^n(\mathbf{y}, C)$ denotes the *n-step transition probability* of moving from $\mathbf{y}$ to the set C in n steps.

The following lemma shows the irreducibility of $\{\mathbf{Y}_{\Gamma_n}\}$ in the compound Poison case when the Z_1 has strictly positive density on $(0, \infty)$.

Lemma 3.4. *If B has distinct eigenvalues, $\lambda_2 < \lambda_1 < 0$ and conditions 2.8 and 2.53 are satisfied, then the Markov chain $\{\mathbf{Y}_{\Gamma_n}\}$ with state-space $\{\mathbf{x} = (x_1, x_2)' : x_2 > \max(\lambda_1 x_1, \lambda_2 x_1)\}$ is φ-irreducible where φ is the restriction of two-dimensional Lebesgue measure to $\mathbb{R}^2_+$.*

Proof. We'll give the proof for COGARCH(2,2) case since the simulations and applications in Chapter 3 are concentrated on COGARCH(2,2) model. We have

$$\mathbf{Y}_{\Gamma_1} = e^{B\Gamma_1}\mathbf{Y}_0 + \mathbf{e}(\alpha_0 + \mathbf{a}' e^{B\Gamma_1}\mathbf{Y}_0) Z_1.$$

Under the conditions stated, $\{\mathbf{Y}_{\Gamma_1}\}$ has strictly positive probability density with respect to φ on the subset of $\mathbb{R}^2_+$

$$S = \{(y_1, y_2) : 0 < y_1 < m\},$$

where

$$m := \max_{t \geq 0}(\mathbf{e}_1' e^{Bt}\mathbf{x}) = \begin{cases} \left(x_1 - \frac{x_2}{\lambda_1}\right)\left[\frac{\lambda_2(x_2 - \lambda_1 x_1)}{\lambda_1(x_2 - \lambda_2 x_1)}\right]^{\frac{\lambda_2}{\lambda_1 - \lambda_2}} \text{ if } x_2 > 0, \\ x_1 \text{ if } x_2 < 0, \end{cases}$$

and $\mathbf{e}_1 = (1, 0)'$, $\mathbf{e}_2 = \mathbf{e} = (0, 1)'$. The time t at which the maximum occurs is

$$t_0 := \begin{cases} \frac{1}{\lambda_1 - \lambda_2} \log\left(\frac{\lambda_2(x_2 - \lambda_1 x_1)}{\lambda_1(x_2 - \lambda_2 x_1)}\right) \text{ if } x_2 > 0, \\ 0 \text{ if } x_2 < 0. \end{cases}$$

Define the mapping $f = (f_1, f_2)' : \mathbb{R}_+^2 \to S$ as follows:

$$y_1 = f_1(t) = \mathbf{e}_1' e^{Bt}\mathbf{x},$$
$$y_2 = f_2(t, z) = \mathbf{e}_2' e^{Bt}\mathbf{x} + (\alpha_0 + \mathbf{a}' e^{Bt}\mathbf{x})z,$$

and define the following subset of S,

$$S_1 = \{(y_1, y_2) : x_1 < y_1 < m \text{ and } y_2 > \mathbf{e}_2' e^{Bt}\mathbf{x} \text{ for some } t \in (0, t_0)\}.$$

(S_1 may be empty, depending on $\mathbf{x}$.) Under the mapping f, each point in $S \setminus S_1$ has a unique inverse image in $(t_0, \infty) \times \mathbb{R}_+$ and the inverse mapping h_1 is differentiable with strictly positive Jacobian:

$$|\mathbf{J}| = \left|\frac{\partial(t, z)}{\partial(y_1, y_2)}\right| = \left|\frac{\partial(y_1, y_2)}{\partial(t, z)}\right|^{-1} = \begin{vmatrix} \frac{df_1}{dt} & 0 \\ \frac{df_2}{dt} & \frac{df_2}{dz} \end{vmatrix}^{-1}$$
$$= (\mathbf{e}e^{Bt}\mathbf{x}(\alpha_0 + \mathbf{a}' e^{Bt}\mathbf{x}))^{-1} > 0 \text{ for any } t \geq 0.$$

Each point in S_1 corresponds to two distinct points, one in $(t_0, \infty) \times \mathbb{R}_+$ and one in $(0, t_0) \times \mathbb{R}_+$. The two inverse mappings, h_1 and h_2 respectively, are differentiable with strictly positive Jacobian. Since the joint density of Γ_1 and Z_1 is strictly positive on $\mathbb{R}_+^2$ we conclude that, conditionally on $\mathbf{Y}_0 = \mathbf{x}$, $\mathbf{Y}_{\Gamma_1}$ has strictly positive density on S. Since the transition density $p(\mathbf{x}, \mathbf{y})$ of $\{\mathbf{Y}_{\Gamma_n}, n = 0, 1, 2, \ldots\}$ is strictly positive for $\mathbf{x}$ as specified and for every $\mathbf{y} \in S$, the two-step transition density $p^{(2)}(\mathbf{x}, \mathbf{z})$ is strictly positive for all $\mathbf{z}$ in the positive quadrant since

$$p^{(2)}(\mathbf{x}, \mathbf{z}) \geq \int_{\mathbf{y} \in S: y_2 > y_0} p(\mathbf{x}, \mathbf{y}) p(\mathbf{y}, \mathbf{z}) d\varphi,$$

$p(\mathbf{x}, \mathbf{y}) > 0$ for all $\mathbf{y} \in S$ and y_0 can be chosen sufficiently large to ensure that $p(\mathbf{y}, \mathbf{z}) > 0$ for all $\mathbf{y}$ such that $y_2 > y_0$. (This follows by the same argument used in the first paragraph, noting that $m \to \infty$ as $x_2 \to \infty$ for each fixed x_1). $\square$

We now state the result by Basrak et. al. (2002) which will be used to show that the process $(\mathbf{Y}_{\Gamma_n})_{n \in \mathbb{N}_0}$ is strongly mixing.

Theorem 3.5. *For the stochastic reccurence equation*

$$\mathbf{Y}_{\Gamma_{n+1}} = C_{n+1}\mathbf{Y}_{\Gamma_n} + \mathbf{D}_{n+1}, \quad n \in \mathbb{N}_0,$$

suppose there exists an $\epsilon > 0$ such that $E\|C_1\|^\epsilon < 1$ and $E|\mathbf{D}_1|^\epsilon < \infty$. If the Markov chain $(\mathbf{Y}_{\Gamma_n})_{n\in\mathbb{N}_0}$ is μ-irreducible, then it is geometrically ergodic and, hence, strongly mixing with geometric rate.

The following is the main result of this section.

Theorem 3.6. *Let $(\mathbf{Y}_t)_{t\geq 0}$ be the strictly stationary state process of a COGARCH(2,2) process that satisfies the conditions of Lemma 3.4. Then $(\mathbf{Y}_t)_{t\geq 0}$ is strongly mixing with geometric rate.*

Proof. By Lemma 3.4, the Markov chain $(\mathbf{Y}_{\Gamma_n})_{n\in\mathbb{N}_0}$ is μ-irreducible. It also satisfies the random recurrence equation (2.14). Recall that we have

$$\begin{aligned} C_i &= (I + Z_i\mathbf{e}\mathbf{a}')e^{BT_i}, \\ \mathbf{D}_i &= \alpha_0 Z_i\mathbf{e}. \end{aligned}$$

The condition $E\|C_1\|^\epsilon < 1$ for $\epsilon > 0$ in some neighborhood of zero is satisfied if $E\log\|C_1\| < 0$ and $E\|C_1\|^\delta < \infty$ for some $\delta > 0$. $E\log\|C_1\| < 0$ was shown during the proof of Theorem 2.8. It is straightforward to check

$$E\|C_1\| \leq E((1 + Z_1\|\mathbf{e}\mathbf{a}'\|)e^{\lambda T_1}) < \infty$$

and the other condition $E|\mathbf{D}_1| = \alpha_0 EZ_1 < \infty$. Hence, by Theorem 3.5, the process $(\mathbf{Y}_{\Gamma_n})_{n\in\mathbb{N}_0}$ is strongly mixing with geometric rate. Denote the mixing rate function of $(\mathbf{Y}_{\Gamma_n})$ as

$$\alpha_k = \sup_{A\in\sigma(\mathbf{Y}_0),\, B\in\sigma(\mathbf{Y}_{\Gamma_k})} |P(A\cap B) - P(A)P(B)| \leq Ca^k. \tag{3.11}$$

By (2.14) and using the fact that T is independent of $(T_i, Z_i)_{i\in\mathbb{N}_0}$ and has the distribution of T_1, we have, for any $t \geq 0$,

$$\sigma(\mathbf{Y}_t) = \sigma(\mathbf{Y}_{\Gamma_{N(t)}}, T) \subset \sigma(\mathbf{Y}_{\Gamma_{N(t)}}, T, Z) = \sigma(\mathbf{Y}_{\Gamma_{N(t)+1}}).$$

Hence, $(\mathbf{Y}_t)_{t\geq 0}$ is also strongly mixing with the mixing rate

$$\begin{aligned}\alpha_t' &= \sup_{A\in\sigma(\mathbf{Y}_0),\, B\in\sigma(\mathbf{Y}_t)} |P(A\cap B) - P(A)P(B)| \\ &< \sup_{A\in\sigma(\mathbf{Y}_0),\, B\in\sigma(\mathbf{Y}_{\Gamma_{N(t)+1}})} |P(A\cap B) - P(A)P(B)| \\ &\leq C' a^{ct},\end{aligned}$$

where the last inequality follows from $\frac{N(t)}{t} \xrightarrow{a.s.} c$, as $t\to\infty$. □

Remark 3.7. If the state vector $(\mathbf{Y}_t)_{t\geq 0}$ is strongly mixing with geometric rate then the volatility process $(V_t)_{t\geq 0}$ is also strongly mixing with geometric rate since strong mixing is preserved under linear transformation as well as the rate. The squared increment process $(G_{rn}^{(r)})^2_{n\in\mathbb{N}}$ also inherits the strong mixing property as well as the rate from the volatility process $(V_t)_{t\geq 0}$. (see Haug, et. al. (2007)) □

Remark 3.8. By Theorem 3.6 and Remark 3.7, the squared increment process $((G_t^{(r)})^2)_{n\in\mathbb{N}}$ is strictly stationary and strongly mixing with geometric rate. If, for example, the jump size distribution is standard normal then the driving compound Poisson process has finite moments of all orders, in particular, it follows that $E(G_t^{(r)})^{8+4\nu} < \infty$. It is straigh-forward to see the condition $\sum_{k=0}^{\infty} \alpha_k^{\nu/(2+\nu)} < \infty$ is satisfied for geometric mixing rate. Hence all the conditions of the Theorem 3.3 are satisfied for the process $((G_t^{(r)})^2)_{n\in\mathbb{N}}$. Given the realization of the squared increment process, we fit an ARMA(q,q) process and get the LSE of the COGARCH parameters that are strongly consistent and asymptotically normal (see Corollary 3.11).

Remark 3.9. The maximum likelihood estimators of $\boldsymbol{\beta}_0$ can also be found by maximizing the Gaussian likelihood

$$L(\boldsymbol{\beta},\sigma^2) = \frac{1}{\sqrt{(2\pi)^n v_0\cdots v_{n-1}}} \exp\Big\{-\frac{1}{2}\sum_{j=1}^{n}(X_j-\hat{X}_j)^2/v_{j-1}\Big\}, \tag{3.12}$$

where $\hat{X}_j$, $j=1,\ldots,n$ are one step predictors and v_{j-1}, $j=1,\ldots,n$ are their corresponding mean squared errors both of which can be calculated recursively from the innovations algorithm (see Brockwell and Davis (1991)). Since the linear innovations (3.8) are equal to $\varepsilon_t = X_t - \hat{X}_t$, $t > 0$ and it can be shown that if $\{X_t\}$ is invertible then $v_n \to \sigma^2$ as $n \to \infty$, hence maximizing the likelihood (3.12) is essentially the same as minimizing the least-squares sum (3.10), i.e. the maximum likelihood estimators are asymptotically equivalent to the LSE. $\square$

The next step is to estimate $\mathbf{a}$ by matching the coefficients $c_1, \dots, c_q$ in the (3.1), the ACF of an ARMA(q,q), with the corresponding coefficients of the ACF of the squared increment process and solving the resulting q second order equations system in terms of the parameter $\mathbf{a}$.

In the COGARCH(2,2) case, the ACF (2.68) is reduced to

$$\begin{aligned}\rho(h) &= e^{\tilde{\lambda}_1 h}\frac{(\mu\alpha_1)^2\tilde{\lambda}_1^{-3} - (\mu\alpha_2)^2\tilde{\lambda}_1^{-1} + 2\mu\alpha_1(\tilde{\lambda}_1 + \tilde{\lambda}_2)\tilde{\lambda}_1^{-2} + 2\mu\alpha_2(\tilde{\lambda}_1 + \tilde{\lambda}_2)\tilde{\lambda}_1^{-1}}{2\rho^{-1}e^{\tilde{\lambda}_1}(e^{\tilde{\lambda}_1} - 1)^{-2}(\tilde{\lambda}_1^2 - \tilde{\lambda}_2^2)(\mathbf{a}'P_0\mathbf{a} + \mathbf{a}'Q_0 + R_0)} \\ &\quad - e^{\tilde{\lambda}_2 h}\frac{(\mu\alpha_1)^2\tilde{\lambda}_2^{-3} - (\mu\alpha_2)^2\tilde{\lambda}_2^{-1} + 2\mu\alpha_1(\tilde{\lambda}_1 + \tilde{\lambda}_2)\tilde{\lambda}_2^{-2} + 2\mu\alpha_2(\tilde{\lambda}_1 + \tilde{\lambda}_2)\tilde{\lambda}_2^{-1}}{2\rho^{-1}e^{\tilde{\lambda}_2}(e^{\tilde{\lambda}_2} - 1)^{-2}(\tilde{\lambda}_1^2 - \tilde{\lambda}_2^2)(\mathbf{a}'P_0\mathbf{a} + \mathbf{a}'Q_0 + R_0)}\end{aligned}$$

The ACVF of an ARMA(2,2) process is calculated as follows:

$$\begin{aligned}\frac{\gamma(h)}{\sigma^2} &= \theta_2\xi_1\xi_2 I_{\{0\}}(h) + \xi_1^{-h-1}\frac{(1+\theta_1\xi_1+\theta_2\xi_1^2)(1+\theta_1\xi_1^{-1}+\theta_2\xi_1^{-2})}{(\xi_1^{-1}-\xi_2^{-1})(1-\xi_1^{-2})(1-\xi_1^{-1}\xi_2^{-1})} \\ &\quad - \xi_2^{-h-1}\frac{(1+\theta_1\xi_2+\theta_2\xi_2^2)(1+\theta_1\xi_2^{-1}+\theta_2\xi_2^{-2})}{(\xi_1^{-1}-\xi_2^{-1})(1-\xi_2^{-2})(1-\xi_1^{-1}\xi_2^{-1})}, \quad h = 1, 2, \dots\end{aligned}$$

Matching the ACF of the squared increment process with

$$\rho(h) = c_1\xi_1^{-h} + c_2\xi_2^{-h}, \quad h = 1, 2, \cdots,$$

the ACF of the ARMA(2,2) process and solving the resulting second degree equation system in $\mathbf{a} = (\alpha_1, \alpha_2)$ gives the parameter.

In the COGARCH(1,2) case, the following equation is obtained:

$$(\mu\alpha_1)^2\Big(\frac{c_2\tilde{\lambda}_1^{-3}}{\xi_1(1-\xi_1^{-1})^2}+\frac{c_1\tilde{\lambda}_2^{-3}}{\xi_2(1-\xi_1^{-1})^2}\Big)+2\mu\alpha_1\Big(\tilde{\lambda}_1+\tilde{\lambda}_2\Big)\Big(\frac{c_2\tilde{\lambda}_1^{-2}}{\xi_1(1-\xi_1^{-1})^2}+\frac{c_1\tilde{\lambda}_2^{-2}}{\xi_2(1-\xi_2^{-1})^2}\Big) = 0.$$

From the ACVF of an ARMA(2,2), we can write

$$\frac{c_1}{c_2} = -\frac{(1+\theta_1\xi_1+\theta_2\xi_1^2)(1+\theta_1\xi_1^{-1}+\theta_2\xi_1^{-2})\xi_2(1-\xi_2^{-2})}{(1+\theta_1\xi_2+\theta_2\xi_2^2)(1+\theta_1\xi_2^{-1}+\theta_2\xi_2^{-2})\xi_1(1-\xi_1^{-2})},$$

hence the solution for α_1 can be written as follows:

$$\begin{aligned}\alpha_1 &= -\frac{2\log(\xi_1\xi_2)}{\mu}\Big(\frac{(1+\theta_1\xi_2+\theta_2\xi_2^2)(1+\theta_1\xi_2^{-1}+\theta_2\xi_2^{-2})}{(\xi_1-1)^{-2}(1-\xi_1^{-2})^{-1}}(\log\xi_1)^{-2} \\ &\quad - \frac{(1+\theta_1\xi_1+\theta_2\xi_1^2)(1+\theta_1\xi_1^{-1}+\theta_2\xi_1^{-2})}{(\xi_2-1)^{-2}(1-\xi_2^{-2})^{-1}}(\log\xi_2)^{-2}\Big) \\ &\quad /\Big(\frac{(1+\theta_1\xi_2+\theta_2\xi_2^2)(1+\theta_1\xi_2^{-1}+\theta_2\xi_2^{-2})}{(\xi_1-1)^{-2}(1-\xi_1^{-2})^{-1}}(\log\xi_1)^{-3} \\ &\quad - \frac{(1+\theta_1\xi_1+\theta_2\xi_1^2)(1+\theta_1\xi_1^{-1}+\theta_2\xi_1^{-2})}{(\xi_2-1)^{-2}(1-\xi_2^{-2})^{-1}}(\log\xi_1)^{-3}\Big).\end{aligned} \tag{3.13}$$

where ξ_1, ξ_2 are the autoregressive roots of the ARMA(2,2) process.

Example 3.10. In this example, we show how the COGARCH coefficients can be calculated from the ACF of the squared increment process. For the COGARCH process with coefficients $\alpha_0 = 1$, $\alpha_1 = 0.1$, $\alpha_2 = 0$, $\beta_1 = 1$, $\beta_2 = 0.2$ and compound Poisson driving process with standard normal jumps and $\mu = EL_1^2 = 1$, the mean of the squared increment process is calculated as $M := E(G_n^{(1)})^2 = 2$ and the ACF of the squared increment process is calculated as

$$\rho(h) = 0.1040e^{-0.1127h} - 0.0811e^{-0.8873h}, \quad h = 1, 2, \ldots \tag{3.14}$$

which is easily shown to be the ACF of an ARMA(2,2) process with parameters

$$\phi_1 = 1.3052, \quad \phi_2 = -0.3679, \quad \theta_1 = -1.2642 \quad \text{and} \quad \theta_2 = 0.3669.$$

Figure 5 shows the ACF of the squared increment process at lags $1, 2, \ldots, 50$.

The autoregressive polynomial

$$\phi(\xi) = 1 - 1.3052\xi^{-1} + 0.3679\xi^{-2},$$

yields the zeros,

$$\xi_1^{-1} = 0.8934 \quad \text{and} \quad \xi_2^{-1} = 0.4118,$$

thus giving, by Theorem 2.19,

$$\tilde\beta_1 = \log\xi_1\xi_2 = 1 \quad \text{and} \quad \tilde\beta_2 = \log\xi_1\log\xi_2 = 0.1.$$

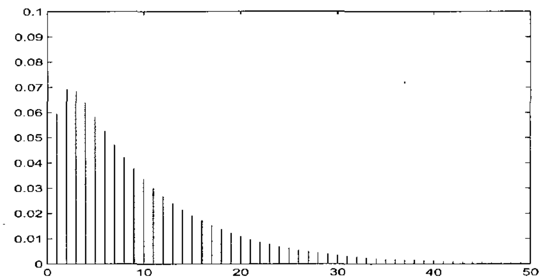

Figure 5: ACF of the squared increment process of the COGARCH(1,2) with parameters $\alpha_0 = 1$, $\alpha_1 = 0.1$, $\beta_1 = 1$, $\beta_2 = 0.2$ and the compound Poisson driving process with standard normal jumps.

Now we can calculate the matrices:

$$P_0 = \begin{pmatrix} 14.6900 & 1.1836 \\ -1.1836 & 0.2854 \end{pmatrix}, \quad Q_0 = \begin{pmatrix} 2.3673 \\ 6.5707 \end{pmatrix}, \quad R_0 = 5.$$

We don't know if we have COGARCH$(1,2)$ or COGARCH$(2,2)$ yet, so we'll find both α_1 and α_2 allowing for the possibility that α_2 is not zero. Writing down the ACF of the squared increment process in terms of the parameter $\mathbf{a}$ as shown below,

$$\begin{aligned}\rho(h) &= \frac{17.2007\alpha_1^2 - 0.2185\alpha_2^2 + 3.8771\alpha_1 - 0.4370\alpha_2}{14.6900\alpha_1^2 + 0.2853\alpha_2^2 + 2.3673\alpha_1 + 6.5707\alpha_2 + 5} e^{-0.1127h} \\ &\quad + \frac{-2.3295\alpha_1^2 + 1.8340\alpha_2^2 - 4.1338\alpha_1 + 3.6680\alpha_2}{14.6900\alpha_1^2 + 0.2853\alpha_2^2 + 2.3673\alpha_1 + 6.5707\alpha_2 + 5} e^{-0.8873h},\end{aligned}$$

and matching the two autocorrelation functions yields the equations

$$-150.7545\alpha_1^2 + 2.3868\alpha_2^2 - 34.9245\alpha_1 + 10.7736\alpha_2 + 5 = 0,$$

and

$$-14.0288\alpha_1^2 + 22.8957\alpha_2^2 - 48.5971\alpha_1 + 51.7914\alpha_2 + 5 = 0,$$

which give the solutions $(0.1, 0)$ and $(-0.2611, -0.3905)$. The latter solution does not satisfy the condition $\alpha_2 \geq 0$ given by Theorem 2.5, so we recover $\alpha_1 = 0.1$ and $\alpha_2 = 0$, the parameters of a COGARCH$(1,2)$ model. If we had assumed that the process was COGARCH$(1,2)$ we could have found $\alpha_1 = 0.1$ directly from (3.13). Further, we find from (2.27) with $\mu = 1$,

$$\beta_1 = \tilde{\beta}_1 = 1 \quad \text{and} \quad \beta_2 = \tilde{\beta}_2 + \alpha_1 = 0.2,$$

and, finally find $\alpha_0 = 1$ from (2.60) with $M = 2$ and $\mu = 1$. $\square$

We now summarize the least-squares estimation procedure proposed in this section by the following corollary which applies in particular to the COGARCH(2,2) process with compound Poisson driving process. In order to avoid overparameterization, we also assume that the driving process satisfies $\mu = EL_1^2 = 1$. For this process define the mapping $Q : \mathbb{R}^5 \to \mathbb{R}^5$ by $(\boldsymbol{\beta}, M) \to Q(\boldsymbol{\beta}, M) = \boldsymbol{\theta} := (\alpha_0, \alpha_1, \alpha_2, \beta_1, \beta_2)$, where

$$\alpha_0 = \frac{M \log \xi_1 \xi_2}{\alpha_1 + \log \xi_1 \log \xi_2}, \tag{3.15}$$

$$\beta_1 = \alpha_2 + \log \xi_1 \log \xi_2, \tag{3.16}$$

$$\beta_2 = \alpha_1 + \log \xi_1 \log \xi_2, \tag{3.17}$$

and ξ_1, ξ_2 are the autoregressive roots of the ARMA(2,2) process, $M := E((G_t^{(r)})^2)$, and $\mathbf{a} = (\alpha_1, \alpha_2)$ is found by solving the resulting equations when matching the ACF of the squared increment process with the ACF of the ARMA(2,2) process, subject to the conditions (2.29), $\alpha_2 \geq 0$ and $\alpha_1 \geq -\alpha_2 \lambda(B)$. In the COGARCH(1,2) case, α_1 is given by (3.13).

Denote the estimators of the COGARCH parameters as $\hat{\boldsymbol{\theta}}_n = Q(\hat{\boldsymbol{\beta}}_n, \hat{M})$ and $\boldsymbol{\theta}_0 = Q(\boldsymbol{\beta}_0, M)$, where $\hat{M}$ is the sample mean of the realizations of the squared increment process. Using the fact that the mapping Q is continuous in $(\boldsymbol{\beta}, M)$ and differentiable at $(\boldsymbol{\beta}_0, M)$ and given the strong consistency and the asymptotic normality of the estimators $\hat{\boldsymbol{\beta}}_n$, we get the strong consistency and the asymptotic normality of the estimators $\hat{\boldsymbol{\theta}}_n$ by applying the delta method, which we summarize in the following.

Corollary 3.11. *Suppose that the conditions of Theorem 3.6 hold, and the driving Lévy process satisfies $EL_1^2 = 1$. Then as $n \to \infty$,*

$$\hat{\boldsymbol{\theta}}_n \xrightarrow{a.s.} \boldsymbol{\theta}_0. \tag{3.18}$$

If, in addition, $(L_t)_{t\geq 0}$ is a Lévy process such that $EL_1^{8+\nu} < \infty$ for some positive constant ν, then as $n \to \infty$,

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \xrightarrow{d} \partial_{(\boldsymbol{\beta},M)}Q(\boldsymbol{\beta}_0, M)N(0,\Sigma), \tag{3.19}$$

where Σ is as in Theorem 3.3.

3.4 The estimation algorithm

In this section, we shall estimate the parameters $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \alpha_2, \beta_1, \beta_2)$ of a COGARCH(2,2) process, based on the results of the Corollary 3.11. We have the data G_i, $i = 0, \ldots, n$, observed at equally spaced time intervals. Let $r = 1$ for simplicity, giving the returns

$$G_i^{(1)} = G_{i+1} - G_i, \quad i = 0, \ldots, n-1.$$

We assume that the conditions in the Corollary 3.11 are satisfied.

Algorithm 3.12. *1. Calculate the sample moment*

$$\hat{M} := \frac{1}{n-1}\sum_{i=0}^{n-1}(G_i^{(1)})^2$$

and the sample autocovariances

$$\hat{\gamma}(h) := \frac{1}{n}\sum_{i=0}^{n-i-1}((G_{i+h}^{(1)})^2 - \hat{M})((G_i^{(1)})^2 - \hat{M}), \quad h = 0, \ldots, 4.$$

Then calculate the sample autocorrelations $\hat{\rho}(h) := \hat{\gamma}(h)/\hat{\gamma}(0)$, $\quad h = 1, \ldots, 4$.

2. Solving the equations

$$\begin{pmatrix} \hat{\rho}(2) & \hat{\rho}(1) \\ \hat{\rho}(3) & \hat{\rho}(2) \end{pmatrix} \begin{pmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \end{pmatrix} = \begin{pmatrix} \hat{\rho}(3) \\ \hat{\rho}(4) \end{pmatrix}$$

yields the preliminary estimators $\hat{\phi}_1$, $\hat{\phi}_2$.

3. Find the LSE estimator $\hat{\boldsymbol{\beta}}$ of ARMA(2,2) by minimizing the least-squares sum (3.10) *using the preliminary estimators $\hat{\phi}_1$, $\hat{\phi}_2$ as initial values.*

4. Finding the zeros of the polynomial $\hat{\phi}(\xi) = 1 - \hat{\phi}_1\xi^{-1} - \hat{\phi}_2\xi^{-2}$ yields $\hat{\xi}_1$ and $\hat{\xi}_2$, *giving,* $\hat{\hat{\beta}}_1 = \log \hat{\xi}_1 \hat{\xi}_2$ *and* $\hat{\hat{\beta}}_2 = \log \hat{\xi}_1 \log \hat{\xi}_2$.

5. Match the ACF of the ARMA(2,2) with the ACF of the squared increments and solve the resulting equations to get $\hat{\mathbf{a}}$, *subject to the conditions* (2.29), $\hat{\alpha}_2 \geq 0$ *and* $\hat{\alpha}_1 \geq \hat{\alpha}_2 \log \hat{\xi}_1$. *In the COGARCH(1,2) case, find* $\hat{\alpha}_1$ *from*

$$\begin{aligned} \hat{\alpha}_1 &= -2\log(\hat{\xi}_1\hat{\xi}_2)\Big(\frac{(1+\hat{\theta}_1\hat{\xi}_2+\hat{\theta}_2\hat{\xi}_2^2)(1+\hat{\theta}_1\hat{\xi}_2^{-1}+\hat{\theta}_2\hat{\xi}_2^{-2})}{(\hat{\xi}_1-1)^{-2}(1-\hat{\xi}_1^{-2})^{-1}}(\log\hat{\xi}_1)^{-2} \\ &\quad - \frac{(1+\hat{\theta}_1\hat{\xi}_1+\hat{\theta}_2\hat{\xi}_1^2)(1+\hat{\theta}_1\hat{\xi}_1^{-1}+\hat{\theta}_2\hat{\xi}_1^{-2})}{(\hat{\xi}_2-1)^{-2}(1-\hat{\xi}_2^{-2})^{-1}}(\log\hat{\xi}_2)^{-2}\Big) \\ &\quad /\Big(\frac{(1+\hat{\theta}_1\hat{\xi}_2+\hat{\theta}_2\hat{\xi}_2^2)(1+\hat{\theta}_1\hat{\xi}_2^{-1}+\hat{\theta}_2\hat{\xi}_2^{-2})}{(\hat{\xi}_1-1)^{-2}(1-\hat{\xi}_1^{-2})^{-1}}(\log\hat{\xi}_1)^{-3} \\ &\quad - \frac{(1+\hat{\theta}_1\hat{\xi}_1+\hat{\theta}_2\hat{\xi}_1^2)(1+\hat{\theta}_1\hat{\xi}_1^{-1}+\hat{\theta}_2\hat{\xi}_1^{-2})}{(\hat{\xi}_2-1)^{-2}(1-\hat{\xi}_2^{-2})^{-1}}(\log\hat{\xi}_2)^{-3}\Big). \end{aligned}$$

Then get the estimators $\hat{\beta}_1 = \hat{\hat{\beta}}_1 + \hat{\alpha}_2$ *(*$\hat{\beta}_1 = \hat{\hat{\beta}}_1$ *in COGARCH(1,2) case) and* $\hat{\beta}_2 = \hat{\hat{\beta}}_2 + \hat{\alpha}_1$.

6. Finally, find the estimator of α_0 *from* $\hat{\alpha}_0 = \hat{M}\hat{\hat{\beta}}_2/\hat{\beta}_2$.

3.5 A simulation study

We now illustrate the estimation procedure by estimating the parameters of a simulated COGARCH$(1,2)$ process where the driving Lévy process is a compound Poisson process. The compound Poisson process is given by

$$L_t = \sum_{i=1}^{N_t} Y_i, \quad t \geq 0,$$

where $N = (N_t)_{t\geq 0}$ is a Poisson process with intensity $c > 0$, and $(Y_i)_{i\in\mathbb{N}}$ are i.i.d. random variables with distribution function F_Y, independent of N. For this driving Lévy process, $\tau_L^2 = 0$ and the Lévy measure of L has the representation $\nu(dy) = cF_Y(dy)$. Let the jump size distribution be a normal with mean 0 and variance σ^2. Then the driving Lévy process has finite moments of all order, in particular,

$$\mu = EL_1^2 = c\sigma^2 \quad \text{and} \quad \rho = EL_1^4 = 3c\sigma^4 < \infty.$$

Let $\alpha_0 = 1$, $\alpha_1 = 0.1$, $\alpha_2 = 0$, $\beta_1 = 1$, $\beta_2 = 0.2$, $c = 1$ and $\sigma^2 = 1$, as in Example 3.10. With $\mathbf{Y}_0 = (10,\ 1)'$ as a starting value, $n = 1,000,000$ equidistant realizations of the log returns $G_i^{(1)}$, $i = 0, \ldots, n-1$ were simulated and Algorithm 3.12 was used to estimate the COGARCH(1,2) coefficients. This procedure was repeated 2000 times. We show the steps of the estimation algorithm in one of the simulations.

Step 1: The sample mean and the sample ACF of the simulated squared returns were calculated:

$$\hat{M} = 2.0116,$$

$$\hat{\rho}(1) = 0.0598, \quad \hat{\rho}(2) = 0.0696, \quad \hat{\rho}(3) = 0.0688 \quad \text{and} \quad \hat{\rho}(4) = 0.0641.$$

Step 2: The equation,

$$\begin{pmatrix} 0.0696 & 0.0598 \\ 0.0688 & 0.0696 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} 0.0688 \\ 0.0641 \end{pmatrix},$$

gives the preliminary estimators $\hat{\phi}_1 = 1.3088$ and $\hat{\phi}_2 = -0.3728$. The histogram of the preliminary estimators of $\boldsymbol{\phi}$ in all 2000 simulations are shown in Figure 6. As discussed in Section 3.1, we see that the histograms are strongly skewed.

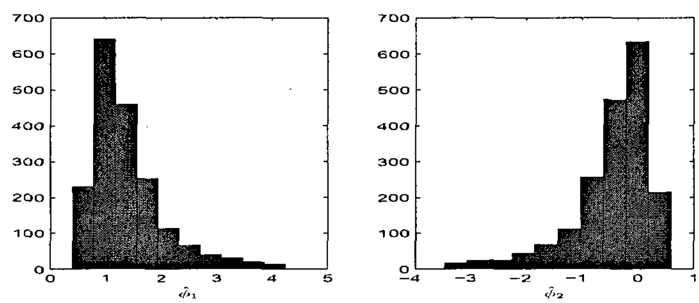

Figure 6: Histogram of the preliminary estimators of ϕ_1 (left) and ϕ_2 (right). The true values are $\phi_1 = 1.3052$ and $\phi_2 = -0.3679$.

Step 3: Using the preliminary estimators as the initial value, the least-square estimators for $\boldsymbol{\beta}$ for the ARMA(2,2) model are calculated as,

$$\hat{\phi}_1 = 1.3047, \quad \hat{\phi}_2 = -0.3768, \quad \hat{\theta}_1 = -1.2681 \quad \text{and} \quad \hat{\theta}_2 = 0.3638.$$

The histograms of the estimated ARMA(2,2) coefficients are shown in Figure 7. The histograms appear to have the shape of normal distributions, confirming that the least-squares estimators are asymptotically normal.

Step 4: Finding the zeros of

$$\phi(\xi) = 1 - 1.3047\xi^{-1} + 0.3768\xi^{-2},$$

gives,

$$\hat{\xi}_1^{-1} = 0.8732 \quad \text{and} \quad \hat{\xi}_2^{-1} = 0.4315,$$

thus giving,

$$\hat{\hat{\beta}}_1 = \log \hat{\xi}_1 \hat{\xi}_2 = 0.9760 \quad \text{and} \quad \hat{\hat{\beta}}_2 = \log \hat{\xi}_1 \log \hat{\xi}_2 = 0.1140.$$

Step 5: Calculate $\hat{\alpha}_1 = 0.0958$. Then calculate

$$\hat{\beta}_1 = \hat{\hat{\beta}}_1 = 0.9760 \quad \text{and} \quad \hat{\beta}_2 = \hat{\hat{\beta}}_2 + \hat{\alpha}_1 = 0.2098.$$

Step 6: Finally, using the sample mean from Step 1 to get

$$\hat{\alpha}_0 = \hat{M}\hat{\hat{\beta}}_2/\hat{\beta}_2 = 1.0930.$$

Table 3.1 summarizes the simulation results after the Algorithm 3.12 is repeated 2000 times. For each parameter, the empirical mean, bias, mean square error (MSE) and mean absolute error (MAE) with corresponding standard errors (in brackets) are calculated. The histogram of the estimated parameters are shown in Figure 8. All histograms show no striking deviation from normal distributions. The estimator for α_0 appears a little skewed and has the largest standard error among all four estimators because of the tail heaviness of the squared increment process.

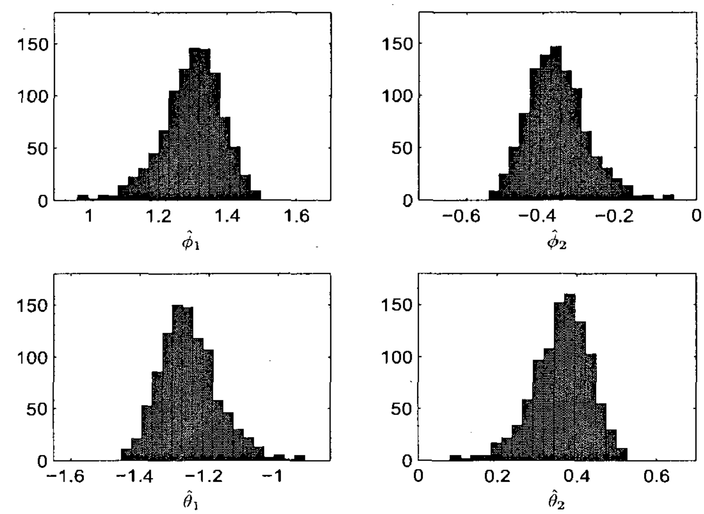

Figure 7: Least-squares estimators of the ARMA parameters. The true values are $\phi_1 = 1.3052$, $\phi_2 = -0.3679$, $\theta_1 = -1.2642$ and $\theta_2 = 0.3669$.

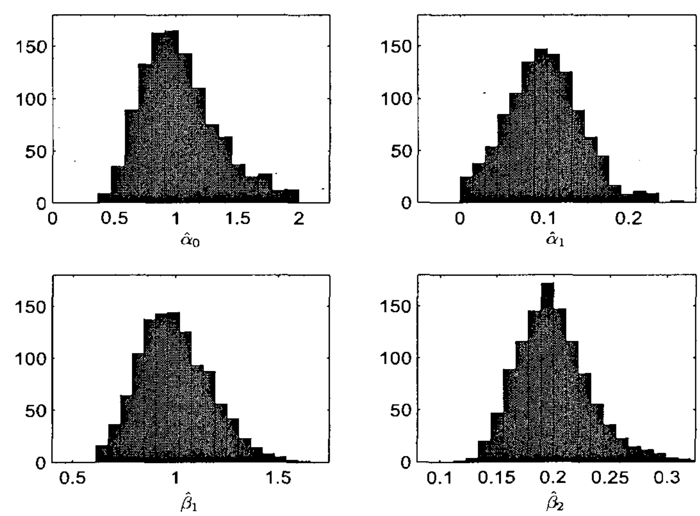

Figure 8: Least-squares estimators of the COGARCH parameters. The true values are $\alpha_0 = 1$, $\alpha_1 = 0.1$, $\beta_1 = 1$ and $\beta_2 = 0.2$.

	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\beta}_2$
Mean	1.0683(0.0100)	0.0961(0.0010)	1.0123(0.0054)	0.1986(0.0009)
Bias	0.0683(0.0100)	-0.0039(0.0013)	0.0123(0.0054)	-0.0014(0.0009)
MSE	0.1099(0.0059)	0.0018(0.00001)	0.0300(0.0014)	0.0009(0.0001)
MAE	0.2509(0.0069)	0.0337(0.0008)	0.1383(0.0033)	0.0242(0.0006)

Table 3.1: Estimated mean, bias, MSE and MAE of the LSE of the COGARCH parameters with corresponding standard deviations (in brackets). The true values are $\alpha_0 = 1$, $\alpha_1 = 0.1$, $\beta_1 = 1$ and $\beta_2 = 0.2$.

3.6 Estimating the volatility

In this section, we show that the state process $\{\mathbf{Y}_t\}$, therefore the volatility process $\{V_t\}$ can be approximated using the observed log returns, the estimated coefficients of the fitted COGARCH(2,2) model and an initial starting value $\mathbf{Y}_0$. The process $\{V_t\}$ is usually called *instantaneous volatility* or *spot volatility*. Since the estimated coefficients of the COGARCH model are used in the approximation, we shall assume that the conditions of the Corollary 3.11 are satisfied so that the estimated coefficients are strongly consistent and asymptotically normal. We start with a fixed $\mathbf{Y}_0$ equal to the mean of the state vector of the fitted model.

From (2.6), we can write that for a small time interval $\frac{1}{h}$, where h is a positive integer,

$$\begin{aligned}\mathbf{Y}_{t+\frac{1}{h}} - \mathbf{Y}_t &= \int_t^{t+\frac{1}{h}} d\mathbf{Y}_s = \int_t^{t+\frac{1}{h}} (B\mathbf{Y}_{s-}ds + \mathbf{e}V_s d[L,L]_s^{(d)}) \\ &= B\int_t^{t+\frac{1}{h}} \mathbf{Y}_{s-}ds + \mathbf{e}\int_t^{t+\frac{1}{h}} V_s d[L,L]_s^{(d)}.\end{aligned}$$

Writing the Euler approximation,

$$\int_t^{t+\frac{1}{h}} \mathbf{Y}_{s-}ds \approx \frac{1}{h}\mathbf{Y}_t$$

and

$$\int_t^{t+\frac{1}{h}} V_s d[L,L]_s^{(d)} = \sum_{t<s\le t+\frac{1}{h}} V_s(\Delta L_s)^2 \approx (G_{t+\frac{1}{h}} - G_t)^2 = \left(G_t^{(\frac{1}{h})}\right)^2$$

we obtain

$$\mathbf{Y}_{t+\frac{1}{h}} \approx (I + \frac{1}{h}B)\mathbf{Y}_t + \mathbf{e}\big(G_t^{(\frac{1}{h})}\big)^2.$$

Using the recursion for $\mathbf{Y}_{t+1}, \dots, \mathbf{Y}_{t+\frac{2}{h}}$, we get

$$\begin{aligned}\mathbf{Y}_{t+1} &\approx (I + \frac{1}{h}B)^h\mathbf{Y}_t + \sum_{i=1}^{h}(I + \frac{1}{h}B)^{h-i}\mathbf{e}\big(G_{t+\frac{i-1}{h}}^{(\frac{1}{h})}\big)^2 \\ &\approx e^B\mathbf{Y}_t + \sum_{i=1}^{h}(I + \frac{1}{h}B)^{h-i}\mathbf{e}\big(G_{t+\frac{i-1}{h}}^{(\frac{1}{h})}\big)^2.\end{aligned}$$

Hence the state process and volatility process at integer times can be estimated recursively from high frequency observations of G by

$$\hat{\mathbf{Y}}_{t+1} = e^{\hat{B}}\hat{\mathbf{Y}}_t + \sum_{i=1}^{h}(I + \frac{1}{h}\hat{B})^{h-i}\mathbf{e}\big(G_{t+\frac{i-1}{h}}^{(\frac{1}{h})}\big)^2 \tag{3.20}$$

and

$$\hat{V}_{t+1} = \hat{\alpha}_0 + \hat{\mathbf{a}}'\hat{\mathbf{Y}}_{t+1}, \quad t = 0, 1, \dots \tag{3.21}$$

where $\hat{\mathbf{Y}}_0$ is an initial starting value and $\hat{B}$ is the matrix B with $\beta_1, \dots, \beta_q$ replaced by $\hat{\beta}_1, \dots, \hat{\beta}_q$. Given $\hat{\mathbf{Y}}_t$ and the observed the log returns $G_t, G_{t+\frac{1}{h}}, \dots, G_{t+1}$ we can estimate the state vector $\hat{\mathbf{Y}}_{t+1}$ and therefore the volatility V_{t+1}. In the result shown below, these observations are availiable since we simulated the process. To apply this estimation in real data we set a unit time interval first, depending on the frequency of the data collected. For example, if we collected 5-minute log returns and then we can choose the unit time to be one hour so that $h = 12$, or we choose the unit time to be 30 minutes so that $h = 6$, etc. In Section 3.7 we show how the volatility is estimated using 5-minute log returns with the unit time chosen as 30 minutes.

Haug et al. (2007) suggested the volatility approximation in the COGARCH(1, 1) case when $h = 1$. In the case of $h = 1$, (3.20) is reduced to

$$\hat{\mathbf{Y}}_{t+1} = e^{\hat{B}}\hat{\mathbf{Y}}_t + \mathbf{e}(G_t^{(1)})^2. \tag{3.22}$$

To show the accuracy of the volatility estimation, $n = 500$ observations of a COGARCH(1, 2) process with estimated coefficients $\alpha_0 = 1.0716$, $\alpha_1 = 0.0958$, $\beta_1 =$

0.0976, $\beta_2 = 0.2098$, obtained in one of the simulations earlier, and compound Poisson driving process with standard normal jumps with intensity $c = 1$ was simulated, using the Matlab code provided in the Appendix. The true parameters were $\alpha_0 = 1$, $\alpha_1 = 0.1$, $\beta_1 = 1$ and $\beta_2 = 0.2$. We know $EL_1 = 0$ and $EL_1^2 = 1$ for this driving process. $h = 10$ was used and the initial value $\mathbf{Y}_0 = (5.1077, 0)'$ is chosen as the mean of the state vector of the fitted model.

Figure 9 shows the simulated squared increments, the true volatility (shown by a blue line) of the COGARCH(1,2) process with parameters $\alpha_0 = 1$, $\alpha_1 = 0.1$, $\beta_1 = 1$ and $\beta_2 = 0.2$ and the estimated volatilities (shown by red lines) based on the simulated values G_t, $t = 0, 0.1, \ldots, 500$ and the estimated coefficients. Since we know from the simulation the values of $G^{\frac{1}{h}}_{t+\frac{i-1}{h}}$ in (3.20) we can use them to estimate the volatilities at integer times. When $h = 10$, the estimated volatility is essentially the same as the true volatility, except for the first few values. As expected, the approximation suggested by Haug et al. (2007) (i.e. with $h = 1$), also shown in Figure 9, is a less accurate approximation to the true volatility. The Matlab code used to approximate the volatility is given in the Appendix.

The goodness-of-fit of the model can be done by a residual analysis. The estimated residuals are computed by $G^{(1)}_{i-1}/\sqrt{\hat{V}_i}$, $i = 1, \ldots, n$. Since the jump size distribution is symmetric around zero, the residuals should have mean that is close to zero and symmetric around zero. The standard deviation of the residuals should be close to 1, the variance of the driving compound process per unit time. If the model is a good fit to the data the estimated residuals should be independent (so, in particular, they and their squares should be uncorrelated). Ljung-Box and McLeod-Li tests were performed to test the correlation of the residuals and their squares. The Ljung-Box test statistic is given by

$$Q_{LB} = n(n+2)\sum_{i=1}^{k} \frac{\hat{\rho}(i)^2}{n-i},$$

where $\hat{\rho}(i)$ is the empirical autocorrelation function of the residuals, and asymp-

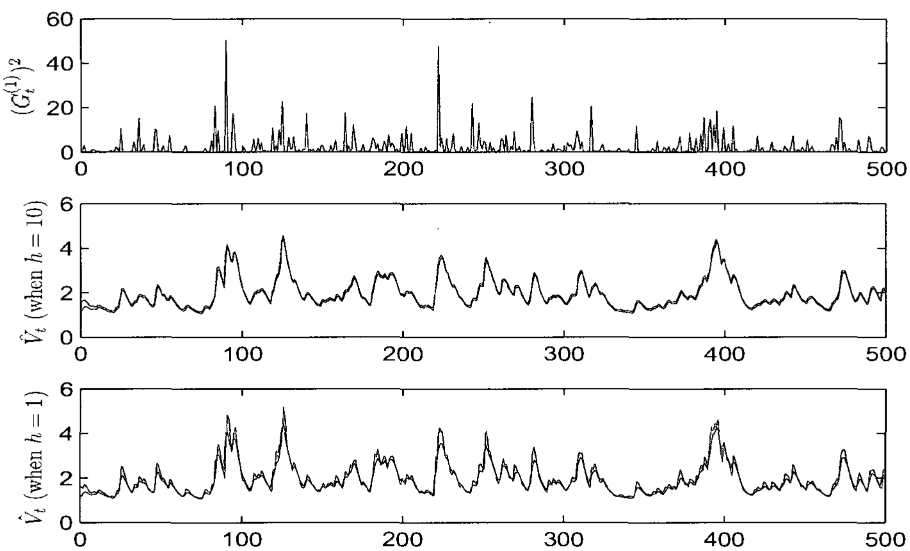

Figure 9: Sample path of the squared increment process (top graph), theoretical (blue line) and the estimated volatility (red line) based on the observations $G_{\frac{1}{n}}, G_{\frac{2}{n}}, \ldots, G_{500}$ and the estimated coefficients. The middle graph is for $h = 10$ and the bottom graph is for $h = 1$. The blue and red lines are virtually indistinguishable when $h = 10$.

totically χ^2-distributed with k degrees of freedom. In the McLeod-Li test statistic, residuals are replaced by their squares. With $k = 50$ lags, the 95th percentile of the chi-square distribution was 56.9424. The Ljung-Box and McLeod-Li statistics were 39.4053 and 39.8425, respectively, providing evidence (as they should) of the appropriateness of the model.

3.7 Real data analysis

We modelled the 30-minute log returns of the Dow Jones Industrial Average recorded from February 12th, 2003 to May 12th, 2006 using the 5-minute returns to estimate the volatility. There was a total of 813 trading days not including the weekends and holidays with 78 5-minute observations per day, resulting in total of $n = 63414$ 5-minute observations. See Figure 10 for the Dow Jones Industrial Average Index for the recorded time. In this example the unit of time is 30 minutes and $h = 6$.

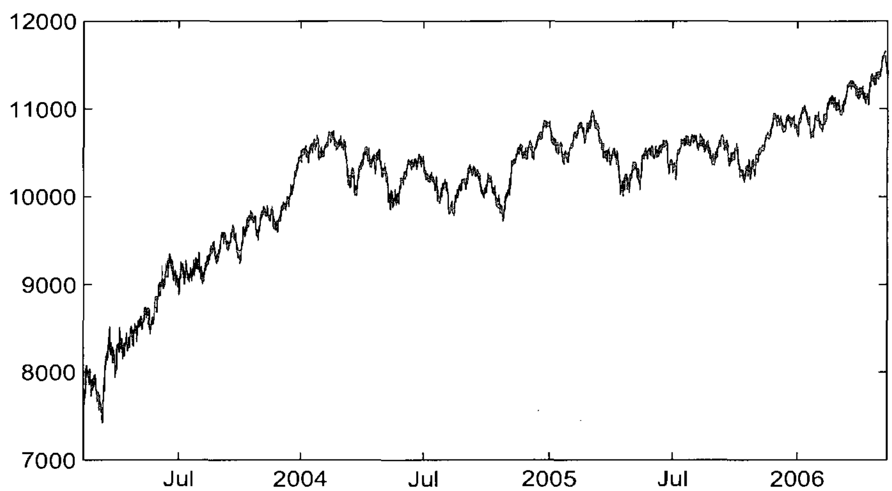

Figure 10: Dow Jones 5-minute data (P_n) from February 12th, 2003 to May 12th, 2006.

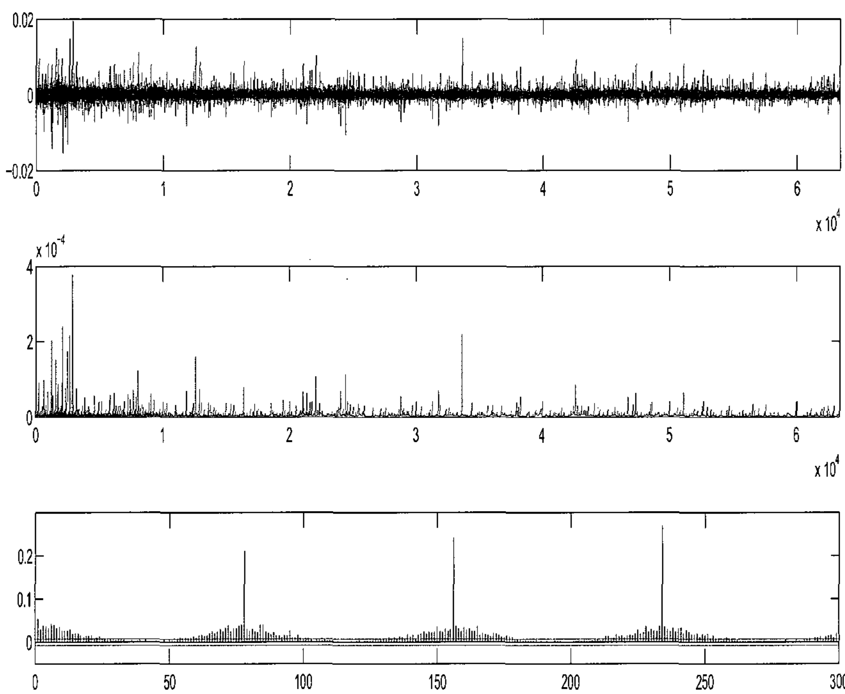

Figure 11: Dow Jones unfiltered log returns (the top graph), their squares (the middle graph) and the ACF of the squared log returns. The ACF clearly contains a seasonal component with period 78.

3.7.1 Filtering the data

The squared increments of the logarithm of the 5-minute returns (known as squared log returns) and their ACF were calculated and are shown in Figure 11. The ACF clearly contains a seasonal component with period 78. This is a daily seasonality because trading is more intense at the beginning and at the end of a day, slows down around noon, and different days do not seem to differ to a high degree. Brodin and Klüppelberg (2006) suggested volatility weighting method to remove the effect of daily seasonality. Their method divides a period (day) into several smaller subperiods (5-minute intervals) and then estimates the seasonality effect in each subperiod in terms of volatility. Then each subperiod is deseasonalized separately. The observed log returns, denoted by $\tilde{x}_n$ as in Brodin and Klüppelberg (2006), is a realization of the process

$$\tilde{x}_n = \mu + v_n x_n, \quad n = 0, \dots, 63413, \tag{3.23}$$

where x_n are the deseasonalized returns, μ is a constant drift and v_n are the seasonality coefficients (volatility weights), estimated by

$$\hat{v}_\tau = \text{median}_{i=1,\dots,N^\tau} |\tilde{x}_{n_i+\tau}|. \tag{3.24}$$

Here we have $N^\tau = 813$ days, during which we have observed our data in the given subperiod $\tau \in \{1, 2, \dots, 78\}$. In Figure 12 we show the estimated seasonality coefficients. The Matlab code for finding the seasonality coefficients and filtering the data can be found in Appendix. The seasonality coefficients also display the fact that trading is more intense at the beginning and at the end of a day.

Finally, the deseasonalized log returns are calculated by

$$x_n = \frac{\tilde{x}_n - \hat{\mu}}{\hat{v}_n}, \quad n = 0, \dots, 63413 \tag{3.25}$$

where $\hat{\mu}$ is the sample mean of the log returns. The squared log returns of the filtered series and its ACF are shown in Figure 13. The filtered data now show no clear seasonal effect.

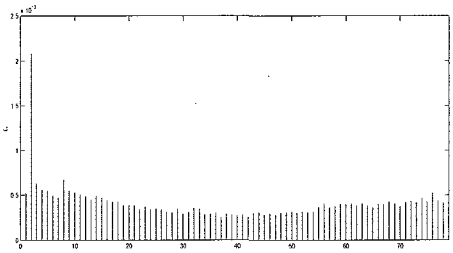

Figure 12: Dow Jones 5-minute data from February 12th, 2003 to May 12th, 2006. Estimated seasonality coefficients (volatility weights) for the 78 subperiods.

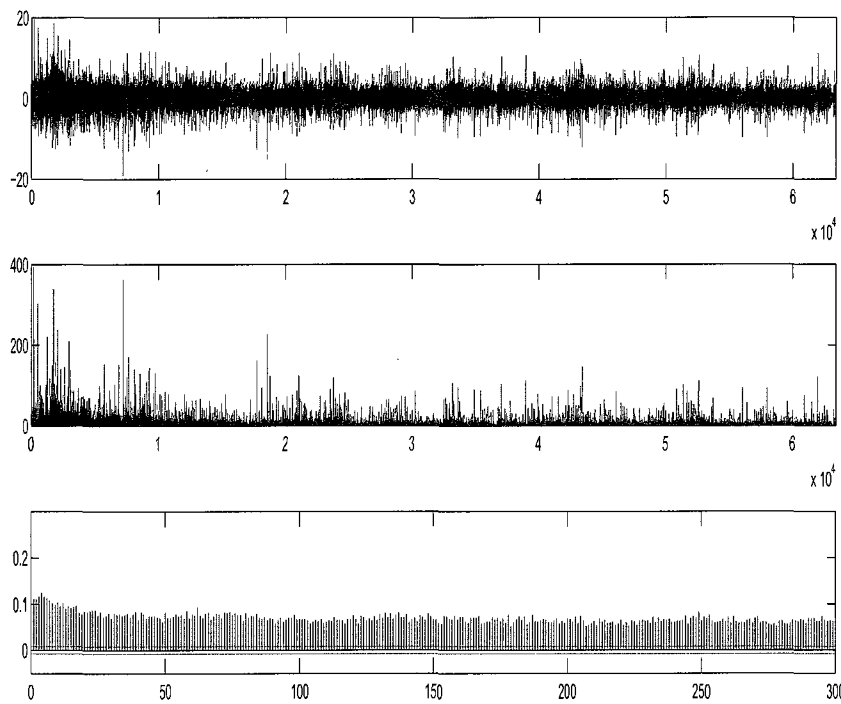

Figure 13: Dow Jones filtered log returns (the top graph), their squares (the middle graph) and the ACF of the squared log returns. The filtered data now show no clear seasonal effect.

3.7.2 Fitting COGARCH$(2,2)$ to squared 30-minute returns

We shall fit a COGARCH$(2,2)$ model to G using the filtered 30-minute log returns $G_0^{(1)}, G_1^{(1)}, \ldots, G_{10568}^{(1)}$ (each of which is the sum of six successive filtered 5-minute log returns) and using the filtered 5-minute returns to estimate the volatilities $V_1, \ldots, V_{10569}$. The unit of time will be 30-minutes with $h = 6$. We assume driving compound Poisson process has jump-rate $c = 2$ and normally distributed jumps with mean zero and variance 0.7071 so that it satisfies $EL_1 = 0$ and $EL_1^2 = 1$.

Minimizing the least-squares sum (3.10) yields the estimated parameters of the ARMA(2,2) model as follows:

$$\hat{\phi}_1 = 1.6548, \quad \hat{\phi}_2 = -0.6552, \quad \hat{\theta}_1 = -1.5961 \quad \text{and} \quad \hat{\theta}_2 = 0.6021.$$

The autoregressive polynomial,

$$\hat{\phi}(\xi) = 1 - 1.6548\xi^{-1} + 0.6552\xi^{-2},$$

yields the zeros,

$$\hat{\xi}_1^{-1} = 0.9986 \quad \text{and} \quad \hat{\xi}_2^{-1} = 0.6561,$$

thus giving the eigenvalues $\tilde{\lambda}_1 = -0.0014$, $\tilde{\lambda}_2 = -0.4214$ and

$$\hat{\hat{\beta}}_1 = \log \hat{\xi}_1 \hat{\xi}_2 = 0.4228, \quad \hat{\hat{\beta}}_2 = \log \hat{\xi}_1 \log \hat{\xi}_2 = 0.0006.$$

The ACF of the ARMA(2,2) process with the above parameters is calculated as,

$$\rho(h) = 0.0985e^{-0.0014h} + 0.0644e^{-0.4214h}, \quad h = 1, 2, \ldots$$

The matrices P_0, Q_0 and R_0 are calculated as follows:

$$P_0 = \begin{pmatrix} 3103.9260 & 1.6013 \\ 1.6013 & 1.0968 \end{pmatrix}, \quad Q_0 = \begin{pmatrix} 1.3539 \\ 3.9274 \end{pmatrix}, \quad R_0 = 3.5.$$

Writing the ACV of the squared increments as a function of of $\mathbf{a}$ and matching it with the ACF of the ARMA(2,2) model yields the equations

$$-2.8505.7894\alpha_1^2 + 1.1550\alpha_2^2 - 34.8905\alpha_1 + 3.9766\alpha_2 + 3.5 = 0,$$

and

$$3261.9098\alpha_1^2 - 26.9585\alpha_2^2 + 57.6452\alpha_1 - 19.7941\alpha_2 + 3.5 = 0,$$

giving the solution $(0.0117, 0.1860)$ for (α_1, α_2). Now we can find

$$\hat{\beta}_1 = \hat{\hat{\beta}}_1 + \hat{\alpha}_2 = 0.6088 \quad \text{and} \quad \hat{\beta}_2 = \hat{\hat{\beta}}_2 + \hat{\alpha}_1 = 0.0122.$$

Finally, using the sample mean $\hat{M} = 3.0458$, we find

$$\hat{\alpha}_0 = \hat{M}\hat{\hat{\beta}}_2/\hat{\beta}_2 = 0.9760.$$

3.7.3 Volatility estimation and goodness-of-fit

Based on the filtered 5-minute log returns and the estimated coefficients, we can now estimate the volatility for 30-minute log returns. The unit of time is 30 minutes and we have $h = 6$ observations per unit interval. The Matlab code (See Appendix) implementing the recursions (3.21) and (3.20) resulted in the estimated volatilities shown in Figure 14. The starting vector $\mathbf{Y}_0$ is chosen to be equal to the mean of the state vector of the fitted model.

We again use the residual analysis to check the goodness-of-fit of the model. The residuals are calculated similarly as before:

$$G_{t-1}^{(1)}/\sqrt{\hat{V}_t}, \quad t = 1, 2, \ldots, 10569.$$

If the model is a good fit to the data then the residuals should be independent, have zero mean and standard deviation one. The sample mean and the sample standard deviation were -0.0232 and 0.9325, respectively, and the Ljung-Box and McLeod-Li test statistics with lags 189 were $Q_{LB} = 222.0025$ and $Q_{ML} = 214.0934$, suggesting a good fit as the the critical value of the chi-square distribution at 0.05 level was 222.0757. The ACF of the filtered squared 30-minute log returns, the residuals and the squared residuals, shown in Figure 15, indicates the model is a very good fit to the data.

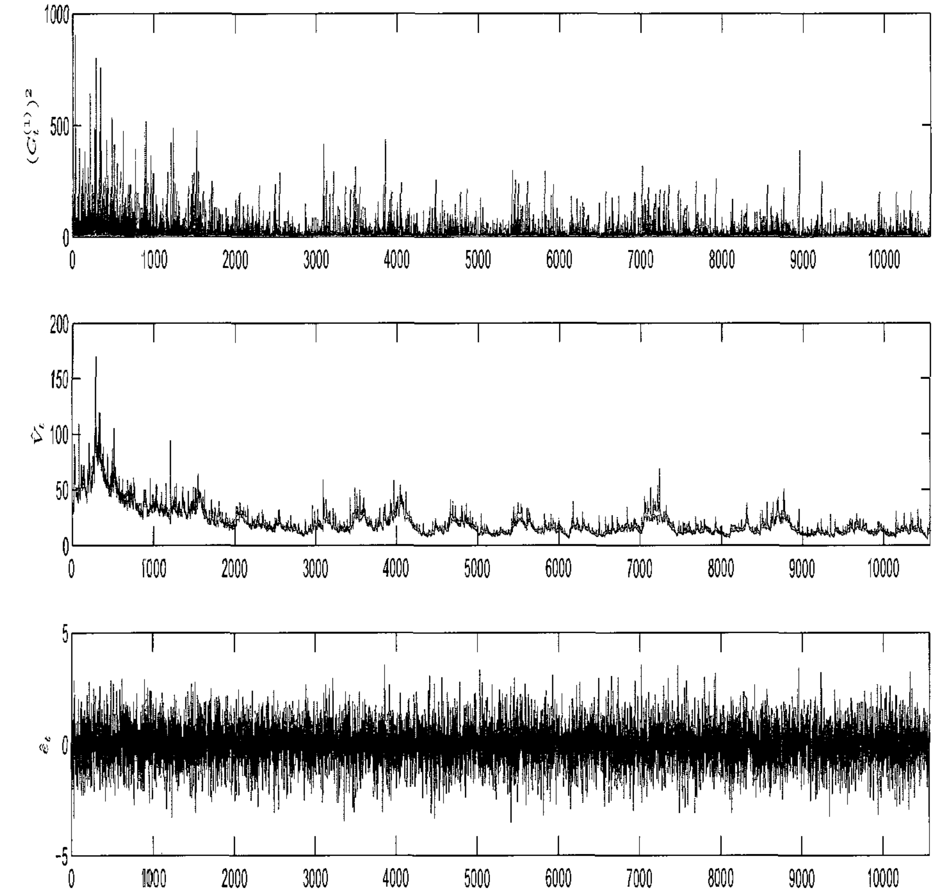

Figure 14: Dow Jones squared filtered 30-minute log returns $(G_t^{(1)})^2$ (top graph), the estimated volatilities $\hat{V}_t$ based on the estimated coefficients and 5-minute log returns (middle graph) and the estimated residuals (bottom graph).

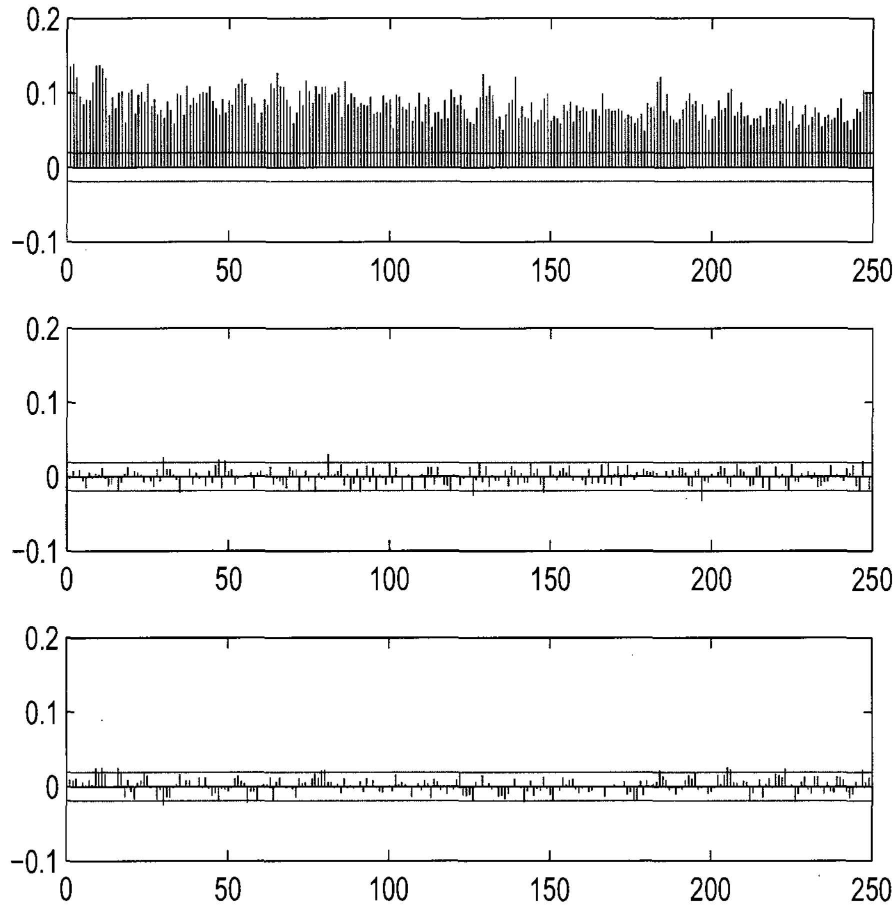

Figure 15: ACF of the filtered squared 30-minute log returns (top graph), the residuals (middle graph) and the squared residuals (bottom graph).

4 Conclusions

4.1 Summary

In financial econometrics, discrete-time GARCH processes are widely used to model the returns observed at regular intervals. Continuous time models are especially useful for the analysis of irregularly spaced data, and high-frequency data. In this paper, a family of continuous time GARCH processes, generalizing the COGARCH$(1,1)$ process of Klüppelberg, et. al. (2004), was introduced and studied. The resulting COGARCH(p,q) processes, $q \geq p \geq 1$, exhibit many of the characteristic features of observed financial time series, such as tail heaviness, volatility clustering and dependence without correlation. As in the discrete time case, the volatility and squared increment process of the COGARCH(p,q) model display a broader range of autocorrelation structures than those of the COGARCH$(1,1)$ process. We established sufficient conditions for the existence of a strictly stationary non-negative solution of the equations for the volatility process and, under conditions which ensure the finiteness of the required moments, determined the autocorrelation functions of both the volatility and squared increment processes. The volatility process was found to have the autocorrelation function of a continuous-time ARMA process while the squared increment process was found to have the autocorrelation function of an ARMA process just as in the discrete time case.

We proposed a least-squares method to estimate the parameters of a COGARCH(2,2) process, making use of the property that the autocorrelation function of the squared increments of the COGARCH(p,q) process is that of an ARMA(q,q) process. We showed that when the driving Lévy process is compound Poisson, then the state process and the squared increments of the COGARCH(2,2) process are strongly mixing with exponential rate, from which it follows that the least-squares estimators are strongly consistent and asymptotically normal. The COGARCH$(2,2)$ model with compound poisson driving process was fitted to the 30-minute log-return series of Dow Jones Industrial average from December 29th, 2003 to May 12th, 2006. For this series we estimated the volatilities and conducted a residual analysis which, as hoped, gave residuals compatible with white noise.

4.2 Future problems

Future problems to be investigated include the following:

- The condition (2.8) established in Theorem 2.4 is necessary and sufficient for stationarity of the state and the volatility processes in the special case $p = q = 1$, but only sufficient for processes with $q > 1$. It would be of interest to investigate the degree to which this condition can be relaxed when $q > 1$.

- The strong mixing property shown in Theorem 3.6 is valid for COGARCH$(2,2)$ processes when the driving Lévy process is compound Poisson. The proof of Lemma 3.4, which shows the φ-irreducibility of the COGARCH$(2,2)$ state process, needs to be generaziled to COGARCH(p,q) processes with $q > 2$ and with general driving Lévy process. This would allow the strong mixing property to be established in the general case and hence to establish asymptotic properties of the parameter estimators in greater generality.

- The COGARCH$(1,1)$ model shows very similar behaviour to that of its discrete-time analogue, the GARCH$(1,1)$ process. Analogous connections between higher order COGARCH and GARCH processes exist and will be the subject of further investigation.

- Comparisons of COGARCH models fitted to observations of the same process made at different frequencies.

References

Applebaum, D. (2004). *Lévy Processes and Stochastic Calculus.* Cambridge University Press, Cambridge.

Barndorff-Nielsen, O.E. (2001). Superposition of Ornstein-Uhlenbeck type processes. *Theory Probab. Appl.* **45**, 175–194.

Barndorff-Nielsen, O.E. and Shephard, N. (2001a). Modelling by Lévy processes for financial econometrics. In: O.E. Barndorff-Nielsen, T. Mikosch and S. Resnick (Eds.), *Lévy Processes: Theory and Applications,* pp. 283–318, Birkhäuser, Boston.

Barndorff-Nielsen, O.E. and Shephard, N. (2001b). Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics (with discussion). *J. Roy. Statist. Soc. Ser. B* **63**, 167–241.

Basrak, B., Davis, R.A. and Mikosch, T. (2002). Regular variation of GARCH processes. *Stochastic Process. Appl.* **99**, 95–115.

Bertoin, J. (1996). *Lévy Processes.* Cambridge University Press, Cambridge.

Bougerol, P. and Picard, N. (1992). Stationarity of GARCH processes and of some nonnegative time series. *J. Econometrics* **52**, 115–127.

Brandt, A. (1986). The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Adv. in Appl. Probab.* **18**, 211–220.

Brockwell, P.J. (2000). Continuous–time ARMA processes. In: C.R. Rao and D.N. Shanbhag (Eds.), *Stochastic processes: theory and methods, Handbook of Statist.* **19**, 249-276, North–Holland, Amsterdam.

Brockwell, P.J. (2001). Lévy-driven CARMA processes. *Ann. Inst. Stat. Mat.* **53**, 113–124.

Brockwell, P.J. (2004). Representations of continuous-time ARMA processes. *J. Appl. Probab.* **41A**, 375–382.

Brockwell, P.J., Chadraa, E., and Lindner, A. (2006). Continuous-time GARCH processes. *The Ann. of Appl. Probab.* **16**, 790–826.

Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods.* 2nd edition, Springer, New York.

Brockwell, P.J., and Marquardt, T. (2005). Fractionally integrated continuous-time ARMA process. *Statistica Sinica*, **15**, 477–494.

Brodin, E. and Klüppelberg, C. (2006). Modeling, estimation and visualization of multivariate dependence for high-frequency data. *Preprint* Munich University of Technology. `www.ma.tum.de/stat/`.

Doukhan, P. (1994) *Mixing. Properties and Examples.* Lecture Notes in Statistics. **85**, Springer Verlag, New York.

Drost, F.C. and Werker, B.J.M. (1996). Closing the GARCH gap: continuous time GARCH modelling. *J. Econometrics* **74**, 31–57.

Duan, J.C. (1997). Augmented GARCH(p,q) process and its diffusion limit. *J. Econometrics* **79**, 97–127.

Fasen, V. (2004). *Lévy Driven MA Processes with Applications in Finance.* Ph.D. thesis, Technical University of Munich.

Fasen, V., Klüppelberg, C. and Lindner, A. (2006). Extremal behavior of stochastic volatility models. *Grossinho, M.d.R., Shiryaev, A., Esquivel, M., and Oliviera, P.E. (Eds.), Stochastic Finance* 107–155, Springer, New York.

Godunov, S.K. (1998). *Modern Aspects of Linear Algebra.* Translations of Mathematical Monographs **175**, American Mathematical Society, Rhode Island.

Goldstein, J.A. (1985). *Semigroups of Linear Operators and Applications.* Oxford Mathematical Monographs, Oxford University Press, Oxford.

Golub, G.H. and Van Loan, C.F. (1989). *Matrix Computations.* 2nd edition, John Hopkins University Press, Baltimore.

Hannan, E.J. (1970). *Multiple Time Series.* Wiley, New York.

Haug, S., Klüppelberg, C., Lindner, A., and Zapp, M. (2005). Estimating the COGARCH(1, 1) model - a first go. Available at `www.ma.tum.de/stat/`.

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

Karatzas, I. and Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus* (Second Edition), New York, Springer-Verlag.

Kesten, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131**, 207–248.

Klüppelberg, C., Lindner, A. and Maller, R. (2004). A continuous time GARCH process driven by a Lévy process: stationarity and second order behaviour. *J. Appl. Probab.* **41**, 601–622.

Klüppelberg, C., Lindner, A. and Maller, R. (2006). Continuous time volatility modelling: COGARCH versus Ornstein-Uhlenbeck models. *Kabanov, Y., Lipster, R. and Stoyanov, J. (Eds.) The Shiryaev Festschrift: From Stochastic Calculus to Mathematical Finance* 393–419, Springer, Berlin.

Lindner, A. and Maller, R. (2005). Lévy integrals and the stationarity of generalised Ornstein–Uhlenbeck processes. *Stoch. Proc. Appl.* **115**, 1701–1722.

Lütkepohl, H. (1996). *Handbook of Matrices*, Wiley, Chichester.

Nelson, D.B. (1990). ARCH models as diffusion approximations. *J. Econometrics* **45**, 7–38.

Protter, P.E. (2004). *Stochastic Integration and Differential Equations.* 2nd edition, Springer, New York.

Rao, C.R. and Rao, M.B. (1998). *Matrix Algebra and its Applications to Statistics and Economics.* World Scientific Publishing, New Jersey.

Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge.

Taylor, H.M. and Karlin, S. (1994). *An Introduction to Stochastic Modeling.* Revised Edition. Academic Press, Boston.

Todorov, V. and Tauchen, G. (2006). Simulation methods for Lévy-driven CARMA stochastic volatility models. *Journal of Business and Economic Statistics*, **24(4)**, 455–469.

Tsai, H. and Chan, K.S. (2004). A note on non-negative continuous-time processes. *J. Roy. Stat. Soc. (B)* **67**, 589–597.

Appendix

Simulating COGARCH(2,2) process

```
%---Input variables:  n,c,alpha_0,alpha_1,alpha_2,beta_1,beta_2
sigma=sqrt(1/c);
a=[alpha_1;alpha_2];
e=[0; 1];
B=[0 1; -beta_2 -beta_1];
%---Simulating jump times
jump_int(1)=-log(1-rand)/c;
T(1)=jump_int(1);
j=1;
for i=1:n
    n_jump(i)=0;
    while T(j)<i
        n_jump(i)=n_jump(i)+1;
        jump_int(j+1)=-log(1-rand)/c;
        T(j+1)=T(j)+jump_int(j+1);
        j=j+1;
    end
end
total_jumps=j-1;
%---Counting the number of jumps
N=ones(n,1);
N(1)=n_jump(1);
for i=2:n
    N(i)=N(i-1)+n_jump(i);
end
```

```
%---Simulating the driving process
dL=sigma*randn(size(1:total_jumps));
L=ones(total_jumps,1);
L(1)=dL(1);
for j=2:total_jumps
    L(j)=L(j-1)+dL(j);
end
%---Computing the processes V and Y
V=ones(total_jumps,1);
Y=ones(2,total_jumps);
Y(:,1)=ones(2,1);
Y(:,1)=expm(B*jump_int(1))*Y(:,1);
Y(:,1)=Y(:,1)+(alpha_0+conj(a')*Y(:,1))*(dL(1))^2*e;
V(1)=alpha_0+conj(a')*Y(:,1);
for j=2:total_jumps
    Y(:,j)=Y(:,j-1);
    Y(:,j)=expm(B*jump_int(j))*Y(:,j);
    Y(:,j)=Y(:,j)+(alpha_0+conj(a')*Y(:,j))*(dL(j))^2*e;
    V(j)=alpha_0+conj(a')*Y(:,j);
end
%---Computing the process G
G=ones(total_jumps,1);
G(1)=sqrt(V(1))*dL(1);
for j=2:total_jumps
    G(j)=G(j-1)+sqrt(V(j))*dL(j);
end
%---Realizations observed at integer times
i=1;
```

```
while N(i)==0
    i=i+1;
end
G_obs(1:i-1)=0;
G_obs(i:n)=G(N(i:n));
L_obs(1:i-1)=0;
L_obs(i:n)=L(N(i:n));
for j=1:i-1
    V_obs(j)=alpha_0+conj(a')*expm(B*j)*ones(2,1);
end
for j=i:n
    V_obs(j)=alpha_0+conj(a')*expm(B*(j-T(N(j))))*Y(:,N(j));
end
```

Least-squares estimation

```
function f = lse(var)

global X;

X=X-mean(X);

phi_1=var(1);

phi_2=var(2);

theta_1=var(3);

theta_2=var(4);

n=length(X);

e=zeros(n,1);

root=roots([-phi_2 -phi_1 1]);

if (abs(root(1)) < 1) || (abs(root(2)) < 1)

    f=inf;

elseif (phi_1^2+4*phi_2<=0 || phi_2>0)

    f=inf;

else

    e(1)=X(1);

    e(2)=X(2)-phi_1*X(1)-theta_1*e(1);

    for i=3:n

        e(i)=X(i)-phi_1*X(i-1)-phi_2*X(i-2)-theta_1*e(i-1)

        -theta_2*e(i-2);

    end

    f=0;

    for i=1:n

        f=f+e(i)^2;

    end

end
```

COGARCH(1, 2): Estimated volatilities and residuals

```
%---Determining the intervals where jump occurred
for i=1:n
    t(i)=floor(T(i))+1;
    index(i)=floor((T(i)-t(i)+1)*h)+1;
end
N_jumps(1)=N(1);
N_jumps(2:n)=N(2:n)-N(1:n-1);
V_hat=ones(n,1);
Y_hat=ones(2,n);
Y_hat(:,1)=[10; 1];
V_hat(1)=2;
resid=ones(n,1);
resid(1)=0;
%---Computing estimated volatilities and residuals
Y_hat(:,1)=(I+B/h)^h*[10; 1];
if N(1)==1
    Y_hat(:,1)=Y_hat(:,1)+(I+B/h)^(h-index(1))*e*(G(1))^2;
elseif N(1)>1
    Y_hat(:,1)=Y_hat(:,1)+(I+B/h)^(h-index(1))*e*(G(1))^2;
        for j=2:N(1)
    Y_hat(:,1)=Y_hat(:,1)+(I+B/h)^(h-index(j))*e*(G(j)-G(j-1))^2;
    end
end
V_hat(1)=alpha_0+conj(a')*Y_hat(:,1);
resid(1)=G_obs(1)/sqrt(V_hat(1));
for i=2:n
    Y_hat(:,i)=(I+B/h)^h*Y_hat(:,i-1);
```

```
    if N_jumps(i)>0
        for j=N(i-1)+1:N(i)
            Y_hat(:,i)=Y_hat(:,i)+(I+B/h)^(h-index(j))*e
            *(G(j)-G(j-1))^2;
        end
    end
    V_hat(i)=alpha_0+conj(a')*Y_hat(:,i);
    resid(i)=(G_obs(i)-G_obs(i-1))/sqrt(V_hat(i));
end
```

Filtering seasonal effect

```
dow5_unfiltered=dlmread('dow5_unfiltered.tsm');

%---Create 78 by 813 matrix

for i=1:78

    for j=1:813

        unfiltered1(i,j)=dow5_unfiltered((j-1)*78+i);

    end

end

%---Find the median absolute value for each column

for i=1:78

    v(i)=median(abs(unfiltered1(i,:)));

end

for i=1:78

    filtered1(i,:)=(unfiltered1(i,:)-mean(dow5_unfiltered))/v(i);

end

%---Create the filtered data as a row matrix

for i=1:78

    for j=1:813

        dow5_filtered((j-1)*78+i)=filtered1(i,j);

    end

end
```

Dow Jones: Estimated volatilities and residuals

```
%---Input:  h,alpha_0,alpha_1,alpha_2,beta_1,beta_2,Y_1,Y_2
a=[alpha_1; alpha_2];
B=[0 1; -beta_2 -beta_1];
k=length(dow30_filtered);
for i=1:k
    dow30_filtered(i)=sum(dow5_filtered((i-1)*6+1:(i-1)*6+6));
end
dow5_filtered_sq=dow5_filtered.^2;
V_hat=ones(k,1);
Y_hat=ones(2,k);
Y_hat(:,1)=expm(B)*[Y_1; Y_2];
for j=1:h
    Y_hat(:,1)=Y_hat(:,1)+(I+B/h)^(h-j)*e*dow5_filtered_sq(j);
end
V_hat(1)=alpha_0+conj(a')*Y_hat(:,1);
resid(1)=dow30_filtered(1)/sqrt(V_hat(1));
for i=2:k
    Y_hat(:,i)=expm(B)*Y_hat(:,i-1);
    for j=1:h
        Y_hat(:,i)=Y_hat(:,i)+(I+B/h)^(h-j)*e
        *dow5_filtered_sq((i-1)*h+j);
    end
    V_hat(i)=alpha_0+conj(a')*Y_hat(:,i);
    resid(i)=dow30_filtered(i)/sqrt(V_hat(i));
end
mean_resid=mean(resid);
std_resid=std(resid);
```

```
resid_sq=resid.^2;
mean_resid_sq=mean(resid_sq);
var_resid_sq=var(resid_sq);
Q_LB=0;
Q_ML=0;
for i=1:189
    rho_resid(i)=(resid(1:n-i)-mean_resid)'*(resid(1+i:n)-mean_resid)/
        n/std_resid^2;
    rho_resid_sq(i)=(resid_sq(1:n-i)-mean_resid_sq)'*(resid_sq(1+i:n)-
        mean_resid_sq)/n/var_resid_sq;
    Q_LB=Q_LB+rho_resid(i)^2/(n-i)*n*(n+2);
    Q_ML=Q_ML+rho_resid_sq(i)^2/(n-i)*n*(n+2);
end
```