LUND UNIVERSITY
Department of Statistics
Supervisor: Prof. Krzysztof Podgórski

Master Thesis
15 ECTS
Spring semester 2019

Modeling asymmetry in volatility response

non-Gaussian innovations approach

Ludvig Göransson

Abstract

This thesis is an explorative note on the non-Gaussian innovations of the volatility process. More specifically, the thesis investigates if the decomposition of the Standard Classical Laplace (SCL) distribution to a difference of two exponential is a valid alternative to modelling the asymmetric volatility processes, taking volatility clustering, the leverage effect and asymmetric response in volatility into account. We derive the probability density functions (pdf) and log-likelihood functions for both asymmetric and non-asymmetric case. The pdf exhibit characteristics suitable for modelling the leverage effect, positive kurtosis and fat tails often observed in financial returns, for both the symmetric and asymmetric case. We derive the likelihood function in the non-asymmetric case with regards to the relevant parameters. The likelihood function evaluated concerning past volatility presents a compact solution with the persistence of volatility, whereas, for the other two parameters, the result is complicated. In the asymmetric case, the log-likelihood function is maximized concerning the parameters of interest through simulation. The convergence rate is fast for the AR and lagged volatility parameters, but slower for the asymmetric parameter, which is hard to observe in this setting.

Keywords: ARCH, GARCH, APARCH, Asymmetric GARCH, non-Gaussian innovations, Laplace distribution, Leverage effect, Stylized facts, Volatility process.

Contents

1 Introduction

In the field of finance, understanding the components of risk assessment can potentially generate substantial financial gains. The research in modelling risk through volatility has increased massively over the last decades, specifically the research in 'stylized facts' of returns. 'Stylized facts' refers to properties often observed in financial gains, where, in the context of this thesis, clustering of volatility, the leverage effect (LE) and leptokurtic behaviour in financial returns are of particular interest. The main focus of this thesis lies in evaluating if the Standard Classical Laplace (SCL) distribution is a valid contender for modelling these three stylized facts mentioned in the volatility process. The critique that these non-normal characteristics have been overlooked in the modelling of volatility and returns by only using Gaussian distributions are vast, Javed and Podgórski (2014) and therein, although the features of financial returns are well documented, see Cryer and Chan (2008) and Hol (2010). With this in mind, the search for generic models using the non-Gaussian distribution for modelling financial returns are of high relevance, and this paper aims at creating a building stone in this field.

Autoregressive Conditional Heteroscedasticity (ARCH) - model is the most famous model to evaluate clusters of volatility. This model is the foundation for modelling the other two 'stylized facts' of interest, the LE and leptokurtosis. There are alternations in many forms proposed since Robert Engle in the late 1970s started to model volatility by conditioning on the variance/standard deviation of past error terms. The majority of the research does not focus on alternating the distribution meant to capture the 'stylized facts' of interest but by varying the model in other ways. This thesis aims at exploring the area of changing distribution with characteristic similar to the 'stylized facts' of interest. We will utilize a generalization and extension of the traditional ARCH - model introduced by Javed and Podgórski (2014), which emphasize the asymmetry often observed in returns in a combination of changing the Gaussian distribution for the SCL distribution.

The Asymmetric Power ARCH model (APARCH), as previously mentioned, has the purpose of emphasizing asymmetry and fat tails, Ding et al. (1993). The power parameter aims at including fat tail behaviour, whereas the asymmetric setting aims at covering the LE. They evaluate the model in Javed and Podgórski (2016), where they argue that the power parameter has no impact. Therefore, a distribution with fatter tails is of interest.

To conclude, this thesis aims at evaluating if a specific non-Gaussian distribution that exhibits the characteristics desired to assess volatility. We utilize the APARCH without taking to much notice in the Power parameter. Next part presents an illustrative example for clarifying an issue with modelling 'stylized facts' with a focus on the previous mentioned LE.

The modelling of asymmetric response in volatility, i.e. the LE, is a delicate matter and should be treated accordingly. It is easy to claim no asymmetry present depending on which method applied. By comparing the most straightforward way of estimating the LE with historical volatility we can see the fragility

$$LE_t = corr\Big(R_t, {R_{t+1}}^2\Big), \tag{1}$$

$$LE_{t,\hat{\sigma}^2} = corr\Big(r_t, \hat{\sigma}^2\Big), \tag{2}$$

where (1) accounts for the correlation between the returns at time t and the squared returns at time $t+1$. This procedure shows the most primitive form of volatility, taking all variation into account. However, this method is not very refined, and it gives a rough measurement of the LE. In (2) $\hat{\sigma}^2$ refers to the variance of the rolling window and r_t, is the logarithmic returns at time t. A negative correlation in (1) or (2) indicates a present LE. Through a rolling window of k number of days, we estimate historical volatility in (2). The historical volatility is more complex than the first volatility measurement, although seen as one of the simplest ways of estimating this effect, and therefore, should be more refined and give a better indication of the LE. The problem with the historical volatility is the rolling window of k days which includes a smoothing effect hiding the LE as can be seen in Table 1.

Leverage Effect

Index	LE_t	$LE_{t,\hat{\sigma}^2}$
S&P 500	-0.1080	0.0392
S&P 1200 Global	-0.0732	-0.0285
S&P 350 Europe	-0.1186	0.0632
S&P GSCI	-0.0295	0.0015
DJIA	-0.1158	0.0560

Table 1: LE for 5 different indices evaluated with primitive and historical volatility. The historical volatility is of a 2 days rolling window.

The LE is present using the more primitive estimate of LE but smoothed out by the historical volatility measure, even for the two days window. The data displays characteristics of LE, but the more advanced measurement is too unrefined to measure the magnitude. Therefore, a less coarse expression for modelling the LE is needed to answer this question.

The comparison of two LE methods is a simple example of the complexity of capturing 'stylized facts' in a proper way and proves that evaluating volatility is a delicate matter.

1.1 Predecessors to the asymmetric power models

Engle (1982) elaborates on the topic of improving forecasts of financial returns by conditioning on the variance of historical error terms. The main argument is that the variance of the residuals is changing over time, which suggests that a suitable model for the data should have a time-varying volatility structure to ensure that the process is stochastic. The resulting ARCH - model was revolutionary as an alternative to previous models, assuming constant variance in the error terms. Robert Engle was later awarded the Nobel Price in Economy, 2003, for his findings regarding the ARCH. The alternations to the ARCH are many. The most famous and frequently practised is the generalized autoregressive conditional heteroscedasticity (GARCH) - model in Bollerslev (1986). The generalization of the ARCH permits more flexibility in the conditional second moments, similar to the extension of the simple autoregressive (AR) process to the autoregressive moving average (ARMA) process. The GARCH process includes the lagged conditional volatility to the ARCH, extending the conditional variance from a linear combination of past sample error terms to linear combination of past sample error terms with an additional lag structure of past volatility. The GARCH is today industry standard, and the alternations are many.

One generalization of the GARCH-model is the Asymmetric Power ARCH (APARCH). The APARCH was, as previously mentioned, introduced in Ding et al. (1993) to take the LE and asymmetric behaviour in financial returns into account as well. The LE discovered by Black (1976) implies asymmetric response in volatility for 'negative' and 'positive' shocks in past returns, i.e. 'negative' and 'positive' shocks of the same amplitude tends to have a different impact on historical volatility. In financial data, the magnitude of decrease in returns is often more profound in comparison to the increase in returns, which suggests that volatility has a negative correlation to returns. Ding et al. (1993) follows the argument and includes this effect when modeling financial returns stated by Nelson (1991) and Engle and Ng (1993). The outcome was a Box-Cox transformation of the conditional standard deviation and absolute asymmetric residuals, under the name of the APARCH. To include the asymmetric effect, 'positive' and 'negative' shocks were given different weights and used to take this 'stylized facts' into account as will be further elaborated in section 2.

Worth mentioning in the APARCH context are two other asymmetric volatility models presented in Nelson (1991) and Glosten et al. (1993). In Nelson (1991), they introduce the Exponential-GARCH (EGARCH). The model is similar to the APARCH but takes the natural logarithm of the conditional volatility estimate instead of the power parameter. In Glosten et al. (1993) they present several alternations of the GARCH where the most popular, the GJR-ARCH model, includes asymmetry by incorporating an indicator function on the squared past error terms to the traditional GARCH. In practice, the volatility measure increase as the indicator function takes the value 1 when $e_t < 0$ otherwise 0. The models are evaluated in a Japanese context and finds that the GJR-GARCH captures the conditional volatility the best, Glosten et al. (1993). They argue that the EGARCH overestimates the effect of 'negative' news, in favour of the GJR-GARCH.

All models mentioned above assume standard Gaussian past noise processes. More recent studies drive the specifications of the models even further and focus on the distribution of the residuals. The argument behind most of the studies is the characteristics of the normal distribution, not taking fat tails and positive kurtosis into account. The APARCH aims at, as will be further elaborated in section 2, emphasizing the asymmetric behaviour often observed in financial returns by the power parameter of the model but not focusing on either the kurtosis or heavy tails.

Documentation of fat tails in financial returns are abundant, which is why using a distribution with fatter tails is of interest assessing volatility measures, Hol (2010). In Javed and Podgórski (2014) the evaluate the APARCH model with generalized asymmetric Laplace (GAL) error terms and compares it to the traditional APARCH with standard normal past noise and the results from APARCH with Normal Inverse Gaussian (NIG) in Jensen and Lunde (2001). The conclusion state that the APARCH is struggling to detect 'stylized facts' unless non-Gaussian error terms are applied. In Javed and Podgórski (2016) the findings from Javed and Podgórski (2014) are further elaborated. They present that a volatility process, including a Non-Gaussian distribution, can take LE and positive kurtosis into better account than previous research. They also ground their findings mathematically, making their model more robust.

1.2 Goal

The main scope of this thesis is to evaluate if an Asymmetric GARCH model with standard classical Laplace (SCL) noise can generate credible results to evaluate simulated volatility, taking 'stylized facts' such as the LE, volatility clustering and leptokurtosis into account.

The purpose of this thesis is threefold. The first follows from the argument of the authors in Javed and Podgórski (2014) and Javed and Podgórski (2016), proving the increased accuracy in modelling volatility and returns from a prior mathematical validation of the model before applying it to real data. They suggest that extensive theoretical research using non-Gaussian error terms in an asymmetric setting of the volatility process is needed. In this thesis, a specification of the gamma random variables in the underlying processes generates the SCL distribution, which gives us a non-Gaussian distribution. The second purpose concerns the properties of the SCL distribution. The first property is that it attains heavy tail and excess kurtosis behaviour, often observed in asset returns - secondly, the density and maximum likelihood functions are known, which makes it a suitable candidate for financial data. The third reason emphasizes the relatively limited amount of alternative models trying to capture asymmetric response in volatility using non-Gaussian innovations. The research question evaluated is:

- Is a modification of the volatility process, utilizing the decomposition of the Laplace distribution to a difference of two exponential random variables, a valid alternative to modelling asymmetric volatility processes?

2 Principles of non-Gaussian innovations

This section aims to clarify the difference between different volatility models. The asymmetric model originates from the GARCH, which is a suitable initial point for the theory of modelling volatility. Subsequent, we present the standard APARCH and last are the GAL innovations elucidated in-depth and applied in an APARCH context. Furthermore, the transition from GAL innovations to the SCL innovations is outlined and at last, which is the model of primary interest.

The GARCH-models is given by the formula

$$\begin{aligned} y_t &= m + ay_{t-1} + \sigma\rho_t e_t, \\ \rho_t &= \alpha_0 + \alpha e_{t-1}^2 + \beta\rho_{t-1} \end{aligned} \tag{3}$$

where the last term of the second equation refers to the generalization of the ARCH, including the lagged volatility. The first part of (3), (4), (7), (9) and (12) follow a similar pattern where y_t is the log normal returns of the asset, m represent the location of the returns, a is the autoregressive parameter and σ represents scaling of volatility. It is the constituents of ρ_t that alters for each of the models. In (3), ρ_t is supposed to account for the long term memory and leptokurtic volatility of the error term, where e_t is a standard normal random variable. A higher value of α in ρ_t represents a stronger correlation with previous observations, whereas a lower value implies the opposite, and zero indicates a random walk process. Finally, β shows the endurance of volatility, where a higher value gives a stronger influence from the previous location in volatility, and vice versa for a lower value.

The APARCH in Ding et al. (1993) is declared below in the second part of (4) where the notations of the APARCH are subsequent to Javed and Podgórski (2014) and given by the formula

$$\begin{aligned} y_t &= m + ay_{t-1} + \sigma\rho_t e_t, \\ \rho_t^\delta &= \alpha_0 + \alpha\rho_{t-1}^\delta[(1-\theta)^\delta|e_{t-1}^+|^\delta + (1+\theta)^\delta|e_{t-1}^-|^\delta] + \beta\rho_{t-1}^\delta \end{aligned} \tag{4}$$

where ρ_t^δ represents the long term memory, leptokurtic and asymmetric volatility of the noise. The error term is, as in (3), a standard normal random variable. The autoregressive parameters, α and β follow the structure from (3).

The asymmetric part of the model is given by θ. The leverage parameter creates the differences in weights given 'positive' and 'negative' shocks, where the 'negative' shocks are $(1+\theta)$ and the 'positive' shocks are $(1-\theta)$. Theoretically, θ can take on any value in the interval $[-1, 1]$. Although, in practice, a positive θ is the only parameter of interest since the LE is negative. Hence, we refer to the asymmetric case as $\theta \geq 0$ in all cases except when we derive the pdf. Following this intuition, the $e_{t-1}^{-\delta}$ and $e_{t-1}^{+\delta}$ represents the random variables that belong to the 'negative' and 'positive' shocks, respectively. These independent identically distributed (i.i.d) random variables are standard normal, Ding et al. (1993). The power parameter in the APARCH model is δ, which is supposed to emphasise leptokurtosis behaviour in the Gaussian distribution. Important notice about the power parameter δ and the location parameter m and their realized effects. In Javed and Podgórski (2014) and Javed and Podgórski (2016) they found no significant difference in putting $\delta = 1$ or $\delta = 2$, arguing that the Box-Cox transformation motivating the inclusion of δ in Ding et al. (1993) is misplaced. The Box-Cox transformation is known for reducing non-normality in the data. However, in this context, the transformation is meant to increase non-normality by creating fatter tails of the standard normal error term. In Javed and Podgórski (2014) and Javed and Podgórski (2016), they highlight that the power parameter, rather than being a transformation parameter taking non-normality into account, is included in the ρ_t process where the effect of δ is unnoticed. The main point is that the power parameter in the APARCH model is somewhat enigmatic, and the choice of $\delta = 1$ or $\delta = 2$ makes no difference in modelling volatility. This implies that modeling volatility with the standard deviation, $\delta = 1$, is impartial to modeling volatility with the variance, $\delta = 2$. Furthermore, they also found a misinterpretation of the location parameter, m. Instead of being a location parameter, it takes the role of a scaling parameter increasing volatility. As a result, the volatility process has problems reflecting small values, generating a gap of values, which the volatility model can not attain. To deal with this controversy, they included a new location parameter, making it possible for the volatility model to assume any value. For simplicity, we will assume that m is a location parameter since this paper does not evaluate the full volatility model.

2.1 Volatility modelling GAL innovations

The next step for the inquiry of asymmetric response in volatility continues in Javed and Podgórski (2014), where they emphasize two specific cases of non-Gaussian innovations possessing the desired features, Generalized Laplace Errors (GAL) and Normal Inverse Gaussian (NIG). The distributions are closed on convolution and, therefore, vital since it allows for time-frequency sampling. This paper mainly focuses on the GAL distribution, which is elaborated below.

Javed and Podgórski (2014) unravels their volatility model with two propositions to obtain GAL innovations, starting with the generalized hyperbolic (GH) distribution, which is given by the formula

$$X = \sigma\gamma^{\frac{1}{2}}Z + \mu\gamma, \tag{5}$$

where X is a normal variance mean mixture random variable and Z is a normal distributed random variable with independent non-negative mixing variable γ. In other words, (5) is the generalized inverse Gaussian (GIG) distribution and consequently, γ are equivalent to $x^{\tau-2}e^{-ax-b/x}$. The second term of the equation constitutes the asymmetry in the shock, where μ is initially set to zero since the shocks are assumed to have the same distribution. The distribution is of importance since the paper aims at investigating asymmetry in response to 'negative' and 'positive' shocks in volatility, not asymmetry in the shocks.

In the second alternation, the GAL distribution is obtained by the difference of two gamma random variables following

$$\sigma\left(\kappa\gamma_t^+ - \frac{\gamma_t^-}{\kappa}\right), \tag{6}$$

where γ_t^+ and γ_t^- are two identically distributed and independent gamma random variables with shape and scale parameter $(1/\tau)$ and (τ) respectively. The τ parameter accounts for the tail behaviour of the distribution. A small τ implies high kurtosis (heavy tails), while the situation is the opposite for a more substantial τ. The structure of γ_t^+ and γ_t^- are the same as for the Gaussian error terms, γ_t^- and γ_t^+ belong to the 'negative' and 'positive' shocks, respectively. The scaling of 'negative' and 'positive' shocks is given by κ, equivalent to μ in (5), and is set equal to 1 with similar reasoning. A $\kappa = 1$ implies that the shocks follow the same distribution.

To apply the desired features of the GAL distribution modeling volatility, the standard normal errors in (4) are replaced by the GAL innovations from (5) and (6). Incorporating the GAL distributed noise from (5) gives

$$\begin{aligned} y_t &= m + a y_{t-1} + \rho_t(\sigma \gamma_t^{\frac{1}{2}} Z_t + \mu \gamma_t), \\ \rho_t^\delta &= 1 + \alpha \rho_{t-1}^\delta \gamma_t^{\frac{\delta}{2}} [(1-\theta)^\delta Z_{t-1}^{+\delta} + (1+\theta)^\delta Z_{t-1}^{-\delta}] + \beta \rho_{t-1}^\delta, \end{aligned} \tag{7}$$

where $\gamma_t^{\frac{\delta}{2}}$ is an i.i.d gamma random variable and the structure follows (4). GAL distributed random variables can be obtained by $\gamma_t^{\frac{1}{2}} Z_t$, given that Z_t follows the normal distribution, which is why the random variable is included. The only difference, except the distribution, between the original APARCH and the APARCH, presented in Javed and Podgórski (2014), above is the intercept of the volatility process, α_0, which is set equal to 1. They also incorporated the recursive innovations in λ_t, creating the following relation

$$\begin{aligned} \rho_t^\delta &= 1 + \rho_{t-1}^\delta \lambda_{t-1}, \\ \lambda_t &= \alpha \gamma_t^{\frac{\delta}{2}} [(1-\theta)^\delta Z_{t-1}^{+\delta} + (1+\theta)^\delta Z_{t-1}^{-\delta}] + \beta, \end{aligned} \tag{8}$$

which concludes the four parameterized model of $\alpha, \beta, \sigma, \theta$ using $\gamma_t^{\frac{1}{2}} Z_t$ to attain the GAL distributed innovations. The previous included parameter α_0 is, as mentioned, put equal to 1, following the notation of Javed and Podgórski (2014) and Javed and Podgórski (2016).

Javed and Podgórski (2014) emphasize the computational advantages in obtaining GAL distribution by avoiding the Bessel function estimating past noise in (6), which gives the expression for volatility with a GAL process accordingly

$$\begin{aligned} y_t &= m + a y_{t-1} + \rho_t \sigma \left(\kappa \gamma_t^+ - \frac{\gamma_t^-}{\kappa} \right), \\ \rho_t^\delta &= 1 + \alpha \rho_{t-1}^\delta [(1-\theta)^\delta \gamma_{t-1}^{+\delta} + (1+\theta)^\delta \gamma_{t-1}^{-\delta}] + \beta \rho_{t-1}^\delta, \end{aligned} \tag{9}$$

where $\rho_t^\delta, \alpha, \beta$ and δ are interpreted subsequently to the intuition of the Gaussian APARCH model of (4). The sequence of γ_t^+ and γ_t^- are, as mentioned, i.i.d random variables with an expected value of 1. The volatility process is the second part of (4) and also simplified by $\alpha_0 = 1$.

Furthermore, the conditions for a strictly stationary process for the parameters in (9) are given by the inequality

$$(1-\theta)^\delta + (1+\theta)^\delta < \frac{1-\beta}{\alpha} \frac{\tau^\delta \Gamma(\tau)}{\Gamma(\tau+\delta)}. \tag{10}$$

For further elaboration of the GAL innovations, see Javed and Podgórski (2014) and Javed and Podgórski (2016).

2.2 Volatility modeling with Standard Classical Laplace innovations

The main interest of this paper is the asymmetric response in volatility from non-Gaussian innovations, taking heavy tails and LE of asset returns into account as well. The previous GAL innovations involve complex computations since the likelihood only can be written in the computationally intensive special functions. In Kotz et al. (2001), they reduce the computational burden with SCL innovations representing the shocks, which they attain by $\tau = 1$ in the gamma process. Important notice is that the SCL distribution includes similar features as the GAL innovations, where the difference of two random variables creates another random variable with different characteristics, making the specification relevant. The approach to generate SCL follows from the relationship between the standard exponential and the SCL, as the relationship between gamma and GAL in section 2.1. The SCL random variable is given by the formula

$$U \stackrel{d}{=} \psi_1 - \psi_2,$$

where U is a SCL random variable and ψ_1 and ψ_2 are two standard exponential random variables. We incorporate the SCL random variable into the volatility process by replacing

γ_t^+ and γ_t^- in 6 by ψ_1 and ψ_2. More formally, the new error terms are given by

$$\sigma\left(\kappa\psi_t^+ - \frac{\psi_t^-}{\kappa}\right), \tag{11}$$

where the structure follow GAL innovations from (6). The process is known as a double exponential process since it is a symmetrical distribution of two mirrored standard exponential distributions, i.e. SCL.

The final full model is essentially a product of previous arguments in section 2 with a specification of the gamma process by $\tau = 1$. The δ is arbitrary put equal to 1, and κ is also set to 1 since we are interested in the asymmetric response in volatility. The error terms are generated by a difference of two standard exponential random variables presented in (11). With all this in mind, the full volatility model in this paper boils down to

$$y_t = m + ay_{t-1} + \rho_t\sigma\left(\kappa\psi_t^+ - \frac{\psi_t^-}{\kappa}\right), \tag{12}$$

$$\rho_t = 1 + \rho_{t-1}[\alpha[(1-\theta)\psi_{t-1}^+ + (1+\theta)\psi_{t-1}^-] + \beta], \tag{13}$$

where the intercept of the volatility process is set equal to 1, ρ_{t-1} represents past volatility, α is the AR-parameter of the process, β is the persistence of past volatility, θ is the asymmetric parameter, and ψ_{t-1}^+ and ψ_{t-1}^- represent exponential random variables generating SCL distribution as previously explained.

The fully parameterized model is a six parameter distribution following

$$\boldsymbol{\theta}^T = [m\ a\ \alpha\ \beta\ \theta\ \sigma], \tag{14}$$

and $\boldsymbol{\theta}^T$ is the vector containing the parameters of interest, which is based on the conditional heteroscedasticity illustrated by ρ_t from (12). The conditional heteroscedasticity is what enables the ML method to estimate the $\boldsymbol{\theta}^T$ effectively, generating an estimate. This thesis focus on the parameters in the middle α, β and θ.

The condition for the strictly stationary volatility process from (10) given $\delta = 1$ yield the constraint for α and β

$$2\alpha < 1 - \beta. \quad (15)$$

Furthermore, constraints of the parameters in the presented models are summarized in table 2.

	Parameters								
Models	m	a	β	α	θ	δ	σ	κ	τ
GARCH	$[-\infty, +\infty]$	$[0, 1]$	> 0	$[-1, 1]$	-	-	> 0	-	
APARCH_{e_t}	$[-\infty, +\infty]$	$[0, 1]$	> 0	$[-1, 1]$	$[0, 1]$	> 0	> 0	-	
APARCH_{γ_t}	$[-\infty, +\infty]$	$[0, 1]$	> 0	$[-1, 1]$	$[0, 1]$	> 0	> 0	$[-\infty, +\infty]$	$[0, 1]$

Table 2: Constraints of the parameters in the models with processes.

Conditions for the strong stationary process that has a finite method of moments for the GARCH are further described in Cryer and Chan (2008), the APARCH in Javed and Podgórski (2014) and Ding et al. (1993) and the non-Gaussian APARCH in Javed and Podgórski (2014).

3 Model selection and fitting

The previous section has been an introduction to the full asymmetric volatility model. This thesis will, as previously mentioned, only focus on the volatility process, (13). Next, we introduce the structure of this section. The first subsection presents the advantages of using the SCL distribution and compares methods applied to similar problems estimating the parameters. The second subsection is subject to disentangle the methodology applied to evaluate the volatility process of section 2.2, by defining the pdfs and the log-likelihood functions for the symmetric and asymmetric case.

3.1 SCL and likelihood

Arguments behind the specification of the gamma distribution by $\tau = 1$ are numerous. The utmost and prominent advantage is that the differences of two exponential processes generate the SCL distribution. Benefits of using the SCL distribution follow from, as mentioned that the maximum likelihood function and densities of the distribution can be tracked and, in comparison to the gamma distribution, easily be evaluated. This will be further elaborated in section 3.2. Another benefit is that there is one less parameter to estimate, since $\tau = 1$. Furthermore, the SCL still exhibit the desired features of leptokurtosis behaviour in comparison to the normal distribution, which is central to this investigation. In conclusion, the model still shares a lot of important features of interest in a more clear and convenient presentation. In Javed and Podgórski (2016) they estimate the parameters in the APARCH with GAL innovations by two different methods, the standard Gaussian maximum likelihood (ML) and a constrained ML. The constraint implies that some of the parameters are predetermined, from empirical observations. They argue that the Monte Carlo (MC) simulations needed to estimate some of the parameters were too inconsistent due to a significant bias in the MC estimate. Javed and Podgórski (2016) also discuss the use of quasi maximum likelihood (QML), although, it is not evaluated. This method implies that the estimates are asymptotically normal distributed, given that the volatility and mean are correctly specified. However, Bollerslev (1986) state that the estimates are not efficient under asymmetric distributions.

The constrained likelihood method evaluated in Javed and Podgórski (2016) is of relevance to this paper, in particular, the theoretical acumen of the standard maximum likelihood for the APARCH without leverage. The leverage effect is excluded by putting $\theta = 0$ in (4). The results illustrate that the model of interest makes sense mathematically, proving the existence of the fourth moment (kurtosis) close to the boundary. Evaluation of the parameters is of importance since this ensures the existence of the model and that we can evaluate the assumption of the finite kurtosis. Evaluating the parameters is, however, harder when we base the MLE on real-world observations, which might take the kurtosis beyond the defined limit. Inquiries on such data are left for further research.

3.2 Likelihood based methods for the volatility process

This section intends to appraise the volatility process. The analysis of fitting the parameters is initiated by simulations of ρ_t following (12). The recursive relation for ρ_t follow the proposal densities of ψ_t^+ and ψ_t^- accordingly

$$\rho_t = 1 + \rho_{t-1}[\alpha[(1-\theta)\psi_{t-1}^+ + (1+\theta)\psi_{t-1}^-] + \beta]. \tag{16}$$

The choice of parameters included in the simulations to generate a volatility process follows the limit for finite moments in Javed and Podgórski (2016) previously presented. This will be evaluated further later on in this section.

The simulated values of ρ_t includes the observed values from which we estimate the parameters. The estimated parameters are then compared to the true values of the parameters. The likelihood function for ρ_t evaluated for this inquiry follows

$$L(\boldsymbol{\theta}; \rho_1, \ldots, \rho_t | \rho_0) = f_{\boldsymbol{\theta}}(\rho_t | \rho_{t-1}, \rho_0) \ldots f_{\boldsymbol{\theta}}(\rho_1 | \rho_0), \tag{17}$$

where ρ_{t-1} and ρ_0 are assumed to exist and the initial value of ρ is known. Since the simulation is a recursive process, the starting value of ρ_0 needs to be assumed. The initial value of the process is known by a legitimate guess Javed and Podgórski (2016), fulfilling assumptions and constraints. This is not a big issue since the process is under the assumption of ergodicity, meaning that the process will converge to its true value given enough iterations and that the model is correctly specified. Which implies that the choice of ρ_0 is arbitrary, putting it equal to 1.

The SCL and (17) pave the ground for evaluating the density of ρ_t to generate the log likelihood function. The volatility in this context is, as mentioned, an autoregressive process where the pdf for $f_{\boldsymbol{\theta}}(\rho_1|\rho_0)$ define the structure for the full likelihood function. Focusing on this in particular, the density $f_{\boldsymbol{\theta}}(\rho_1|\rho_0)$ is attained from the volatility process by the formula

$$\rho_1 = 1 + \rho_0[\alpha[(1-\theta)\psi_0^+ + (1+\theta)\psi_0^-] + \beta], \tag{18}$$

where the expression contains the parameters α_0, θ, α and β. The initial value and α_0 for the process is, as mentioned, given. Furthermore, θ will be subsumed by the summation of the exponential random variables ψ_0^+ and ψ_0^- in the expression for ρ_1 and therefore impossible to estimate. The reason is that the simulated volatility process is built upon i.i.d random variables, making the dependence structure disappear. The parameter θ will be observed in the log returns, although outside the scope of this paper. With all this in mind, the pdf only depends upon the random variables ψ^+ and ψ^-.

To ease up the algebraic calculations, simplifications of 18) are incorporated to find the generic pdf and log-likelihood function for the volatility process. The formula gives the expressions

$$\rho = a + b\ \psi^+ + c\ \psi^-, \tag{19}$$

where, $\rho = \rho_1$, $a = 1 + \beta\rho_0$, $b = \alpha(1-\theta)\rho_0$ and $c = \alpha(1+\theta)\rho_0$. This alternation creates a similar structure and lets $\psi^- \in Exp(1)$, $\psi^+ \in Exp(1)$ be independent random variables with the relation from (19). Then the explicit cumulative density function can be attained accordingly

$$
\begin{aligned}
F_\rho(\rho) &= \mathbf{P}(\rho < \rho) = \mathbf{P}(a \ + \ b\ \psi^+ \ + \ c\ \psi^- < \rho) = \mathbf{P}\left(\psi^+ < \frac{\rho - c\ \psi^- - a}{b}\right) \\
&= \int_0^{\frac{\rho-a}{c}} \int_0^{\frac{\rho - c\psi^- - a}{b}} e^{-\psi^+ - \psi^-}\ d\psi^+\ d\psi^- \\
&= \int_0^{\frac{\rho-a}{c}} e^{-\psi^-}\left[1 - e^{-\frac{(\rho - c\psi^- - a)}{b}}\right] d\psi^- \\
&= \int_0^{\frac{\rho-a}{c}} e^{-\psi^-} - e^{-\psi^-(1-c/b) - \frac{(\rho-a)}{b}} d\psi^- \\
&= -e^{-\frac{\rho-a}{c}} + \frac{e^{-\frac{(\rho-a)}{c}(1-c/b) - \frac{(\rho-a)}{b}}}{1-c/b} - \frac{e^{-\frac{(\rho-a)}{b}}}{1-c/b} + 1 \\
&= \frac{e^{-\frac{(\rho-a)}{c} - \frac{(\rho-a)}{b} - \frac{(\rho-a)}{b}}}{1-c/b} - \frac{e^{-\frac{\rho-a}{b}}}{1-c/b} - e^{-\frac{\rho-a}{c}} + 1 \\
&= \frac{e^{-\frac{(\rho-a)}{c}}}{1-c/b} - \frac{e^{-\frac{(\rho-a)}{b}}}{1-c/b} - e^{-\frac{(\rho-a)}{c}} + 1 \\
&= \frac{be^{-\frac{(\rho-a)}{c}}}{b-c} - \frac{be^{-\frac{(\rho-a)}{c}}}{b-c} - \frac{(b-c)e^{-\frac{(\rho-a)}{b}}}{b-c} + 1 \\
&= \frac{ce^{-\frac{(\rho-a)}{c}} - be^{-\frac{(\rho-a)}{b}}}{b-c} + 1, \qquad\qquad \text{for } b \neq 0
\end{aligned}
\tag{20}
$$

where the relation to the pdf pursue,

$$
F'_{\boldsymbol{\rho}}(\rho) = f_{\boldsymbol{\rho}}(\rho) = \begin{cases} \dfrac{e^{-\frac{(\rho-a)}{b}} - e^{-\frac{(\rho-a)}{c}}}{b-c}, & b \neq c \\[2ex] \dfrac{(\rho-a)e^{-\frac{(\rho-a)}{c}}}{c^2}, & b = c. \end{cases}
\tag{21}
$$

The original expressions for a, b and c in (18) are substituted and the density of the volatility process follows

$$
f_{\boldsymbol{\theta}}(\rho_1|\rho_0) = \begin{cases} \dfrac{e^{-\frac{(\rho_1-(1+\beta\rho_0))}{\alpha(1-\theta)\rho_0}} - e^{-\frac{(\rho_1-(1+\beta\rho_0))}{\alpha(1+\theta)\rho_0}}}{\rho_0\alpha((1-\theta)-(1+\theta))}, & \theta \neq 0 \\[2ex] \frac{(\rho_1-\beta\rho_0-1)e^{-\frac{(\rho_1-(1+\beta\rho_0))}{\alpha\rho_0}}}{(\alpha\rho_0)^2}, & \theta = 0. \end{cases}
\tag{22}
$$

Finally, the generic pdf for the volatility process conditioning on the previous term accordingly

$$f_{\boldsymbol{\theta}}(\rho_t|\rho_{t-1},\rho_0) = \begin{cases} \dfrac{e^{-\frac{(\rho_t-(1+\beta\rho_{t-1}))}{\alpha(1+\theta)\rho_{t-1}}} - e^{-\frac{(\rho_t-(1+\beta\rho_{t-1}))}{\alpha(1-\theta)\rho_{t-1}}}}{2\alpha\theta\rho_{t-1}} & \theta \neq 0, \\ \frac{(\rho_t-\beta\rho_{t-1}-1)e^{-\frac{(\rho_t-(1+\beta\rho_{t-1}))}{\alpha\rho_{t-1}}}}{(\alpha\rho_{t-1})^2} & \theta = 0, \end{cases} \tag{23}$$

where, ρ_t, ρ_{t-1} and ρ_0 represent the observed volatility at time t, $t-1$ and the initial value of the process respectively. The pdf is divided into two cases, where $\theta \neq 0$ and $\theta = 0$. The first expression includes asymmetric behaviour, while the second does not. Important to emphasize is that the asymmetry refers to the response in volatility, not the distribution itself. The parameter θ aims to accentuate or restrain certain features of the distribution, taking asymmetric response in volatility into account.

In the asymmetric setting, θ and α has the most significant impact on the shape and scaling since the parameters are in the denominator of the exponential expressions. A $\theta > 0.5$ indicates that one of the exponential terms are highly correlated to the volatility process at $t-1$ while the other will not, given an $\alpha > 0$, which implies that there is an asymmetric effect present. The β parameter reflects how much of past volatility that is affecting the volatility at time point t. The inclusion of β is merely an effect of the generalization from the GARCH generating a more flexible structure, although quite restrained considering that the pdf is olny defined for $\rho_t > (1+\beta\rho_{t-1})$.

The second equation in (23) is reduced to the parameters α and β, since $\theta = 0$. The expression presents features resembling the gamma distribution. The scaling and shape are mostly effected by α since the denominator of the full expression are quadratic. β has a similar impact, as previously explained in the asymmetric case. Furthermore, to evaluate the parameter space, the log-likelihood function is derived and evaluated for $\theta \leq 0$ and $\theta = 0$.

Starting with the case of $\theta \leq 0$ the log-likelihood function is evaluated with respect to $\boldsymbol{\theta} = [\beta\ \alpha\ \theta]$. The log-likelihood function is attained via the upper expression for the pdf in (23) following

$$
\begin{aligned}
l(\boldsymbol{\theta};\rho_t,\ldots,\rho_1|\rho_0) &= log\prod_{t=1}^{T}\left[\frac{e^{-\frac{(\rho_t-\beta\rho_{t-1}-1)}{\alpha(1+\theta)\rho_{t-1}}}-e^{-\frac{(\rho_t-\beta\rho_{t-1}-1)}{\alpha(1-\theta)\rho_{t-1}}}}{2\alpha\theta\rho_{t-1}}\right]\\
&= -\sum_{t=1}^{T}log(2\alpha\theta\rho_{t-1})+\sum_{t=1}^{T}log\left[e^{-\frac{(\rho_t-\beta\rho_{t-1}-1)}{\alpha(1+\theta)\rho_{t-1}}}\left[1-e^{-\left(\frac{(\rho_t-\beta\rho_{t-1}-1)}{\alpha(1-\theta)\rho_{t-1}}-\frac{(\rho_t-\beta\rho_{t-1}-1)}{\alpha(1+\theta)\rho_{t-1}}\right)}\right]\right]\\
&= -\sum_{t=1}^{T}log(2\alpha\theta\rho_{t-1})-\sum_{t=1}^{T}\left[\frac{\rho_t-\beta\rho_{t-1}-1}{\alpha(1+\theta)\rho_{t-1}}\right]+\\
&\quad\sum_{t=1}^{T}log\left[1-e^{-\left(\frac{(\rho_t-\beta\rho_{t-1}-1)}{\alpha(1-\theta)\rho_{t-1}}-\frac{(\rho_t-\beta\rho_{t-1}-1)}{\alpha(1+\theta)\rho_{t-1}}\right)}\right]. \qquad (24)
\end{aligned}
$$

The first term presents the contracted form of the autoregressive parameter α and the asymmetry parameter θ. The numerator of the second expression represents the sum of the volatility process from time point t to T, subdued by the endurance of volatility. The denominator in (24) constituents of the actual weight for the innovations. However, this is not an issue since the density is symmetric meaning that the negative weight can be extracted from (24) generating another expression with corresponding constituents.

A closer look at (24) we can see that it display strong resemblance to the expression for logistic regression given by the formula

$$
l(\alpha,\beta)=\sum_{i=1}^{N}y_i(\alpha+f_1(X_{i1})\ +\ \ldots\ +\ f_p(X_{ip}))-log(1+e^{(\alpha+f_1(X_{i1})+\ldots+f_p(X_{ip}))}). \qquad (25)
$$

The resemblance between (25) and (24) implies that the parameters in the volatility process might be estimated by utilizing the methods for the logistic distribution. The methods are complicated, but the area is well studied, and there are many robust features. Although, the investigation of this inquiry stretches beyond the scope of this note.

The log-likelihood for the special case where $\theta = 0$ is derived via the bottom expression in (23) following

$$
\begin{aligned}
l(\boldsymbol{\theta};\rho_1,\ldots,\rho_t|\rho_0) &= log\left[\prod_{t=1}^{T}\frac{(\rho_t-\beta\rho_{t-1}-1)e^{-\frac{(\rho_t-\beta\rho_{t-1}-1)}{\alpha\rho_{t-1}}}}{(\alpha\rho_{t-1})^2}\right] \qquad (26)\\
&= \sum_{t=1}^{T}log(\rho_t-\beta\rho_{t-1}-1)-2\sum_{t=1}^{T}log(\alpha\rho_{t-1})-\sum_{t=1}^{T}\frac{(\rho_t-\beta\rho_{t-1}-1)}{\alpha\rho_{t-1}}.
\end{aligned}
$$

The first term presents the endurance of past volatility in the process emphasized by -1. Similar to the first term in the asymmetric case, the second term includes the autoregressive parameter. The last term presents similar features to the normal distribution without the quadratic form. What is apparent from the final log-likelihood for $\theta = 0$, is the simplicity of the function, which is of utmost interest.

4 Evaluation of the volatility process

This section assesses the previously presented volatility process with derived pdf and log-likelihood. The optimization problem consists of evaluating the functions by systematically choosing input values from within an allowed set and computing the value of the function. The sets in this setting represent the volatility process and the different parameters of interest. We use sets that intentionally does not fulfil the assumptions and constraints to see where the functions break down. We apply this approach to both the asymmetric and symmetric case for the pdf and the log-likelihood functions.

4.1 Initial computations of the volatility process

We evaluate the volatility process for a finite fourth moment by sets alternating α and β, where α and β varies between 0.01-1 by 0.01 and 0.1-1 by 0.1, respectively. The results present a big influence of α in the process. The process is more sensitive to large α in comparison to a large β in line with the constraint in (15). The process starts to generate non-finite features for $\alpha > 0.55$, with extreme spikes in volatility. An $\alpha < 0.1$ generates a process with no features of volatility clustering, making these uninteresting as well. Therefore, the main focus of evaluating the α was between $0.11 - 0.25$. The increase of β, increase the spikes for the α values of interest. Although, the processes look very similar in the α interval, which is to expect since the simulated exponential random variables do not change for the change in α and β. However, the choice falls upon a $\beta = 0.3$ and an $\alpha = 0.19$ since the spikes are not too unrealistic, the process exhibit volatility clustering features and the condition of (10) is fulfilled, see figure 1.

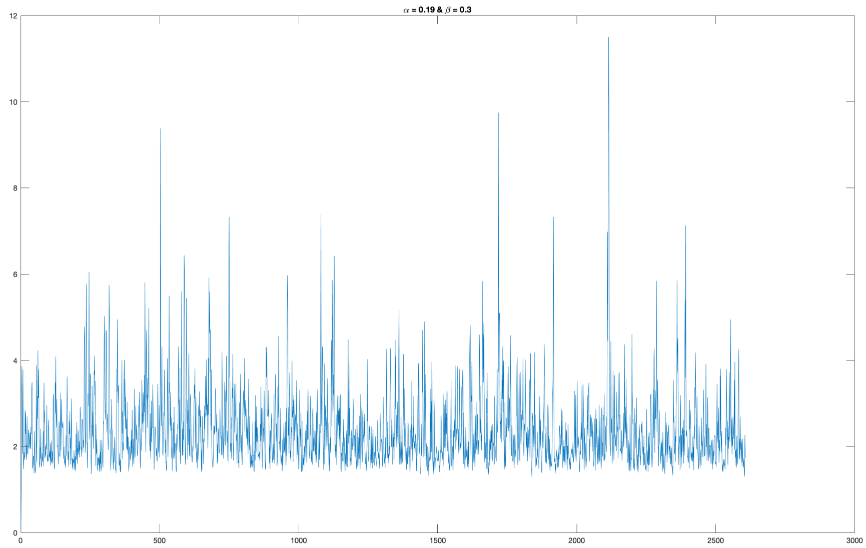

Figure 1: The figure presents the stationary volatility processes computed for $\alpha = 0.19$ and $\beta = 0.3$.

4.2 Computing the probability density function

The densities of the volatility process are evaluated by a stationary set of ρ_t varying the parameters α, β and θ for $\theta \geq 0$ and similar for α and β in the case of $\theta = 0$, Givens and Hoeting (2012). The pdf is given by the formula

$$f_{\boldsymbol{\theta}}(\rho_t|\rho_{t-1}, \rho_0) = \begin{cases} \dfrac{e^{-\frac{(\rho_t-(1+\beta\rho_{t-1}))}{\alpha(1+\theta)\rho_{t-1}}} - e^{-\frac{(\rho_t-(1+\beta\rho_{t-1}))}{\alpha(1-\theta)\rho_{t-1}}}}{2\alpha\theta\rho_{t-1}}, & \theta \neq 0 \\ \\ \frac{(\rho_t-\beta\rho_{t-1}-1)e^{-\frac{(\rho_t-(1+\beta\rho_{t-1}))}{\alpha\rho_{t-1}}}}{(\alpha\rho_{t-1})^2}, & \theta = 0, \end{cases} \tag{27}$$

which is the result of previous derived densities. The generated density functions for $\theta \geq 0$ and $\theta = 0$ are presented in figure 2, 3, 4 and 5.

I start to evaluate the pdf for a set of α, where the model is defined under the constraint of $2\alpha < 1 - \beta$. The density present interesting features of positive kurtosis and extreme fat tails, which are desirable modelling volatility and financial returns. A smaller α indicates thinner tails in comparison to a larger α, which generates heavier tails. The figure also present features of less leptokurtosis for larger α's and vice versa for a lower α. Furthermore, the asymmetry in this distribution exists by default because of its composition, with features of the gamma distribution. However, relating this to the definition of the pdf for $\theta \geq 0$ we expect an increase in asymmetry as α increase which is apparent in figure 2.

Next, We evaluate β in the pdf under the same constraint as α. Also, the volatility process is not defined for negative values, which implies that $\rho_t - (1 + \beta\rho_{t-1}) > 0$ must hold. Therefore, a shift in the density is apparent. Centring all the densities around 0 takes this issue into account. As mentioned, β represents the persistence of past volatility which is what we can observe in figure 3. A larger value for β present allowance of more extreme values, which is more probable for a higher lagged structure whereas lower values for β implies the opposite with less leptokurtosis.

The final parameter θ can be viewed as a scaling parameter of the asymmetric effect. An increase of θ increases leptokurtosis for the heavy tail and reduces the probability of extreme values. In the context of the pdf, a higher θ will increase leptokurtosis for the part of the density containing $(1 + \theta)$ and the opposite for the part of the density containing $(1 - \theta)$, see figure 4. Corollary a higher θ imply that the exponential expression including $(1 - \theta)$ converges towards zeros very fast since the term generates a large value in the exponent of the negative exponential, given a certain α, see figure 4 and the opposite for the part including $(1 + \theta)$. The results confirm that the θ parameter can represent the scaling of the asymmetry in the distributions, in line with Javed and Podgórski (2014) and Kotz et al. (2001).

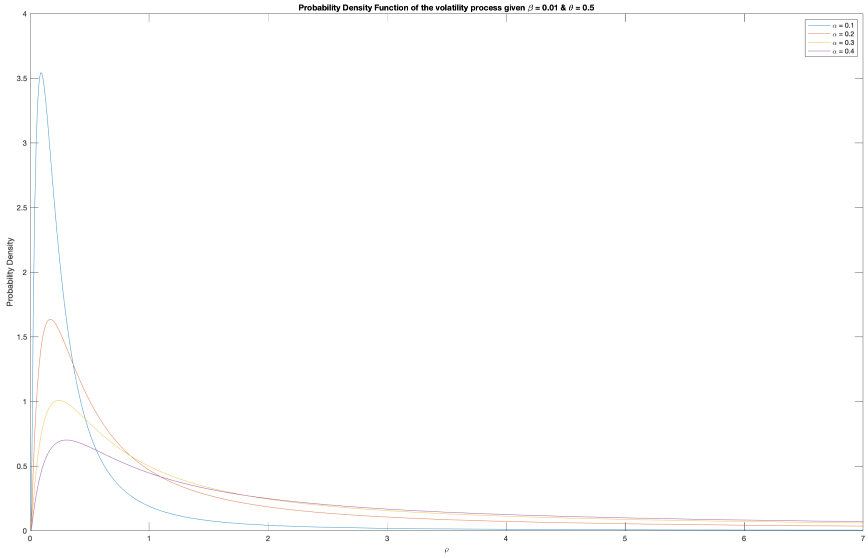

Figure 2: Pdf of the volatility process for a set of α and ρ given a specific value of β and θ, under the assumptions $\theta \geq 0$ and $2\alpha > 1 - \beta$.

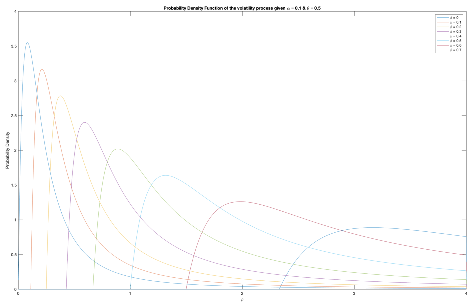

Figure 3: Pdf of the volatility process for a set of β and ρ given a specific value of α and θ, under the assumptions $\theta \geq 0$ and $2\alpha > 1 - \beta$.

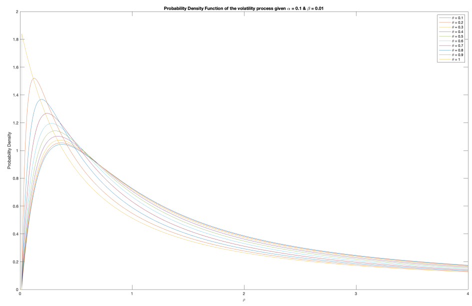

Figure 4: Pdf of the volatility process for a set of θ and ρ given a specific value of α and β, under the assumptions $\theta \geq 0$ and $2\alpha > 1 - \beta$.

Next, we evaluate the pdf for the particular case of $\theta = 0$, varying α and β. Figure 5 present densities for a fixed β and a set of α and ρ. In the figure, we can see similar features to the pdf with $\theta \geq 0$, with skewness to the right and strong leptokurtosis for lower values of α. The distributions does not present as heavy tails as the pdf for $\theta \geq 0$, which is to expect since we do not emphasize this feature for $\theta = 0$. From figure 5 we also conjecture behaviour very similar to the previous case as the distribution becomes more heavy-tailed when α takes on larger values. The distributions for $\theta = 0$ assume less positive kurtosis in comparison to $\theta \geq 0$. The densities for a fixed α and a set of β is comparable to $\theta \geq 0$ in figure 3, where the densities wander with heavier leptokurtosis as beta increases.

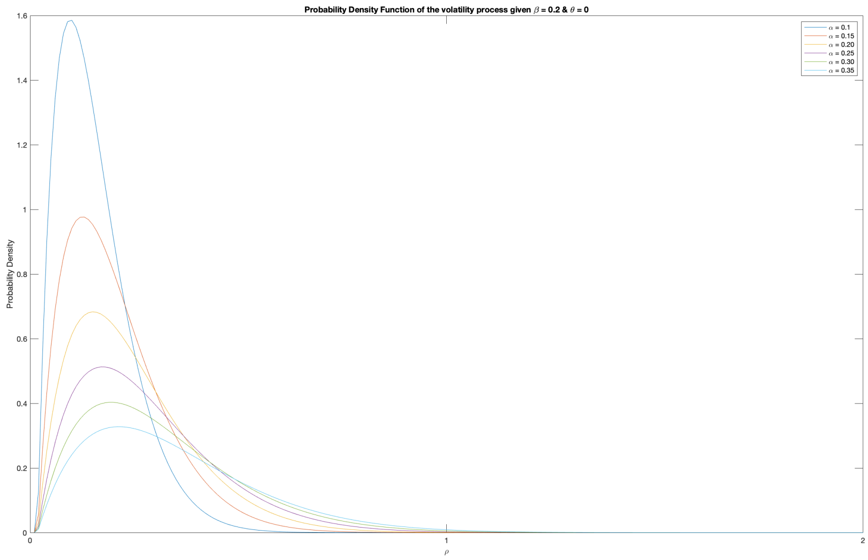

Figure 5: Pdf of the volatility process for a set of α and ρ given a specific value of β, under the assumptions $\theta = 0$ and $2\alpha > 1 - \beta$.

4.3 Exploring the log-likelihood for the parameter space

The parameter space for the log-likelihood functions are evaluated for both $\theta \geq 0$ and the special case of $\theta = 0$. We evaluate the first function through computation concerning the parameters of interest, similar to the earlier presented pdf. The latter case is evaluated through analytical solutions for the parameters of interest. Worth mentioning for the case of $\theta = 0$ because of its attractive feature of analytical solution is that it can be used as a benchmark and compared to the estimates. However, the comparison is outside the scope of this thesis.

The log-likelihood for the asymmetric is derived in (24) and given by the formula

$$l(\boldsymbol{\theta}; \rho_1, \dots, \rho_t | \rho_0) = \tag{28}$$

$$-\sum_{t=1}^{T} log(2\alpha\theta\rho_{t-1}) - \sum_{t=1}^{T} \left[\frac{\rho_t - \beta\rho_{t-1} - 1}{\alpha(1+\theta)\rho_{t-1}} \right] + \sum_{t=1}^{T} log \left[1 - e^{-\left(\frac{(\rho_t - \beta\rho_{t-1} - 1)}{\alpha(1-\theta)\rho_{t-1}} - \frac{(\rho_t - \beta\rho_{t-1} - 1)}{\alpha(1+\theta)\rho_{t-1}} \right)} \right].$$

The function is maximized concerning the parameters α, β and θ, one at the time. The same set of the underlying stochastic process, ρ, is used for all parameters. The parameter values of α and β are chosen under the constraint of $2\alpha < 1 - \beta$. We also assume that $\theta \geq 0$.

Figure 6 presents the estimate $\hat{\alpha}$, which displays interesting features of convergence towards the true value of $\alpha = 0.19$. Figure 6 also implies we can assume that $\hat{\alpha}$ is the unique maximum likelihood estimator.

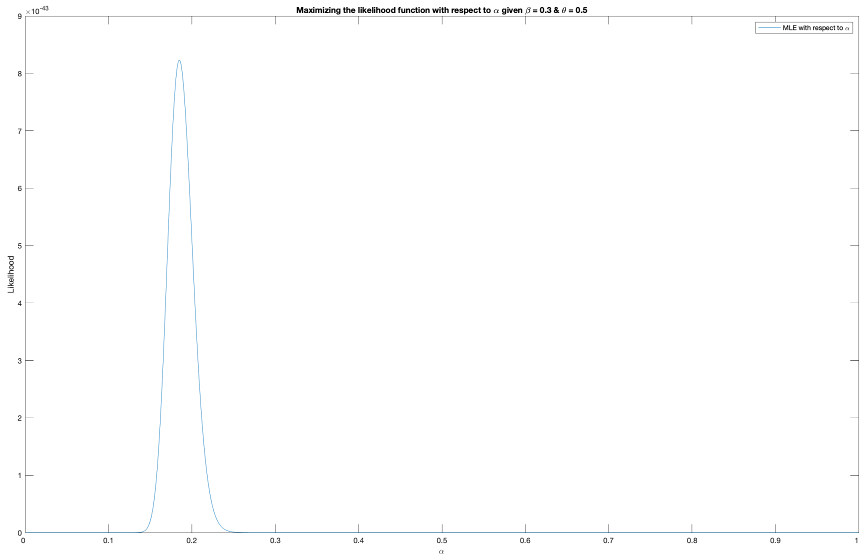

Figure 6: Maximum likelihood estimator $\hat{\alpha}$ given $\beta = 0.3$ and $\theta = 0.5$.

The next parameter is β. Figure 7 display salient qualities, similar to $\hat{\alpha}$ with convergence towards the true value. Soon after the likelihood function has reached its maximum, the model is not defined since $\rho_t > (1 + \beta\rho_{t-1})$ is no longer fulfilled.

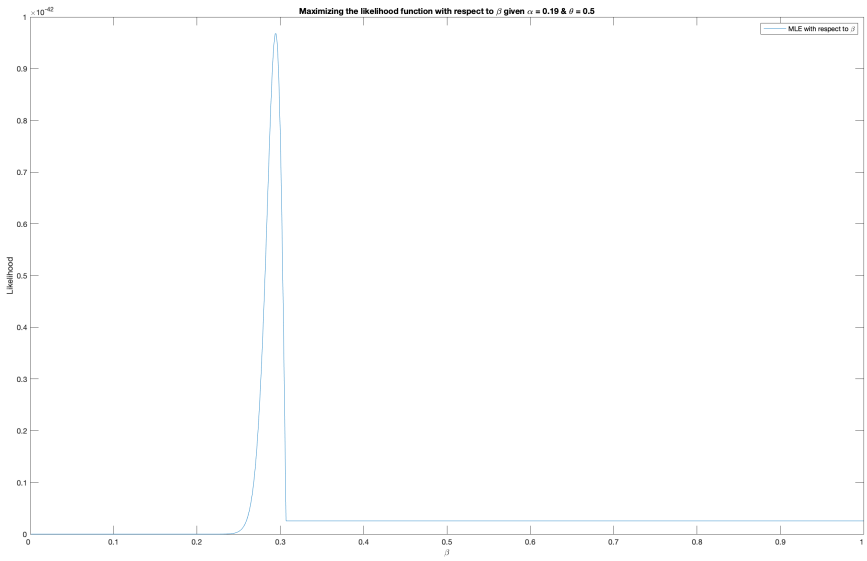

Figure 7: MLE of the parameter β given $\alpha = 0.19$ and $\theta = 0.5$. The estimate is evaluated for $n = 100$.

Finally, we assess the last estimator of interest, $\hat{\theta}$. In comparison to the other two estimators, figure 8 present no evidence of convergence towards the real value of θ. Earlier, we stated that we can observe θ in the full model for the returns, but that it might be hard in the volatility process, which is apparent from the figure. It is, as mentioned, because the asymmetric component $[(1-\theta)\phi^{+}_{t-1} + (1+\theta)\phi^{-}_{t-1}]$ is assimilated in to one homogeneous part of the volatility process.

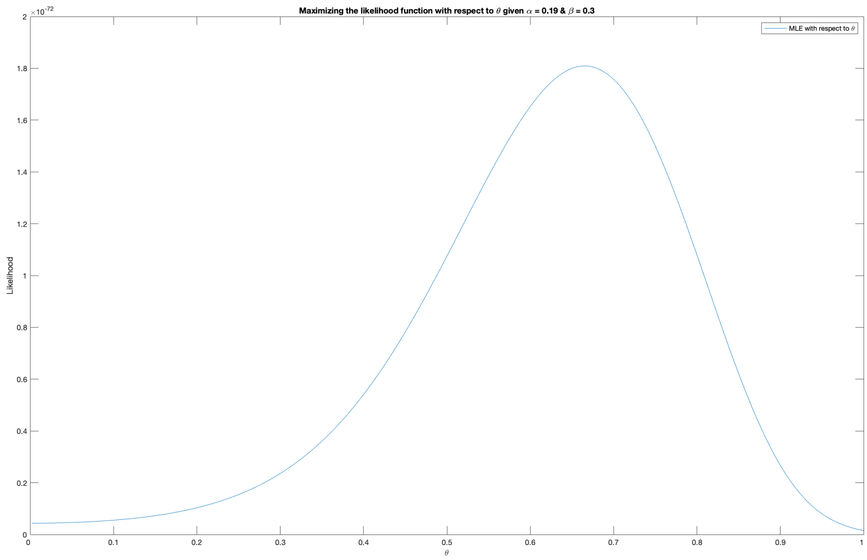

Figure 8: MLE of the parameter θ given $\alpha = 0.19$ and $\beta = 0.3$.The estimate is evaluated for $n = 100$.

Next, we assess the analytical solution to the log-likelihood function of $\theta = 0$ concerning α and β. Partial differentiating the log-likelihood function (26) with respect to α and β, putting the expressions equal to zero to find for which expression for the parameters that maximizes the function. The formula gives the partial differentiation with respect to α

$$\frac{\partial l(\boldsymbol{\theta}; \rho_1, \ldots, \rho_t|\rho_0)}{\partial \alpha} = \frac{\partial}{\partial \alpha}\left(\sum_{t=1}^{T} log(\rho_t - \beta\rho_{t-1} - 1) - 2\sum_{t=1}^{T} log(\alpha\rho_{t-1}) - \sum_{t=1}^{T}\frac{(\rho_t - \beta\rho_{t-1} - 1)}{\alpha\rho_{t-1}}\right) = 0$$

$$\begin{aligned} &= -\frac{2T}{\alpha} + \sum_{t=1}^{T}\frac{(\rho_t - \beta\rho_{t-1} - 1)}{\alpha^2\rho_{t-1}} = 0 \\ \alpha &= \sum_{t=1}^{T}\frac{(\rho_t - \beta\rho_{t-1} - 1)}{2T\rho_{t-1}} \\ \alpha &= \frac{1}{2T}\sum_{t=1}^{T}\frac{\rho_t}{\rho_{t-1}} - \frac{1}{2T}\sum_{t=1}^{T}\frac{1+\beta\rho_{t-1}}{\rho_{t-1}}. \end{aligned} \tag{29}$$

The α that maximizes the function is divided into two parts, both dependent on past volatility and the constant $\frac{1}{2T}$. The first part represents the sum over time of volatility at time t divided by volatility at $t-1$. The second part of the expression is 1 plus the endurance in volatility at time $t-1$ divided by past volatility at $t-1$ summed over T steps.

The next parameter of interest is β. The parameter is maximized in a similar matter by partial differentiating the log-likelihood concerning β. The scheme can be seen bellow

$$\begin{aligned}
\frac{\partial l(\boldsymbol{\theta};\rho_1,\ldots,\rho_t|\rho_0)}{\partial\beta} &= \frac{\partial}{\partial\beta}\left(\sum_{t=1}^{T} log(\rho_t-\beta\rho_{t-1}-1)-2\sum_{t=1}^{T} log(\alpha\rho_{t-1})-\sum_{t=1}^{T}\frac{(\rho_t-\beta\rho_{t-1}-1)}{\alpha\rho_{t-1}}\right)=0\\
&= \frac{\partial}{\partial\beta}\left(\sum_{t=1}^{T} log(\rho_t-\beta\rho_{t-1}-1)-\sum_{t=1}^{T}\frac{(\rho_t-\beta\rho_{t-1}-1)}{\alpha\rho_{t-1}}\right)=0\\
&= \sum_{t=1}^{T}\left(\frac{\rho_{t-1}}{\rho_t-\beta\rho_{t-1}-1}\right)-\frac{T}{\alpha}=0 \qquad (30)
\end{aligned}$$

where the solution to the log-likelihood function with respect to β is in the denominator of the first expression. The denominator as a whole is previously described as one of the constraints of the model, $\rho > (1+\beta\rho_{t-1})$, which needs to be fulfilled for the model to be defined. The numerator is volatility at $t-1$. The second part is T over α which will suppress the first part since both α and T is positive. It is hard to say the outcome of the expression but can be evaluated using an iterative operation.

An iterative approach is applied to ensure that the values of α and β converges towards reasonable values for all values of β and α, given a stationary volatility process. The initial values of α and β are predetermined, then the expressions are sequentially updated for each iteration. The parameters converge after approximately 50 iterations towards 0.1862 and 0.2987, which implies that (29) and (30) are correct.

5 Conclusion

Starting with the research question, the simulated volatility process from the two exponential random variables exhibit volatility clustering and stochastic behaviour. The pdf for $\theta \neq 0$ present exciting features in line with the 'stylized facts' of interest, in particular, the asymmetric response in returns. The pdf also presented strong positive kurtosis, which is an attractive feature when modelling financial returns. The results and reasoning for the case of $\theta = 0$ imply similar characteristics but not as strong, as to be expected since we do not emphasize the asymmetry in the volatility process.

The simulated maximum likelihood estimates for $\theta \geq 0$ showed fast convergence for α and β. For θ, we concluded that the asymmetry is observed in the real returns and cannot be evaluated in the volatility process. The paper also outlines the analytic solutions for the parameters that maximize the log-likelihood for $\theta = 0$. Since the analytical results are hard to interpret, especially for β, we introduce an iterative operation. The results present fast convergence of the parameters and conclude that the results applies for all α and β.

This exploratory paper found compelling features utilizing the SCL distribution, which are suitable for the 'stylized facts' often found in financial returns with many precise results. However, further research is needed to fully assess the volatility process and the following full model predicting returns model using non-Gaussian innovations. Another interesting approach, as previously mentioned, would be to explore the parameters of interest by applying logistic regression. It would also be fruitful to examine the comparison between the analytical and computed solutions.

As mentioned, this paper aims at being a building stone exploring non-Gaussian innovations which it has and given theoretical insights. These results can be utilized in future research to benchmark the analysis of the complete and general models for assessing volatility.

References

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the Business and Economics Section of the American Statistical Association*, pages 177 – 181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307 – 327.

Cryer, J. D. and Chan, K.-S. (2008). *Time Series Analysis: With Application in R.* Springer-Verlag New York Inc., 2nd edition.

Ding, Z., Granger, C. W., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1):83 – 106.

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

Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The Journal of Finance*, 48(5):1749–1778.

Givens, G. H. and Hoeting, J. A. (2012). *Computational Statistics.* John Wiley and Sons, Inc, 2nd edition.

Glosten, L., Jagannathan, R., and Runkle, D. (1993). A relationship between the expected value and the volatililty of the nominal excess return on stocks. *The Journal of Finance*, 48(5):1779–1801.

Hol, E. M. J. H. (2010). *Empirical Studies on Volatility in International Stock Markets.* Springer-Verlag New York Inc., 6th edition.

Javed, F. and Podgórski, K. (2014). Effect for volatility with generalized Laplace error. *Economic Quality Control*, 29(2):157 – 166.

Javed, F. and Podgórski, K. (2016). Tail behavior and dependence structure in the APARCH model. *Journal of Time Series Econometrics*, 9(2).

Jensen, M. and Lunde, A. (2001). The NIG-S and ARCH model: a fat-tailed, stochastic, and autoregressive conditional heteroskedastic volatility model. *The Econometrics Journal*, 4(4):319–342.

Kotz, S., Kozubowski, T., and Podgórski, K. (2001). *The Laplace Distribution and Generalizations*. Springer.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *The Econometric Society*, 59(2):347 – 370.