Haug, Klüppelberg, Lindner, Zapp:

Estimating the COGARCH(1,1) model - a first go

Sonderforschungsbereich 386, Paper 458 (2005)

Estimating the COGARCH(1,1) model - a first go

S. Haug* C. Klüppelberg* A. Lindner* M. Zapp*

October 21, 2005

Abstract

We suggest moment estimators for the parameters of a continuous time GARCH(1,1) process based on equally spaced observations. Using the fact that the increments of the COGARCH(1,1) process are ergodic, the resulting estimators are consistent. We investigate the quality of our estimators in a simulation study based on the compound Poisson driven COGARCH model. The estimated volatility with corresponding residual analysis is also presented.

2000 MSC Subject Classifications: primary: 91B70, 91B84
secondary: 62F10, 62F12

JEL Classification: C23, C52

Keywords: continuous time GARCH process, GARCH process, Lévy process, moment estimators, stochastic volatility, volatility estimation

*Center for Mathematical Sciences, Munich University of Technology, D-85747 Garching, Germany, emails: {haug , cklu , lindner , zapp}@ma.tum.de, www.ma.tum.de/stat/

1 Introduction

The GARCH(1,1) process is a model widely used by practitioners in the financial industry. It is defined as

$$Y_n = \sigma_n \epsilon_n \quad \text{with} \quad \sigma_n^2 = \beta + \lambda Y_{n-1}^2 + \delta \sigma_{n-1}^2 \,, \quad n \in \mathbb{N}, \tag{1.1}$$

where $\beta > 0, \lambda, \delta \geq 0$. This model captures some of the most prominent features in financial data, in particular in the volatility process. Empirical studies show that volatility changes randomly in time, has heavy or semi-heavy tails and clusters on high levels. These stylized features are modelled by the GARCH family as has been shown for the GARCH(1,1) process in detail by Mikosch and Starica [8].

The modern treatment of stochastic volatility models is mostly in continuous time. Approaches to create a continuous time GARCH model go back to Nelson [10] and we refer to Drost and Werker [3] for an overview. Such processes are diffusion limits to discrete time GARCH models, where, unfortunately, many of the above features of the GARCH process are wiped out in the limit; see Fasen, Klüppelberg and Lindner [4]. Since empirical work indicates upwards jumps in the volatility, a model driven by a Lévy process seems a natural approach. In Klüppelberg, Lindner and Maller [6, 7] such a model was suggested by iterating the volatility equation in (1.1) and replacing the noise variables ϵ_n by the jumps $\Delta L_t = L_t - L_{t-}$ of a Lévy process $L = (L_t)_{t \geq 0}$. A reparameterization, setting $\eta = -\log \delta$ and $\varphi = \lambda/\delta$, yields the following continuous time GARCH(1,1) model, where the parameter space is given by $\beta, \eta > 0$ and $\varphi \geq 0$.

The COGARCH(1,1) process $G = (G_t)_{t \geq 0}$ is defined as the solution to the SDEs

$$dG_t = \sigma_t \, dL_t, \quad t \geq 0 \,, \tag{1.2}$$

$$d\sigma_{t+}^2 = (\beta - \eta \sigma_t^2) \, dt + \varphi \, \sigma_t^2 d[L, L]_t^{(d)}, \quad t \geq 0 \,, \tag{1.3}$$

where $[L, L]_t^{(d)} = \sum_{0 < s \leq t} (\Delta L_s)^2$ is the discrete part of the quadratic variation process $[L, L]$ of the Lévy process L, $G_0 := 0$ and σ_0^2 is taken to be independent of L. Throughout we assume that L is càdlàg, and we denote by ν_L the Lévy measure of L, which is assumed to be non-zero, and by $\tau_L^2 \geq 0$ the variance of the Brownian motion component of L (see Sato [12] for the basic definitions and notations concerning Lévy processes). Whereas the process G is taken as being càdlàg, for the volatility process we assume càglàd sample paths.

The quantity σ_t^2 is called the *instantaneous volatility* or *spot volatility*, which is assumed to be stationary and latent. In contrast to classical stochastic volatility models, it is not independent of the process, which drives the price process. On the contrary, L drives both, the volatility and the price process. Note that G jumps at the same times as L does and has jump size $\Delta G_t = \sigma_t \Delta L_t$, and that ΔL_t is independent of $\sigma_t = \sigma_{t-}$.

If our data consist of returns over intervals of time of length $r > 0$, denote

$$G_t^{(r)} := G_t - G_{t-r} = \int_{(t-r,t]} \sigma_s \, dL_s \,, \quad t \geq r \,,$$

and $(G_{nr}^{(r)})_{n \in \mathbb{N}}$ describes an equidistant sequence of such non-overlapping returns of length r. Calculating the corresponding quantity for the volatility yields

$$\begin{aligned} \sigma_{rn}^{2(r)} &:= \sigma_{rn}^2 - \sigma_{r(n-1)}^2 = \int_{(r(n-1),rn]} \left((\beta - \eta \sigma_s^2) \, ds + \varphi \, \sigma_s^2 \, d[L,L]_s^{(d)} \right) \\ &= \beta r - \eta \int_{(r(n-1),rn]} \sigma_s^2 \, ds + \varphi \int_{(r(n-1),rn]} \sigma_s^2 \, d[L,L]_s^{(d)} \,. \end{aligned} \tag{1.4}$$

It is also worth noting that the stochastic process

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

is the discrete part of the quadratic variation $[G,G]_t = \int_0^t \sigma_s^2 \, d[L,L]_s$ of G, so that $\int_{(r(n-1),rn]} \sigma_s^2 \, d[L,L]_s^{(d)}$ in (1.4) corresponds to the jump part of the quadratic variation of G accumulated during $(r(n-1), rn]$.

The goal of this paper is to estimate the model parameters β, η, φ. Moreover, we shall present a simple estimate of the volatility. We would like to mention that Müller [9] developed an MCMC estimation procedure for the COGARCH(1,1) model, which works also for irregularly spaced observations.

An important role is played by the *auxiliary process*

$$X_t = \eta t - \sum_{0<s\leq t} \log(1 + \varphi \, (\Delta L_s)^2) \,, \quad t \geq 0 \,. \tag{1.5}$$

The stationary volatility process has, for instance, the representation

$$\sigma_t^2 = \left(\beta \int_0^t e^{X_s} ds + \sigma_0^2 \right) e^{-X_{t-}} \,, \quad t \geq 0 \,, \tag{1.6}$$

with $\beta > 0$ and $\sigma_0^2 \stackrel{d}{=} \beta \int_0^\infty e^{-X_t} dt$, independent of L. The auxiliary process $(X_t)_{t \geq 0}$ itself is a spectrally negative Lévy process of bounded variation with drift η, no Gaussian component (i.e. $\tau_X^2 = 0$), and Lévy measure ν_X given by

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

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

$$\Psi(s) = -\eta s + \int_{\mathbb{R}} ((1 + \varphi \, x^2)^s - 1) \, \nu_L(dx) \,, \quad s \geq 0 \,. \tag{1.7}$$

For fixed $s \geq 0$ the Laplace transform $\mathbb{E}e^{-sX_t}$ is finite for one and hence all $t > 0$, if and only if the integral appearing in (1.7) is finite. This is equivalent to $\mathbb{E}|L_1|^{2s} < \infty$. Stationarity of the volatility process is in particular implied by the existence of some $s > 0$ such that $\Psi(s) \leq 0$.

One of the advantages of the COGARCH is that its second order structure is well-known. In the following result we give the moments of $G_n^{(r)}$, which are independent of n by stationarity: expressions (1.8) and (1.10) have been already proved in Proposition 5.1 of Klüppelberg et al. [6], there however under some additional assumptions such as bounded variation of L for (1.10). In Appendix A we shall give a different proof under less restrictive assumptions and also calculate the fourth moment of G.

Proposition 1.1. *Suppose that the Lévy process $(L_t)_{t\geq 0}$ has finite variance and zero mean, and that $\Psi(1) < 0$. Let $(\sigma_t^2)_{t\geq 0}$ be the stationary volatility process, so that $(G_t)_{t\geq 0}$ has stationary increments. Then $\mathbb{E}(G_t^2) < \infty$ for all $t \geq 0$, and for every $t, h \geq r > 0$ it holds*

$$\mathbb{E}(G_t^{(r)}) = 0\,, \quad \mathbb{E}(G_t^{(r)})^2 = \frac{\beta r}{|\Psi(1)|}\mathbb{E}(L_1^2)\,, \quad \mathrm{Cov}\,(G_t^{(r)}, G_{t+h}^{(r)}) = 0. \tag{1.8}$$

If further $\varphi > 0$, $\mathbb{E}(L_1^4) < \infty$ and $\Psi(2) < 0$, then $\mathbb{E}(G_t^4) < \infty$ for all $t \geq 0$ and, if additionally the Lévy measure ν_L of L is such that $\int_{\mathbb{R}} x^3 \nu_L(dx) = 0$, then it holds for every $t, h \geq r > 0$

$$\begin{aligned}\mathbb{E}(G_t^{(r)})^4 &= 6\mathbb{E}(L_1^2)\frac{\beta^2}{\Psi(1)^2}(2\eta\varphi^{-1} + 2\tau_L^2 - \mathbb{E}(L_1^2))\left(\frac{2}{|\Psi(2)|} - \frac{1}{|\Psi(1)|}\right)\left(r - \frac{1 - e^{-r|\Psi(1)|}}{|\Psi(1)|}\right) \\ &\quad + \frac{2\beta^2}{\varphi^2}\left(\frac{2}{|\Psi(2)|} - \frac{1}{|\Psi(1)|}\right) r + 3\frac{\beta^2}{\Psi(1)^2}(\mathbb{E}(L_1^2))^2 r^2 \end{aligned} \tag{1.9}$$

and

$$\begin{aligned}\mathrm{Cov}\,((G_t^{(r)})^2, (G_{t+h}^{(r)})^2) &= \frac{\beta^2}{|\Psi(1)|^3}(2\eta\varphi^{-1} + 2\tau_L^2 - \mathbb{E}(L_1^2))\mathbb{E}(L_1^2)\left(\frac{2}{|\Psi(2)|} - \frac{1}{|\Psi(1)|}\right) \\ &\quad \times \left(1 - e^{-r|\Psi(1)|}\right)\left(e^{r|\Psi(1)|} - 1\right) e^{-h|\Psi(1)|}. \end{aligned} \tag{1.10}$$

Lemma 1.2. *Under the conditions of Proposition 1.1 the process $((G_{nr}^{(r)})^2)_{n\in\mathbb{N}}$ has for each fixed $r > 0$ the autocorrelation structure of an* ARMA(1,1) *process.*

Proof. Denote by $\gamma(h) = \mathrm{Cov}\,((G_{nr}^{(r)})^2, (G_{(n+h)r}^{(r)})^2)$, $h \in \mathbb{N}_0$, the autocovariance function and by $\rho(h) = \mathrm{Corr}\,((G_{nr}^{(r)})^2, (G_{(n+h)r}^{(r)})^2)$, $h \in \mathbb{N}_0$, the autocorrelation function of the discrete time process $((G_{nr}^{(r)})^2)_{n\in\mathbb{N}}$. Then

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

Moreover, for $h = 1$ we get

$$\rho(1) = \frac{\gamma(1)}{\mathbb{V}\mathrm{ar}\,(G_r^2)}\,.$$

Recalling the autocorrelation function of an ARMA(1,1) process (see e.g. Brockwell and Davis [2], Exercise 3.16), we identify $e^{-r|\Psi(1)|}$ as the autoregressive root ϕ. The moving average root θ can be determined by matching $\rho(1) = (1+\phi\theta)(\phi+\theta)/(1+\theta^2+2\phi\theta)$. $\square$

Remark 1.3. From Corollary 4.4 of Klüppelberg et al. [6] we know the moments $\mathbb{E}(\sigma^{2k})$ of the stationary volatility process ($k \in \mathbb{N}$), which exist if and only if $\mathbb{E}(L_1^{2k}) < \infty$ and $\Psi(k) < 0$. In particular, if $\mathbb{E}(L_1^4) < \infty$ and $\Psi(2) < 0$, then for $t, h \geq 0$

$$\mathbb{E}(\sigma_t^2) \;=\; \frac{\beta}{|\Psi(1)|} \quad \text{and} \quad \mathbb{E}(\sigma_t^4) \;=\; \frac{2\beta^2}{|\Psi(1)\Psi(2)|}, \tag{1.11}$$

$$\mathbb{C}\mathrm{ov}\,(\sigma_t^2, \sigma_{t+h}^2) \;=\; \beta^2\left(\frac{2}{|\Psi(1)\Psi(2)|} - \frac{1}{\Psi(1)^2}\right) e^{-h|\Psi(1)|} \;=\; \mathbb{V}\mathrm{ar}\,(\sigma_t^2)\, e^{-h|\Psi(1)|}\,. \tag{1.12}$$

Econometric literature suggests that volatility is quite persistent, which would imply that $e^{-|\Psi(1)|}$ is close to 1; i.e. $\Psi(1) < 0$ near 0. This should be kept in mind, when estimating the model parameters.

2 Method of moment estimation

2.1 Identifiability of the model parameters

We aim at estimation of the model parameters (β, η, φ) from a sample of equally spaced returns, matching empirical autocorrelation function and moments to their theoretical counterparts given in Proposition 1.1. In our next result we show that the parameters are identifiable by this estimation procedure for driving Lévy processes L as in Proposition 1.1 for which the variance of L and the variance τ_L^2 of the Brownian motion component in L are known. For the sake of simplicity we set $r = 1$ and $\mathbb{V}\mathrm{ar}\,(L_1) = 1$.

Theorem 2.1. *Suppose $(L_t)_{t\geq 0}$ is a Lévy process such that $\mathbb{E}L_1 = 0$, $\mathbb{V}\mathrm{ar}\,L_1 = 1$, the variance τ_L^2 of the Brownian motion component of L is known with $0 \leq \tau_L^2 < \mathbb{V}\mathrm{ar}\,L_1 = 1$), $\mathbb{E}(L_1^4) < \infty$ and $\int_{\mathbb{R}} x^3\,\nu_L(dx) = 0$. Assume also that $\Psi(2) < 0$, and denote by $(G_n^{(1)})_{n\in\mathbb{N}}$ the stationary increment process of the* COGARCH *process with parameters $\beta, \eta, \varphi > 0$. Let $m_1, m_2, k, k_\rho, p > 0$ be constants such that*

$$\begin{aligned}
\mathbb{E}(G_n^{(1)})^2 &\;=\; m_1,\\
\mathbb{E}(G_n^{(1)})^4 &\;=\; m_2,\\
\gamma(h) \;=\; \mathbb{C}\mathrm{ov}\,((G_n^{(1)})^2, (G_{n+h}^{(1)})^2) &\;=\; k\,e^{-hp}\,, \quad h \in \mathbb{N}\,,\\
\rho(h) \;=\; \mathbb{C}\mathrm{orr}\,((G_n^{(1)})^2, (G_{n+h}^{(1)})^2) &\;=\; k_\rho e^{-hp}\,, \quad h \in \mathbb{N}\,,
\end{aligned}$$

with $k_\rho = k/(m_2 - m_1^2)$. Define

$$\begin{aligned} M_1 &:= m_2 - 3m_1^2 - 6\frac{1-p-e^{-p}}{(1-e^p)(1-e^{-p})}\,k\,, \\ M_2 &:= \frac{2kp}{M_1(e^p-1)(1-e^{-p})}. \end{aligned}$$

Then $M_1, M_2 > 0$, and the parameters β, η, φ are uniquely determined by m_1, m_2, k and p and are given by the formulas

$$\begin{aligned} \beta &= p\,m_1\,, && (2.1) \\ \varphi &= p\sqrt{1+M_2} - p, && (2.2) \\ \eta &= p\sqrt{1+M_2}\,(1-\tau_L^2) + p\,\tau_L^2 = p + \varphi(1-\tau_L^2)\,. && (2.3) \end{aligned}$$

Proof. Since $r = \mathbb{E}(L_1^2) = 1$, we obtain from Proposition 1.1

$$\begin{aligned} m_1 &= \frac{\beta}{|\Psi(1)|}, && (2.4) \\ m_2 &= 6\,\frac{\beta^2}{|\Psi(1)|^3}\left(2\eta\varphi^{-1} + 2\tau_L^2 - 1\right)\left(\frac{2}{|\Psi(2)|} - \frac{1}{|\Psi(1)|}\right)\left(|\Psi(1)| - 1 + e^{-|\Psi(1)|}\right) \\ &\quad + \frac{2\beta^2}{\varphi^2}\left(\frac{2}{|\Psi(2)|} - \frac{1}{|\Psi(1)|}\right) + 3\frac{\beta^2}{\Psi(1)^2}, && (2.5) \\ p &= |\Psi(1)|, && (2.6) \\ k &= \frac{\beta^2}{|\Psi(1)|^3}\left(2\eta\varphi^{-1} + 2\tau_L^2 - 1\right)\left(\frac{2}{|\Psi(2)|} - \frac{1}{|\Psi(1)|}\right)\left(1 - e^{-|\Psi(1)|}\right)\left(e^{|\Psi(1)|} - 1\right) && (2.7) \end{aligned}$$

Then (2.4) and (2.6) immediately give (2.1). Inserting (2.7) in (2.5) and using (2.4) and (2.6), we obtain

$$m_2 = 6\frac{p-1+e^{-p}}{(1-e^{-p})(e^p-1)}\,k + \frac{2m_1^2p^2}{\varphi^2}\left(\frac{2}{|\Psi(2)|} - \frac{1}{p}\right) + 3m_1^2.$$

By definition of M_1 and (A.5), we see that

$$M_1 = \frac{2m_1^2p^2}{\varphi^2}\left(\frac{2}{|\Psi(2)|} - \frac{1}{p}\right) = \frac{2m_1^2p^2}{\varphi^2}\,\frac{\varphi^2}{|\Psi(2)|p}\int_{\mathbb{R}^4} x^4\,\nu_L(dx) > 0,$$

so that

$$\frac{2}{|\Psi(2)|} - \frac{1}{p} = \frac{M_1\varphi^2}{2m_1^2p^2}.$$

Inserting this in (2.7) gives

$$k = \frac{2\eta\varphi^{-1} + \tau_L^2 - 1}{p^3}\,\frac{M_1\varphi^2}{2}\,(1-e^{-p})(e^p-1),$$

so that

$$0 < pM_2 = \frac{2kp^2}{M_1(e^p - 1)(1 - e^{-p})} = \frac{2\eta\varphi^{-1} + 2\tau_L^2 - 1}{p}\varphi^2 = \left(2 + \frac{\varphi}{p}\right)\varphi,$$

where we used

$$p = |\Psi(1)| = \eta - \varphi(\mathbb{E}(L_1^2) - \tau_L^2) \tag{2.8}$$

from (1.7). Solving this quadratic equation in φ gives (2.2), which together with (2.8) implies (2.3). $\square$

We conclude from (2.1)-(2.3) that our model parameter vector (β, η, φ) is a continuous function of the first two moments m_1, m_2 and the parameters of the autocorrelation function p and k_ρ. Hence, by continuity, consistency of the moments implies immediately consistency of the corresponding plug-in estimates for (β, η, φ).

2.2 The estimation algorithm

The parameters are estimated under the following assumptions:

(H1) We have equally spaced observations G_n, $n = 0, \ldots, N$, giving return data $G_n^{(1)} = G_n - G_{n-1}$, $n = 1, \ldots, N$.

(H2) $\mathbb{E}L_1 = 0$ and $\mathbb{V}\mathrm{ar}\,(L_1) = 1$, i.e. σ^2 can be interpreted as the volatility.

(H3) The variance τ_L^2 of the Brownian motion component of L is known and in $[0, 1)$.

(H4) $\int_{\mathbb{R}} x^3 \, \nu_L(dx) = 0$, $\mathbb{E}(L_1^4) < \infty$ and $\Psi(2) < 0$.

We estimate m_1, m_2 by their empirical counterparts and the parameters p, k_ρ of the autocorrelation function using a least squares estimate; see Seber and Wild [13] for theoretical background.

Algorithm 2.2. *(1) Calculate moment estimators*

$$\widehat{m}_1 := \frac{1}{N}\sum_{n=1}^{N}(G_n^{(1)})^2 \quad \text{and} \quad \widehat{m}_2 := \frac{1}{N}\sum_{n=1}^{N}(G_n^{(1)})^4 .$$

(2) For fixed $h_{\max} \geq 2$ and a compact set $K \subset \mathbb{R}_+^2$ minimize

$$\sum_{h=1}^{h_{\max}} (\widehat{\rho}(h) - k_\rho e^{-ph})^2$$

with respect to k_ρ and p, where $\widehat{\rho}(h)$ is the empirical autocorrelation function at lag h as defined in (2.12). This yields estimators $\widehat{k}_\rho$ and $\widehat{p}$.

(3) Calculate $\widehat{k} = \widehat{k}_\rho(\widehat{m}_2 - \widehat{m}_1^2)$ *and insert* $\widehat{m}_1, \widehat{m}_2, \widehat{k}$ *and* $\widehat{p}$ *into (2.1)-(2.3). This yields estimators for* β, η, φ*:*

$$\widehat{\beta} = \widehat{p}\,\widehat{m}_1\,, \tag{2.9}$$

$$\widehat{\varphi} = \widehat{p}\sqrt{1+\widehat{M_2}} + \widehat{p}\,, \tag{2.10}$$

$$\widehat{\eta} = \widehat{p}\sqrt{1+\widehat{M_2}}\,(1-\tau_L^2) + \widehat{p}\,\tau_L^2 = \widehat{p} + \widehat{\varphi}\,(1-\tau_L^2). \tag{2.11}$$

In part (2) of the above algorithm we fitted the autocorrelation function. Alternatively, we could also have based our least squares estimation on the autocovariance function. It turned out, however, that the estimators chosen as above are considerably more accurate. The reason for this is that k_ρ is independent of β (it cancels out). This improves the estimation as the estimator for β has the largest error among all our estimators.

Underlying the least squares estimation above is a non-linear regression model. If the errors were i.i.d., then under certain regularity conditions these estimators would be consistent and asymptotically normal. In our case, the errors are not to be expected to be i.i.d., quite contrary, the dependence structure inherent in the correlation estimates $\widehat{\rho}(h)$ will be quite complicated. For the discrete time GARCH(1,1) model this has been investigated in Mikosch and Starica [8].

In Theorem 2.1 it was shown that M_1 and M_2 are strictly positive. This does not imply a-priori that the empirical estimates $\widehat{M_1}$ and $\widehat{M_2}$ are strictly positive and that $\sqrt{1+\widehat{M_2}}$ is well-defined. As we shall show in the next section, the COGARCH(1,1) model is ergodic and mixing, which suffices to prove strong consistency of the above estimators. In particular, $\widehat{M_2}$ will be strictly positive for large samples sizes and the afore mentioned problem does not occur.

2.3 Consistency of moment estimators

The proof of the mixing property of the COGARCH increments and hence of the following theorem can be found in Appendix B. Since we are only interested in applying it to show consistency of the moment estimator, we do not state and prove it under the most general assumptions possible.

Theorem 2.3. *Suppose that* $(L_t)_{t\geq 0}$ *is such that* $\mathbb{E}(L_1^4) < \infty$ *and the parameters of the COGARCH process are such that* $\Psi(2) < 0$*. Let* $(\sigma_t^2)_{t\geq 0}$ *be the strictly stationary volatility process given as solution to* (1.3)*. Then the process* $(G_{rn}^{(r)})_{n\in\mathbb{N}}$ *is strictly stationary and ergodic for every* $r > 0$*.*

Theorem IV.2.2 of Hannan [5] ensures that empirical moments and the empirical covariance function converge almost surely under strict stationarity and ergodicity to their theoretical counterparts.

We apply this to the process $(G_n^{(1)})^2_{n\in\mathbb{N}}$. For given data $(G_1^{(1)})^2, (G_2^{(1)})^2, \ldots, (G_N^{(1)})^2$ the empirical moments and autocovariance and autocorrelation functions are given by

$$\begin{aligned}
\widehat{m}_1 &= \frac{1}{N}\sum_{n=1}^{N}(G_n^{(1)})^2 \\
\widehat{m}_2 &= \frac{1}{N}\sum_{n=1}^{N}(G_n^{(1)})^4 \\
\widehat{\gamma}(h) &= \frac{1}{N}\sum_{n=1}^{N-h}((G_{n+h}^{(1)})^2 - \widehat{m}_1)((G_n^{(1)})^2 - \widehat{m}_1)\,, \quad h \geq 1 \\
\widehat{\gamma}(0) &= \widehat{m}_2 - \widehat{m}_1^2 \\
\widehat{\rho}(h) &= \widehat{\gamma}(h)/\widehat{\gamma}(0)\,, \quad h \geq 0\,.
\end{aligned} \tag{2.12}$$

All these empirical moments converge almost surely to their theoretical counterparts; i.e. as $N \to \infty$,

$$\widehat{m}_1 \stackrel{\text{a.s.}}{\to} \mathbb{E}(G_t^{(1)})^2\,, \quad \widehat{m}_2 \stackrel{\text{a.s.}}{\to} \mathbb{E}(G_t^{(1)})^4\,, \quad (\widehat{\gamma}(1),\ldots,\widehat{\gamma}(h_{\max})) \stackrel{\text{a.s.}}{\to} (\gamma(1),\ldots,\gamma(h_{\max})),$$
$$\text{and} \quad (\widehat{\rho}(1),\ldots,\widehat{\rho}(h_{\max})) \stackrel{\text{a.s.}}{\to} (\rho(1),\ldots,\rho(h))\,. \tag{2.13}$$

In the next Theorem it is shown that also $\widehat{k}_\rho \stackrel{\text{a.s.}}{\to} k_\rho$ and $\widehat{p} \stackrel{\text{a.s.}}{\to} p$, so that consistency of the moment estimator follows.

Theorem 2.4. *Let $(G_t)_{t\geq 0}$ be the* COGARCH(1,1) *process with strictly stationary volatility process $(\sigma_t^2)_{t\geq 0}$ given by (1.2) and (1.3). Assume that $(L_t)_{t\geq 0}$ satisfies the assumptions of Theorem 2.1. Let K be a compact subset of $\mathbb{R}_+ \times \mathbb{R}_+$ containing the true value $\boldsymbol{\theta}^0 := (k_\rho^0, p^0)$. Then the moment estimators $\widehat{\beta}, \widehat{\eta}, \widehat{\varphi}$ as defined in Algorithm 2.2 by (2.9), (2.10), and (2.11) are strongly consistent.*

Proof. Define for appropriate $h_{\max} << N$ and $\boldsymbol{\theta} = (k_\rho, p) \in K$

$$m_N(\boldsymbol{\theta}) := \begin{pmatrix} \widehat{\rho}(1) - \rho_{\boldsymbol{\theta}}(1) \\ \vdots \\ \widehat{\rho}(h_{\max}) - \rho_{\boldsymbol{\theta}}(h_{\max}) \end{pmatrix},$$

where $\rho_{\boldsymbol{\theta}}(h) = k_\rho e^{-ph}$. Since $((G_n^{(1)})^2)_{n\in\mathbb{N}}$ is strictly stationary and ergodic by Theorem 2.3, we know from (2.13) that for all $\boldsymbol{\theta} \in K$,

$$m_N(\boldsymbol{\theta}) \xrightarrow{a.s.} m(\boldsymbol{\theta}), \qquad N \to \infty,$$

where

$$m(\boldsymbol{\theta}) := \begin{pmatrix} \rho_{\boldsymbol{\theta}^0}(1) - \rho_{\boldsymbol{\theta}}(1) \\ \vdots \\ \rho_{\boldsymbol{\theta}^0}(h_{\max}) - \rho_{\boldsymbol{\theta}}(h_{\max}) \end{pmatrix}.$$

Observing that

$$q(\boldsymbol{\theta}) := m(\boldsymbol{\theta})^T m(\boldsymbol{\theta})$$

has a unique minimum at $\boldsymbol{\theta} = \boldsymbol{\theta}^0$ which is equal to 0, we conclude

$$q_N(\boldsymbol{\theta}^0) \stackrel{\text{a.s.}}{\to} 0, \qquad N \to \infty, \tag{2.14}$$

where $q_N(\boldsymbol{\theta}) := m_N(\boldsymbol{\theta})^T m_N(\boldsymbol{\theta})$. For finite N we have $0 \le q_N(\widehat{\boldsymbol{\theta}}_N) \le q_N(\boldsymbol{\theta}^0)$, with $\widehat{\boldsymbol{\theta}}_N := \arg\min_{\boldsymbol{\theta}\in K} q_N(\boldsymbol{\theta})$. From (2.14) then follows that

$$q_N(\widehat{\boldsymbol{\theta}}_N) := m_N(\widehat{\boldsymbol{\theta}}_N)^T m_N(\widehat{\boldsymbol{\theta}}_N) \stackrel{\text{a.s.}}{\to} 0, \qquad N \to \infty. \tag{2.15}$$

Next, observe that

$$\begin{aligned}
|q_N(\widehat{\boldsymbol{\theta}}_N) - q(\widehat{\boldsymbol{\theta}}_N)| &= \left| \sum_{h=1}^{h_{\max}} \Big(\widehat{\rho}^2(h) - \rho^2_{\boldsymbol{\theta}^0}(h) + 2\rho_{\widehat{\boldsymbol{\theta}}_N}(h)\{\rho_{\boldsymbol{\theta}^0}(h) - \widehat{\rho}(h)\} \Big) \right| \\
&\le \sum_{h=1}^{h_{\max}} \Big(|\widehat{\rho}(h)| + |\rho_{\boldsymbol{\theta}^0}(h)| + 2|\rho_{\widehat{\boldsymbol{\theta}}_N}(h)| \Big) \, |\rho_{\boldsymbol{\theta}^0}(h) - \widehat{\rho}(h)| \\
&\le 4 \sum_{h=1}^{h_{\max}} |\rho_{\boldsymbol{\theta}^0}(h) - \widehat{\rho}(h)| \stackrel{\text{a.s.}}{\to} 0\,, \quad N \to \infty,
\end{aligned}$$

by (2.13), where we used that $|\widehat{\rho}(h)| \le 1$, see Brockwell and Davis [2], Problem 7.11. Together with (2.15) this implies

$$q(\widehat{\boldsymbol{\theta}}_N) \stackrel{\text{a.s.}}{\to} 0 = q(\boldsymbol{\theta}^0),$$

and since q has its only minimum at $\boldsymbol{\theta}^0$ with value $q(\boldsymbol{\theta}^0) = 0$, it follows that $\widehat{\boldsymbol{\theta}}_N \stackrel{\text{a.s.}}{\to} \boldsymbol{\theta}^0$. The estimators $\widehat{\beta}, \widehat{\eta}$ and $\widehat{\varphi}$ are uniquely determined by (2.9) – (2.11). Moreover, (β, η, φ) is a continuous function of the first two moments m_1, m_2 and the parameters of the autocorrelation function p and k_ρ. This shows together with the above conclusions strong consistency:

$$\widehat{\beta} \stackrel{\text{a.s.}}{\to} \beta, \quad \widehat{\eta} \stackrel{\text{a.s.}}{\to} \eta \quad \text{and} \quad \widehat{\varphi} \stackrel{\text{a.s.}}{\to} \varphi\,. \qquad \square$$

2.4 Compound Poisson COGARCH(1,1) process

This section is devoted to the compound Poisson COGARCH(1,1) process, which corresponds to a compound Poisson driving process L given by

$$L_t = \sum_{k=1}^{N_t} Y_k\,, \quad t \ge 0\,,$$

where $N = (N_t)_{t\ge 0}$ is a Poisson process with intensity $c > 0$, and $(Y_k)_{k\in\mathbb{N}}$ are i.i.d. random variables, independent of N. We introduce a generic random variable Y with the same distribution function as the Y_k, denoted by F_Y. For this model (H3) is clearly satisfied, with $\tau_L^2 = 0$. The Lévy measure of L has the representation $\nu(dx) = cF_Y(dx)$. This allows us to calculate the Laplace exponent from (1.7) getting

$$\Psi(s) = -\eta\, s + c \int_{\mathbb{R}} \big((1+\varphi\, y^2)^s - 1\big)\, F_Y(dy)\,.$$

From this we obtain

$$\Psi(1) = -\eta + \varphi c\mathbb{E}(Y^2) \quad \text{and} \quad \Psi(2) = -2\eta + 2\varphi c\mathbb{E}(Y^2) + \varphi^2 c\mathbb{E}(Y^4)\,.$$

Since Theorem 2.1 requires $\mathbb{E}(L_1) = 0$ and $\mathbb{V}\mathrm{ar}\,(L_1) = \mathbb{E}(L_1^2) = 1$, we must have $\mathbb{E}(Y^2) = 1/c$ yielding $p = |\Psi(1)| = \eta - \varphi$. The conditions $\mathbb{E}(L_1^4) < \infty$ and $\int_{\mathbb{R}} x^3\, \nu_L(dx) = 0$ translate into $\mathbb{E}(Y^4) < \infty$ and $\mathbb{E}(Y^3) = 0$, respectively. Moreover, we obtain $\Psi(2) = 2(\varphi - \eta) + \varphi^2\mathbb{E}(Y^4)/(\mathbb{E}(Y^2)) = -2p + \varphi^2\mathbb{E}(Y^4)/(\mathbb{E}(Y^2))$. Then the condition $\Psi(2) < 0$ translates into $\varphi^2 < 2p/(c\mathbb{E}(Y^4))$.

These conditions are satisfied for a driving compound Poisson process with jump intensity $c = 1$ and standard normally distributed jumps. The model parameters are chosen as $\beta = 1, \eta = 0.05$ and $\varphi = 0.04$. As starting value for σ_0^2 we choose the theoretical mean of the stationary model corresponding to the above parameters given by $\mathbb{E}(\sigma_\infty^2) = 10$.

In Figure 1 we plotted simulated sample paths for the time interval $[0, 3\,000]$ of the driving Lévy process L, the volatility process σ, the COGARCH(1,1) process G, and the differenced COGARCH $G^{(1)}$, respectively. All four sample paths have been simulated with the same random seed. As can be seen (G_t) looks similar to (L_t), they only differ by the jump sizes. Also the volatility clustering, which is observed in real data, can be rediscovered in this simulation.

In the next section we shall investigate the quality of the estimators given by Algorithm 2.2 in a simulation study. An important problem in financial data is the occurrence of volatility jumps. Recall that in our model the volatility jumps exactly when the price itself jumps. As jumps in a compound Poisson model with moderate frequency are rare, we should be able to estimate the jump rate c from the discretized data $G_n^{(1)}$. This is shown in our next result. The analysis is based on $z(N)$, the number of intervals, where G does not change; i.e. $z(N) = \sum_{t=1}^N 1_{\{G_t^{(1)}=0\}}$. This implies immediately that one needs a fine enough observation grid.

Proposition 2.5. *Let $(L_t)_{t\geq 0}$ be a compound Poisson process with continuous jump distribution F_Y and intensity $c > 0$. Then*

$$\widehat{c} = -\log\left(\frac{z(N)}{N}\right) \stackrel{\text{a.s.}}{\to} c\,, \quad N \to \infty\,,$$

and

$$\sqrt{N}(\widehat{c} - c) \stackrel{d}{\to} N(0, e^c(1 - e^{-c}))\,, \quad N \to \infty\,, \tag{2.16}$$

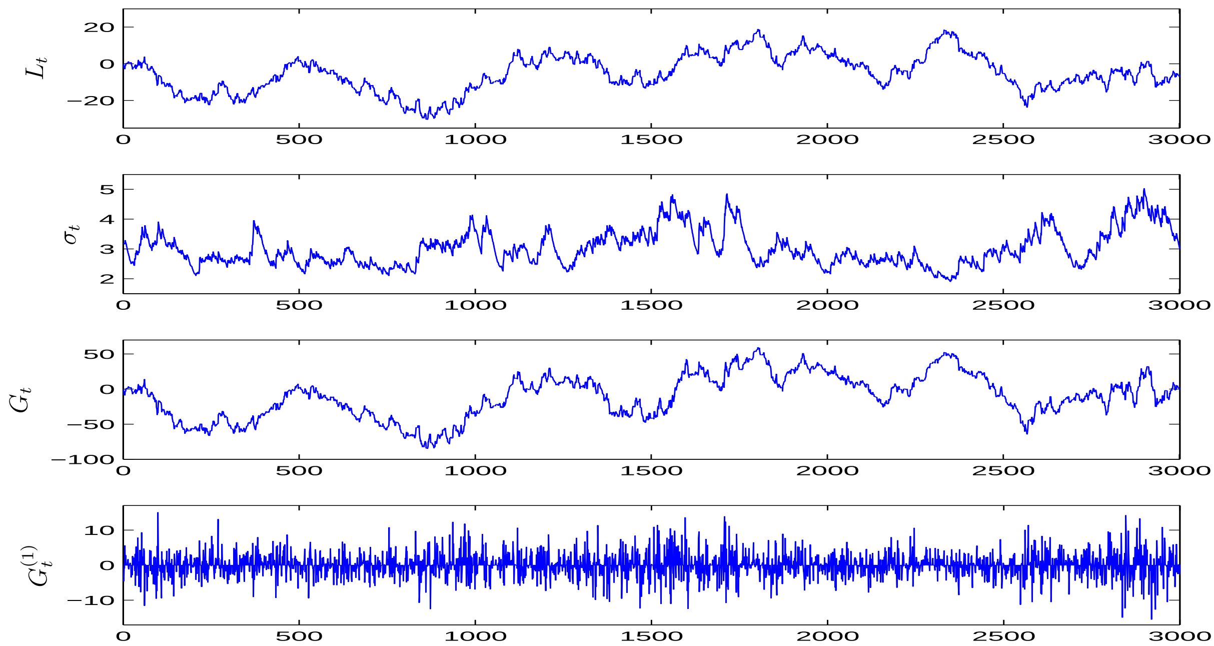

Figure 1: Simulated compound Poisson process $(L_t)_{0\le t\le 3\,000}$ with Poisson rate $c = 1$ and $N(0,1)$-distributed jumps (*first*), the volatility process (σ_t) (*second*), the corresponding COGARCH process (G_t) with parameters $\beta = 1, \eta = 0.05$ and $\varphi = 0.04$ (*third*) and the differenced COGARCH process $(G_t^{(1)})$ of order 1 (*last*).

where $N(\mu, \sigma^2)$ denotes the standard normal distribution with mean μ and variance σ^2.

Proof. Denote by S_n the number of jumps in the interval $(n-1, n]$. Then the $S_n,\ n = 1, \ldots, N$, are i.i.d. Poisson distributed with parameter c. Therefore, the indicator variables $1_{\{S_n=0\}},\ n = 1, \ldots, N$, are also i.i.d. Since F_Y is continuous, we have

$$1_{\{S_n=0\}} = 1_{\{G_n^{(1)}=0\}} \qquad a.s.\,, \quad n = 1, \ldots, N\,.$$

By the strong law of large numbers, we get

$$\frac{1}{N}\sum_{n=1}^{N} 1_{\{G_n^{(1)}=0\}} \xrightarrow{a.s.} \mathbb{E}(1_{\{S_1=0\}}) = P(S_1 = 0) = e^{-c}, \qquad T \to \infty\,,$$

and therefore

$$-\log\left(\frac{z(N)}{N}\right) \xrightarrow{a.s.} c, \qquad N \to \infty.$$

Moreover, as $1_{\{G_n^{(1)}=0\}}$, $n = 1, \dots, N$, are i.i.d., the central limit theorem applies giving

$$\frac{z(N) - Ne^{-c}}{\sqrt{Ne^{-c}(1-e^{-c})}} \xrightarrow{d} N(0,1)\,, \quad N \to \infty\,.$$

Invoking the delta-method to $-\log(\frac{z(N)}{N})$ (e.g. Brockwell and Davis [2], Proposition 6.4.1), using $\sqrt{e^{-c}(1-e^{-c})/N} \to 0$ as $N \to \infty$ and the fact that $-\log(\cdot)$ is differentiable (at e^{-c}), we obtain (2.16). □

Remark 2.6. The central limit theorem of Proposition 2.5 allows us to construct confidence intervals for the jump rate c. Using (2.16) and

$$\frac{N}{z(N)}\left(1 - \frac{z(N)}{N}\right) \xrightarrow{P} e^{c}(1-e^{-c}), \qquad N \to \infty,$$

we apply Slutzky's theorem to get

$$\frac{-\log(\frac{z(N)}{N}) - c}{\sqrt{\frac{1}{z(N)}(1 - \frac{z(N)}{N})}} \xrightarrow{d} N(0,1). \tag{2.17}$$

Solving (2.17) with respect to c, we get a $100(1-\alpha)\%$ confidence interval

$$\left[-\log\left(\frac{z(N)}{N}\right) - q_{1-\frac{\alpha}{2}}\sqrt{\frac{1}{z(N)} - \frac{1}{N}}, -\log\left(\frac{z(N)}{N}\right) + q_{1-\frac{\alpha}{2}}\sqrt{\frac{1}{z(N)} - \frac{1}{N}}\,\right],$$

where $q_{1-\frac{\alpha}{2}}$ is the $(1-\frac{\alpha}{2})$-quantil of the standard normal distribution.

3 Simulation study

In this section we investigate the behaviour of the moment estimators of Algorithm 2.2. As the driving Lévy process L we choose a compound Poisson process as in Section 2.4 with standard normally distributed jump sizes Y_k. Then all conditions of Theorem 2.1 on L are satisfied. To satisfy (H2) we have to choose the jump rate $c = 1$. Next we have to choose parameters β, η and φ. As indicated in Remark 1.4 the autocovariance function of $(G^{(1)})^2$ should not decrease too fast. From Proposition 1.1 we know that this is implied by $\Psi(1) < 0$ close to zero. Moreover, Theorem 2.1 requires $\Psi(2) < 0$. Setting $\beta = 0.1$, $\eta = 0.05$ and $\varphi = 0.04$ gives $\Psi(1) = -0.01$ and $\Psi(2) = -0.0152$ which are satisfactory values. To apply Algorithm 2.2 we also have to choose h_{max}. Numerical experience (see Zapp [14] for details) has shown that h_{max} equal to 150 is a good choice for a time series length of 3 000 observations.

3.1 Estimation results

We simulate 1 000 samples of 3 000 equidistant observations of $G^{(1)}$. Table 3.1 summarizes the outcome of our simulation study concerning the parameters β, η and φ.

The empirical mean of all the estimated parameter values $\widehat{\beta}$, $\widehat{\eta}$ and $\widehat{\varphi}$ is shown in the first line, with the empirical standard deviations in brackets. We also estimated bias, mean square error (MSE), mean absolute error (MAE), again with the corresponding standard deviation in brackets. The estimators $\widehat{\eta}$ and $\widehat{\varphi}$ show a better quality than $\widehat{\beta}$. This is not surprising as it is a well-known phenomenon that the drift of a model (1.3) is hard to estimate. One needs a very large sample for a precise estimator. The estimation results concerning the jump rate c and $v_{Y_1}^2$, the variance of the jumps Y_k, are shown in Table 3.2 showing satisfactory performance. Again we calculated the empirical mean, MSE and MAE with corresponding empirical standard deviations. As one can see from (2.3) and (2.6) $\Psi(1)$ is equal to $\Psi(1) = -\eta + \varphi$. Thus these two parameters give important characteristics of the model concerning stationarity and the rate p of decrease of the autocovariance and autocorrelation function. From (2.10) it is also clear, that the estimated parameters will always correspond to a stationary model, since $\widehat{p} > 0$ and thus $\widehat{\psi(1)} = -\widehat{\eta} + \widehat{\varphi} < 0$.

	$\widehat{\beta}$	$\widehat{\eta}$	$\widehat{\varphi}$
Mean	0.0984 (0.0014)	0.0447 (0.0004)	0.0344 (0.0003)
Bias	-0.0016 (0.0014)	-0.0053 (0.0004)	-0.0056 (0.0003)
MSE	0.0019 (1.3e-5)	0.0002 (1.0e-5)	0.0001 (0.9e-5)
MAE	0.0340 (0.0008)	0.0111 (0.0002)	0.0081 (0.0002)

Table 3.1: Estimated mean, bias, MSE and MAE for $\widehat{\beta}, \widehat{\eta}$ and $\widehat{\varphi}$ and corresponding estimated standard deviations in brackets. The true values are $\beta = 0.1$, $\eta = 0.05$ and $\varphi = 0.04$.

	$\widehat{c}$	$\widehat{v}_{Y_1}^2$
Mean	1.0007 (0.0007)	0.9999 (0.0007)
Bias	0.0007 (0.0007)	-0.8e-04 (0.0007)
MSE	0.0006 (0.2e-4)	0.0006 (0.2e-4)
MAE	0.0192 (0.0004)	0.0192 (0.0004)

Table 3.2: Estimated mean, bias, MSE and MAE for $\widehat{c}$ and $\widehat{v}_{Y_1}^2$ and corresponding estimated standard deviations in brackets. The true values are $c = 1$ and $v_{Y_1}^2 := \mathbb{V}\mathrm{ar}\,(Y_1) = 1$.

3.2 Estimation of the volatility σ_t^2

Recall from (1.4) for $r = 1$,

$$\sigma_n^2 = \sigma_{n-1}^2 + \beta - \eta \int_{(n-1,n]} \sigma_s^2 ds + \varphi \sum_{n-1<s\le n} \sigma_s^2 (\Delta L_s)^2 , \quad n \in \mathbb{N} . \tag{3.1}$$

Since σ_s is latent and ΔL_s is usually not observable, we have to approximate the integral and the sum on the right hand side. For the integral we use a simple Euler approximation

$$\int_{(n-1,n]} \sigma_s^2 ds \approx \sigma_{n-1}^2 , \quad n \in \mathbb{N} .$$

As we observe G only at integer times we approximate

$$\sum_{n-1<s\le n} \sigma_s^2 (\Delta L_s)^2 \approx (G_n - G_{n-1})^2 = (G_n^{(1)})^2 , \quad n \in \mathbb{N} .$$

An estimate of the volatility process $(\sigma_t^2)_{t\ge 0}$ can therefore be calculated recursively by

$$\widehat{\sigma}_n^2 = \widehat{\beta} + (1 - \widehat{\eta})\widehat{\sigma}_{n-1}^2 + \widehat{\varphi}\,(G_n^{(1)})^2 , \quad n \in \mathbb{N} . \tag{3.2}$$

Note that together with $G_n^{(1)} = \widehat{\sigma}_{n-1}\epsilon_n$, ϵ_n i.i.d $\sim (0,1)$, $n \in \mathbb{N}$, this defines a discrete time GARCH(1,1) model and we have to require that $0 < \eta < 1$. The estimator (3.2) is plotted in Figure 2 together with the theoretical $(\sigma_t^2)_{t\ge 0}$ for one simulation.

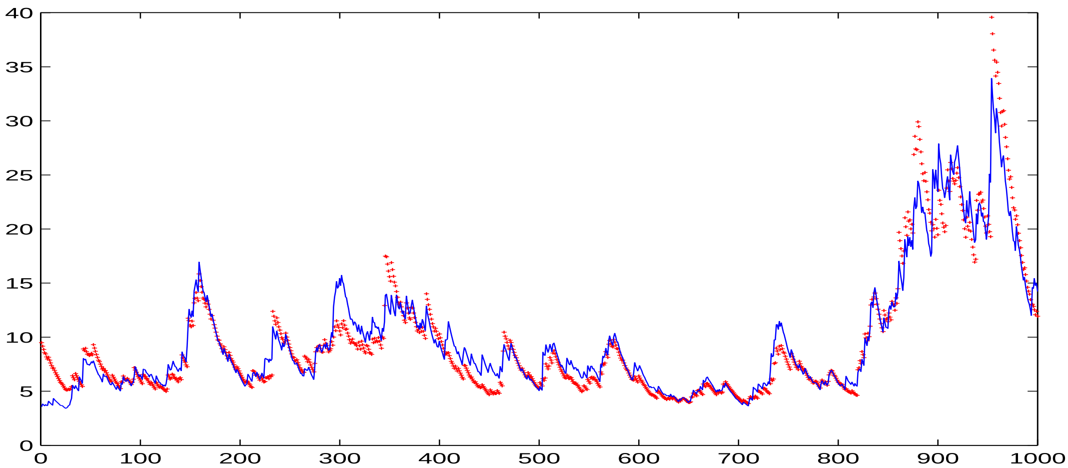

Figure 2: Sample paths of σ_t^2 (solid line) and $\widehat{\sigma}_t^2$ (+) of one simulation.

In this section we investigate the goodness of fit of our estimation method by a residual analysis. The estimated residuals are given by $G_n^{(1)}/\widehat{\sigma}_{n-1}$ for $n = 1, \dots, N$. Since we assumed a symmetric jump distribution with zero mean, trivally the residuals should be symmetric around zero and their mean should be close to zero. Furthermore, if the volatility has been estimated correctly, we expect the standard deviation to be close to 1. Consequently, we estimated mean, bias, MSE, MAE and the corresponding standard deviations for the mean, the standard deviation and the skewness of the residuals $G_n^{(1)}/\widehat{\sigma}_{n-1}$ based on 1 000 simulations. The results are reported in Table 3.3 and indicate a reasonable fit.

	mean$(G_n^{(1)}/\widehat{\sigma}_{n-1})$	std$(G_n^{(1)}/\widehat{\sigma}_{n-1})$	skewness$(G_n^{(1)}/\widehat{\sigma}_{n-1})$
Mean	0.0006 (0.0006)	1.0118 (0.0003)	0.0028 (0.0047)
Bias	0.0006 (0.0006)	0.0118 (0.0003)	0.0028 (0.0047)
MSE	0.0003 (0.1e-4)	0.0002 (0.7e-5)	0.0224 (0.0011)
MAE	0.0147 (0.0003)	0.0127 (0.0002)	0.1176 (0.0029)

Table 3.3: Estimated mean, bias, MSE and MAE for the mean, standard deviation and skewness of the residuals with corresponding estimated standard deviations in brackets.

For one simulation we calculated a smoothed histogram for the residuals. The corresponding plot together with the fitted standard normal distribution can be seen on the left in Figure 3. For the same simulation the estimated autocorrelation function of the squared residuals $(G_n^{(1)})^2/\widehat{\sigma}_{n-1}^2$ can be seen on the right in Figure 3.

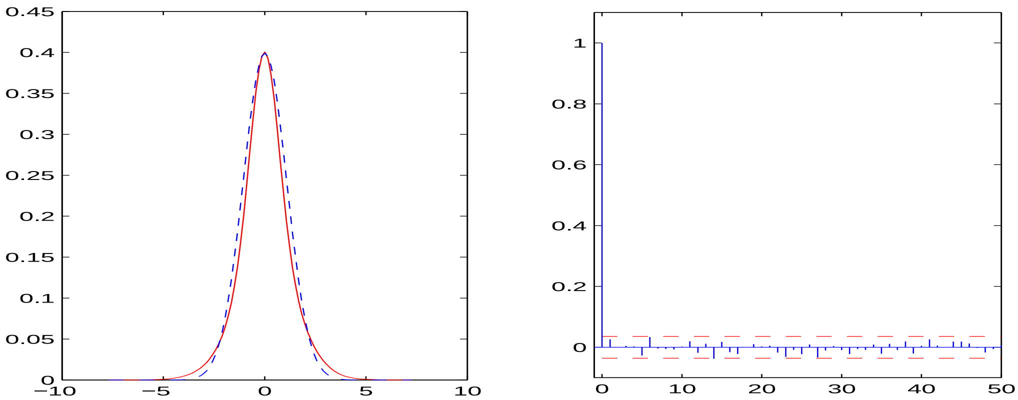

Figure 3: *Left:* Smoothed histogram of the residuals $G_n^{(1)}/\widehat{\sigma}_{n-1}$ for one simulation (solid line), together with the density of the standard normal distribution (dotted line). *Right:* Sample autocorrelation function of the squared residuals $(G_n^{(1)})^2/\widehat{\sigma}_{n-1}^2$ for the same simulation.

Acknowledgement

We would like to thank Thierry Jeantheau for his advice and interest in this work; in particular, MZ profited immensely from various discussions. CK thanks Jean Jacod for a lecture on stochastic analysis. SH and AL gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft, SH through the "Sonderforschungsbereich 386" and AL for a research grant number Li 1026/2-1.

Appendix

A Calculating the moments

Proof of Proposition 1.1. Since L has finite variance and zero mean, it is a square integrable martingale. Further, $\Psi(1) < 0$ implies $\mathbb{E}(\sigma_t^2) = \frac{\beta}{|\Psi(1)|} < \infty$ by (1.11), and it follows easily from the properties of the stochastic integral that

$$\mathbb{E}(G_t^2) = \mathbb{E}[G,G]_t = \mathbb{E}\int_0^t \sigma_s^2\, d[L,L]_s = \mathbb{E}[L,L]_1 \int_0^t \mathbb{E}(\sigma_s^2)\, ds\,,$$

giving that $\mathbb{E}(G_t^2)$ is finite and has the form specified in (1.8). The remaining equations in (1.8) are shown as in Proposition 5.1 of [6].

Suppose that $\varphi > 0$, $\mathbb{E}(L_1^4) < \infty$ and $\Psi(2) < 0$. Then $\mathbb{E}(G_t^4)$ is finite by the Burkholder-Davis-Gundy inequality, cf. Protter [11], p. 222, since

$$\mathbb{E}\left([G,G]_t^2\right) = \mathbb{E}\left(\int_0^t \sigma_s^2\, d[L,L]_s\right)^2$$

is finite as a consequence of $\mathbb{E}(\sigma_t^4) < \infty$ and $\mathbb{E}(L_1^4) < \infty$.

Now suppose additionally that $\int_{\mathbb{R}} x^3\, \nu_L(dx) = 0$. To calculate the value of $\mathbb{E}(G_t^4)$, observe that by integration by parts,

$$G_t^2 = 2\int_0^t G_{s-}\, dG_s + [G,G]_t = 2\int_0^t G_{s-}\sigma_s\, dL_s + \int_0^t \sigma_s^2\, d[L,L]_s, \tag{A.1}$$

$$\begin{aligned} G_t^4 &= 2\int_0^t G_{s-}^2\, dG_s^2 + [G^2,G^2]_t \\ &= 4\int_0^t G_{s-}^3\sigma_s\, dL_s + 2\int_0^t G_{s-}^2\sigma_s^2\, d[L,L]_s \\ &\quad +4\int_0^t G_{s-}^2\sigma_s^2\, d[L,L]_s + \int_0^t \sigma_s^4\, d\big[[L,L],[L,L]\big]_s \\ &\quad +4\int_0^t G_{s-}\sigma_s^3\, d\big[[L,L],L\big]_s. \end{aligned} \tag{A.2}$$

Taking expectations in (A.2), the first and the last summand vanish due to the assumptions $\mathbb{E}L_1 = 0$ and $\int_{\mathbb{R}} x^3\,\nu_L(dx) = 0$, respectively, so that

$$\mathbb{E}(G_t^4) = 6\mathbb{E}(L_1^2)\int_0^t \mathbb{E}(G_{s-}^2\sigma_s^2)\,ds + \int_{\mathbb{R}} x^4\,\nu_L(dx)\int_0^t \mathbb{E}(\sigma_s^4)\,ds. \tag{A.3}$$

The expression $\mathbb{E}(G_{s-}^2\sigma_s^2)$ was already calculated in the proof of Proposition 5.1 in [6], however, under additional assumptions which required in particular bounded variation of L. The following calculations do not require these restrictions.

Let $Y_t := \int_0^t G_{s-}\sigma_s\,dL_s$, $t \geq 0$. Then $\mathbb{E}(Y_t) = 0$ for all $t \geq 0$, and integration by parts and substituting from (1.3) give

$$\begin{aligned}
Y_t\sigma_{t+}^2 &= \int_0^t Y_{s-}\,d\sigma_{s+}^2 + \int_0^t \sigma_s^2\,dY_s + [\sigma_+^2, Y]_t \\
&= \int_0^t Y_{s-}(\beta - \eta\sigma_s^2)\,ds + \int_0^t Y_{s-}\varphi\sigma_s^2\,d[L,L]_t^{(d)} \\
&\quad + \int_0^t \sigma_s^3 G_{s-}\,dL_s + \left[\int_0^{\cdot}(\beta - \eta\sigma_s^2)\,ds + \int_0^{\cdot}\varphi\sigma_s^2\,d[L,L]_s^{(d)}, \int_0^{\cdot} G_{s-}\sigma_s\,dL_s\right]_t.
\end{aligned}$$

Taking expectations gives

$$\begin{aligned}
E(Y_t\sigma_{t+}^2) &= \big(\varphi(\mathbb{E}(L_1^2) - \tau_L^2) - \eta\big)\int_0^t E(Y_{s-}\sigma_s^2)\,ds + \mathbb{E}\int_0^t \varphi\sigma_s^3 G_{s-}\,d\sum_{0<u\leq s}(\Delta L_u)^3 \\
&= \big(\varphi(\mathbb{E}(L_1^2) - \tau_L^2) - \eta\big)\int_0^t E(Y_s\sigma_{s+}^2)\,ds,
\end{aligned}$$

where we used that $\int_{\mathbb{R}} x^3\,\nu_L(dx) = 0$ and that $Y_{s-}\sigma_s^2 = Y_s\sigma_{s+}^2$ almost surely for fixed s. Solving this integral equation and using that $Y_0 = 0$ implies $E(Y_0\sigma_{0+}^2) = 0$, it follows that $E(Y_t\sigma_{t+}^2) = 0$ for all $t \geq 0$. Substituting

$$\int_0^t \sigma_s^2\,d[L,L]_s = \int_0^t \sigma_s^2\tau_L^2\,ds + \varphi^{-1}\left(\sigma_{t+}^2 - \sigma_0^2 - \int_0^t(\beta - \eta\sigma_s^2)\,ds\right)$$

from (1.3), equations (A.1) and (1.12) now give

$$\begin{aligned}
\mathbb{E}(G_t^2\sigma_{t+}^2) &= \mathbb{E}\left(\sigma_{t+}^2\int_0^t \sigma_s^2\,d[L,L]_s\right) \\
&= (\tau_L^2 + \varphi^{-1}\eta)\int_0^t \mathbb{E}(\sigma_t^2\sigma_s^2)\,ds + \varphi^{-1}\mathbb{E}(\sigma_t^4) - \varphi^{-1}\mathbb{E}(\sigma_t^2\sigma_0^2) - \varphi^{-1}\beta\mathbb{E}(\sigma_t^2)t \\
&= (\tau_L^2 + \varphi^{-1}\eta)\mathbb{V}\mathrm{ar}\,(\sigma_0^2)\frac{1 - e^{-t|\Psi(1)|}}{|\Psi(1)|} + \varphi^{-1}\mathbb{V}\mathrm{ar}\,(\sigma_0^2)(1 - e^{-t|\Psi(1)|}) \\
&\quad + \big((\tau_L^2 + \eta\varphi^{-1})(\mathbb{E}(\sigma_0^2))^2 - \beta\varphi^{-1}\mathbb{E}(\sigma_0^2)\big)\,t.
\end{aligned} \tag{A.4}$$

Using (1.11), (1.12) and $\Psi(1) = -\eta + \varphi\,(\mathbb{E}(L_1^2) - \tau_L^2)$ then leads to

$$\mathbb{E}(G_t^2\sigma_{t+}^2) = \frac{\beta^2}{\psi(1)^2}\left(\frac{2}{|\Psi(2)|} - \frac{1}{|\Psi(1)|}\right)(2\eta\varphi^{-1} + 2\tau_L^2 - \mathbb{E}(L_1^2))(1 - e^{-t|\Psi(1)|}) + \frac{\beta^2}{\Psi(1)^2}\mathbb{E}(L_1^2)t.$$

This then implies (1.9), where we used (A.3), (1.12) and the fact that

$$\int_{\mathbb{R}} x^4 \, \nu_L(dx) = \frac{\Psi(2) - 2\Psi(1)}{\varphi^2} \tag{A.5}$$

by (1.7).

For the autocorrelation of the squared increments, observe that by equation (5.4) of [6] we have

$$\mathbb{C}\mathrm{ov}\,((G_t^{(r)})^2, (G_{t+h}^{(r)})^2) = \left(\frac{e^{r|\Psi(1)|} - 1}{|\Psi(1)|} \right) \mathbb{E}(L_1^2) \mathbb{C}\mathrm{ov}\,(G_r^2, \sigma_r^2)\, e^{-h|\Psi(1)|} \tag{A.6}$$

(in [6] this was stated under the additional assumption that L is a quadratic pure jump process (i.e. $\tau_L^2 = 0$), but it can be seen that the proof given there holds true also for L having a Brownian motion component). This then implies (1.10) by (A.4), (1.8) and (1.11). $\square$

B Ergodicity of $(G_{rn}^{(r)})_{n\in\mathbb{N}}$

Proof of Theorem 2.3. For the ease of notation suppose that $r = 1$ and denote

$$Z_n := G_n^{(1)} = \int_{(n-1,n]} \sigma_s \, dL_s, \quad n \in \mathbb{N}.$$

We shall first show that the process $(Z_n)_{n\in\mathbb{N}}$ is mixing, i.e. for any Borel sets $\overline{U}, \overline{V}$ of $\mathbb{R}^{\mathbb{N}}$ it holds

$$\lim_{n\to\infty} P((Z_k)_{k\in\mathbb{N}} \in \overline{U}, (Z_{k+n})_{k\in\mathbb{N}} \in \overline{V}) = P((Z_k)_{k\in\mathbb{N}} \in \overline{U}) \; P((Z_k)_{k\in\mathbb{N}} \in \overline{V}). \tag{B.1}$$

Since the Borel sets in $\mathbb{R}^{\mathbb{N}}$ and the distribution of $(Z_k)_{k\in\mathbb{N}}$ are generated by cylinder sets, it suffices to prove (B.1) only for special cylinder sets. Thus, for $p, q \in \mathbb{N}$ define

$$\begin{aligned} U &= [u_1, u_1'] \times \ldots \times [u_p, u_p'], \\ V &= [v_1, v_1'] \times \ldots \times [v_q, v_q'], \end{aligned}$$

with $u_i \le u_i', v_j \le v_j'$. Then we have to show that

$$\begin{aligned} &\lim_{n\to\infty} P\left((Z_1, \ldots Z_p, Z_{1+n}, \ldots, Z_{q+n}) \in U \times V\right) \\ &\quad = P\left((Z_1, \ldots, Z_p) \in U\right) \; P\left((Z_1, \ldots, Z_q) \in V\right). \end{aligned} \tag{B.2}$$

Now define for $s \ge p$

$$A_s \;:=\; e^{-(X_{s-} - X_{p-})} \quad \text{and} \quad B_s \;:=\; \beta e^{-(X_{s-} - X_{p-})} \int_p^s e^{(X_r - X_{p-})} \, dr\,,$$

where $(X_s)_{s\geq 0}$ is the auxiliary process of (1.5). Then by the strong Markov property of Lévy processes and representation (1.6) we obtain

$$\sigma_s^2 = A_s\sigma_p^2 + B_s, \quad s \geq p,$$

and (A_s, B_s) is independent of $(L_t)_{0\leq t\leq p}$. With

$$C_s := \sigma_p^2 \frac{A_s}{\sigma_s + \sqrt{B_s}}$$

we have the representation

$$\sigma_s = \sqrt{B_s} + (\sigma_s - \sqrt{B_s}) = \sqrt{B_s} + \frac{A_s\,\sigma_p^2}{\sigma_s + \sqrt{B_s}} = \sqrt{B_s} + C_s.$$

Define

$$Y_n := \int_{(n-1,n]} \sqrt{B_s}\,dL_s \quad \text{and} \quad R_n := \int_{(n-1,n]} C_s\,dL_s\,, \quad n-1 \geq p\,.$$

We first show that $R_n \xrightarrow{P} 0$ as $n \to \infty$. Since

$$R_n = \sigma_p^2 \int_{(n-1,n]} \frac{A_s}{\sigma_s + \sqrt{B_s}}\,dL_s, \quad n-1 \geq p,$$

it follows from the facts that σ_s^2 is bounded from below by β/η (see KLM [7], Prop 3.4) and that $\mathbb{E}(e^{-sX_t}) = e^{t\Psi(s)}$ that

$$\begin{aligned}
&\mathbb{E}\left|\int_{(n-1,n]} \frac{A_s}{\sigma_s + \sqrt{B_s}}\,dL_s\right| \\
&\leq\ |\mathbb{E}(L_1)| \int_{(n-1,n]} \mathbb{E}\left(\frac{A_s}{\sigma_s + \sqrt{B_s}}\right) ds + \left(\mathbb{E}([L,L]_1) \int_{(n-1,n]} \mathbb{E}\left(\frac{A_s}{\sigma_s + \sqrt{B_s}}\right)^2 ds\right)^{1/2} \\
&\leq\ |\mathbb{E}(L_1)|\sqrt{\eta/\beta} \int_{n-1}^{n} e^{(s-p)\Psi(1)}\,ds + \left(\mathbb{E}([L,L]_1)\eta/\beta \int_{n-1}^{n} e^{(s-p)\Psi(2)}\,ds\right)^{1/2}
\end{aligned}$$

which converges to 0 as $n \to \infty$ since $\Psi(2) < 0$, so that $R_n \xrightarrow{P} 0$ as $n \to \infty$. Let $\varepsilon > 0$. Then, since $Z_n = Y_n + R_n$, we can estimate

$$\begin{aligned}
&P\big((Z_1,\ldots,Z_p) \in U, Y_{1+n} \in [v_1+\varepsilon, v_1'-\varepsilon], \ldots, Y_{q+n} \in [v_q+\varepsilon, v_q'-\varepsilon]\big) \\
&\qquad -P\left(\exists\, j \in \{1,\ldots,q\} : |R_{j+n}| \geq \varepsilon\right) \\
&\leq\ P\left((Z_1,\ldots,Z_p) \in U,\ (Z_{1+n},\ldots,Z_{q+n}) \in V\right) \\
&\leq\ P\big((Z_1,\ldots,Z_p) \in U, Y_{1+n} \in [v_1-\varepsilon, v_1'+\varepsilon], \ldots, Y_{q+n} \in [v_q-\varepsilon, v_q'+\varepsilon]\big) \\
&\qquad +P\left(\exists\, j \in \{1,\ldots,q\} : |R_{j+n}| \geq \varepsilon\right).
\end{aligned}$$

Recalling that $R_n \xrightarrow{P} 0$ as $n \to \infty$ and that Y_n is independent of $(L_s)_{0\le s\le p}$ for $n \ge p+1$ we obtain

$$\begin{aligned} & P\left((Z_1,\dots,Z_p)\in U\right) \liminf_{n\to\infty} P\left(Y_{1+n}\in[v_1+\varepsilon, v_1'-\varepsilon],\dots,Y_{q+n}\in[v_q+\varepsilon,v_q'-\varepsilon]\right) \\ & \quad \le \liminf_{n\to\infty} P\left((Z_1,\dots,Z_p)\in U, (Z_{1+n},\dots,Z_{q+n})\in V\right) \\ & \quad \le \limsup_{n\to\infty} P\left((Z_1,\dots,Z_p)\in U, (Z_{1+n},\dots,Z_{q+n})\in V\right) \\ & \quad \le P\left((Z_1,\dots,Z_p)\in U\right) \\ & \qquad \times \limsup_{n\to\infty} P\left(Y_{1+n}\in[v_1-\varepsilon, v_1'+\varepsilon],\dots,Y_{q+n}\in[v_q-\varepsilon,v_q'+\varepsilon]\right). \end{aligned}$$

Similarly, we obtain

$$\begin{aligned} & \limsup_{n\to\infty} P\left(Y_{1+n}\in[v_1-\varepsilon, v_1'+\varepsilon],\dots,Y_{q+n}\in[v_q-\varepsilon,v_q'+\varepsilon]\right) \\ & \quad \le \limsup_{n\to\infty} P\left(Z_{1+n}\in[v_1-2\varepsilon, v_1'+2\varepsilon],\dots,Z_{q+n}\in[v_q-2\varepsilon,v_q'+2\varepsilon]\right) \\ & \quad = P\left(Z_1\in[v_1-2\varepsilon, v_1'+2\varepsilon],\dots,Z_q\in[v_q-2\varepsilon,v_q'+2\varepsilon]\right) \end{aligned}$$

and a similar estimate for $\liminf$. Now (B.2) follows from

$$\begin{aligned} & \lim_{\varepsilon\to 0} P\left(Z_1\in[v_1-2\varepsilon, v_1'+2\varepsilon],\dots,Z_q\in[v_q-2\varepsilon,v_q'+2\varepsilon]\right) \\ & \quad = P\left(Z_1\in[v_1,v_1'],\dots,Z_q\in[v_q,v_q']\right). \end{aligned}$$

Thus we have proved that the process $(Z_n)_{n\in\mathbb{N}}$ is mixing. By Theorem 3.2.6 of Ash and Gardner [1] this implies that $(Z_n)_{n\in\mathbb{N}}$ is ergodic, giving the result. $\square$

References

[1] Ash, R.B. and Gardner, M.F. (1975) *Topics in Stochastic Processes.* Academic Press, New York.

[2] Brockwell, P.J. and Davis, R.A. (1987) *Time Series: Theory and Methods.* Springer, New York.

[3] Drost, F.C. and Werker, B.J.M. (1996) Closing the GARCH gap: continuous time GARCH modelling. *J. Econometrics* **74**, 31-57.

[4] Fasen, V., Klüppelberg, C. and Lindner, A. (2004) Extremal behavior of stochastic volatility models. In: Grossinho, M.d.R., Shiryaev, A.N., Esquivel, M. and Oliviera, P.E. (Eds.) *Stochastic Finance.* Springer, New York. To appear.

[5] Hannan, E.J. (1970) *Multiple Time Series.* Wiley, New York.

[6] Klüppelberg, C., Lindner, A. and Maller, R. (2004) A continuous time GARCH process driven by a Lévy process: stationarity and second order behaviour. *J. Appl. Probab.***41** (3), 601-622.

[7] Klüppelberg, C., Lindner, A. and Maller, R. (2005) Continuous time volatility modelling: COGARCH versus Ornstein-Uhlenbeck models, in: Kabanov, Y., Lipster, R. and Stoyanov, J. (Eds.), *From Stochastic Calculus to Mathematical Finance. The Shiryaev Festschrift*. Springer. To appear.

[8] Mikosch, T. and Starica, C. (2000) Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. *Ann. Statist.* **28**, 1427-1451.

[9] Müller, G. (2005). MCMC estimation of the COGARCH(1,1) model. In preparation.

[10] Nelson, D.B. (1990) ARCH models as diffusion approximations. *J. Econometrics* **45**, 7-38.

[11] Protter, P.E. (2004) *Stochastic Integration and Differential Equations*, second ed., Springer, Berlin.

[12] Sato, K.-I. (1999) *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge.

[13] Seber, G.A.F. and Wild, C.J. (1989) *Nonlinear Regression.* Wiley, New York.

[14] Zapp, M. (2004) *Schätzverfahren in einem neuen* COGARCH-*Modell.* Diploma Thesis. Munich University of Technology. Available at
http://www-m4.ma.tum.de/Diplarb