MATHEMATICAL MODELING AND COMPUTING, Vol. 8, No. 3, pp. 379–390 (2021)

Quasi-maximum likelihood estimation of the Component-GARCH model using the stochastic approximation algorithm with application to the S&P 500

Settar A.[1], Fatmi N. I.[1], Badaoui M.[1,2]

[1] *IPIM, National Schools of Applied Sciences, Khouribga, Sultan Moulay Slimane University, Morocco*
[2] *LaMSD, Higher School of Technology, Oujda, Mohammed First University, Morocco*

(Received 25 April 2021; Revised 7 May 2021; Accepted 17 May 2021)

The component GARCH (CGARCH) is suitable to better capture the short and long term of the volatility dynamic. Nevertheless, the parameter space constituted by the constraints of the non-negativity of the conditional variance, stationary and existence of moments, is only ex-post defined via the GARCH representation of the CGARCH. This is due to the lack of a general method to determine a priori the relaxed constraints of non-negativity of the CGARCH(N) conditional variance for any $N \geqslant 1$. In this paper, a CGARCH parameter space constructed from the GARCH(1,1) component parameter spaces is provided a priori to identifying its GARCH form. Such a space fulfils the relaxed constraints of the CGARCH conditional variance non-negativity to be pre-estimated ensuring the existence of a QML estimation in the sense of the stochastic approximation algorithm. Simulation experiment as well as empirical application to the S&P500 index are presented and both show the performance of the proposed method.

Keywords: *component GARCH, conditional variance, stochastic approximation, Kalman filter, quasi-maximum likelihood.*

2010 MSC: 62L20, 62M10, 60G12, 91B05, 91G60 **DOI:** 10.23939/mmc2021.03.379

1. Introduction

One of the stylized facts common to the financial return series which is related to the time series dependencies in volatility is that returns show a little serial correlation while the squared returns are highly serially correlated. The introduction of the ARCH model [1] generalized to the GARCH model [2] allowed to account this feature. A process ε_t is called GARCH(p, q) if it satisfies :

$$\varepsilon_t = \sigma_t \eta_t, \quad \eta_t \sim \mathrm{iid}(0,1), \tag{1}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \quad t \in \mathbb{Z} \tag{2}$$

with $\omega > 0$, $\alpha_i \geqslant 0$ for $i = 1, \ldots, p$ and $\beta_j \geqslant 0$ for $j = 1, \ldots, q$.

In particular, one has the autocorrelation function for the GARCH(1,1) process at lag k, when the fourth moment exists [3], as follows:

$$\rho_k = \left(\alpha_1 + \frac{\alpha_1^2 \beta_1}{1 - 2\alpha_1\beta_1 - \beta_1^2}\right)(\alpha_1 + \beta_1)^{k-1}. \tag{3}$$

Which is approximately given [4] by

$$\rho_k \approx \left(\alpha_1 + \frac{1}{3}\beta_1\right)(\alpha_1 + \beta_1)^{k-1}.$$

Nevertheless, it is clear that the autocorrelation function still decreases exponentially which is in contrast to the long memory property that have been pointed out in several empirical studies. Indeed, it has been revealed that the sample autocorrelation function of the S&P500 absolute returns decreases very fast at the beginning, and then decreases very slowly and remains significantly positive [5], which is different from an exponentially decreasing function. Works of Andersen and Bollerslev [6], Andersen et al. [7], Bollerslev and Wright [8], Karanasos [9] are among others that empirically highlighted the presence of the long memory property. Ding et al. [5], Ding and Granger [4] claim that a such empirical behaviour reveals the existence of different components of volatility dominating different periods. Some components of volatility can have a very significant effect in the short term, but fall very fast. Others may have a relatively smaller short term effect, but last for a long time. In this sense, Ding and Granger [4] introduce the CGARCH (Component GARCH) model as a new specification of volatility by decomposing it into several GARCH component specifications, allowing some to capture the long-term dynamics of volatility, and others to capture its short-term fluctuations.

Thus, it is said that ε_t is a CGARCH(N) process in the sense of Ding and Granger [4] if it verifies:

$$\begin{aligned} \varepsilon_t &= \sigma_t \eta_t, \quad \eta_t \sim \text{iid}(0,1), \\ \sigma_t^2 &= \sum_{i=1}^{N} w_i \sigma_{i,t}^2, \quad \sum_{i=1}^{N} w_i = 1, \\ \sigma_{i,t}^2 &= \omega_i + \alpha_i \varepsilon_{t-1}^2 + \beta_i \sigma_{i,t-1}^2, \quad i = 1, \ldots, N. \end{aligned} \tag{4}$$

Following the specification (4), all the statistical properties of the GARCH model hold for the CGARCH because indeed, any CGARCH(N) composed of N GARCH(1,1), is expressed as a restricted GARCH(N,N). In this sense, Karanasos [9] has managed to make explicit the specification GARCH(N,N) of a CGARCH(N) composed of a IGARCH(1,1) and $(N-1)$ GARCH(1,1) (Lemma 3.1 in [9]).

Let us focus on the non-negativity constraints of the CGARCH conditional variance. So, we consider the CGARCH model of Ding and Granger [4] in the unweighted form used by Maheu [10], that is

$$\begin{aligned} \varepsilon_t &= \left(\sum_{i=1}^{N} \sigma_{i,t}^2 \right)^{\frac{1}{2}} \eta_t = \sigma_t \eta_t, \quad \eta_t \sim \text{iidN}(0,1), \\ \sigma_{i,t}^2 &= \omega_i + \alpha_i \varepsilon_{t-1}^2 + \beta_i \sigma_{i,t-1}^2, \quad i = 1, \ldots, N. \end{aligned} \tag{5}$$

Obviously, non-negativity constraints of σ_t^2 are necessary to be imposed. They are generally identified from the GARCH(N,N) specification associated with CGARCH(N). For instance, for $N = 2$, the CGARCH(2) is expressed as a GARCH(2,2) as follows

$$\sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + a_2 \varepsilon_{t-2}^2 + b_1 \sigma_{t-1}^2 + b_2 \sigma_{t-2}^2,$$

where $a_0 = \omega_1(1-\beta_2)+\omega_2(1-\beta_1)$, $a_1 = (\alpha_1+\alpha_2)$, $a_2 = -(\alpha_1\beta_2+\alpha_2\beta_1)$, $b_1 = \beta_1+\beta_2$ and $b_2 = -\beta_1\beta_2$.

Thus, to keep $\sigma_t^2 > 0$, it is sufficient in the sense of Bollerslev [2] that the CGARCH parameters verify $a_0 > 0$ and $a_i \geqslant 0$ (resp. $b_j \geqslant 0$) for $i = 1, 2$ (resp. $j = 1, 2$). These constraints can be relaxed in the sense of Nelson and Cao [10] to obtain the following inequality set [11]

$$0 < \alpha_2 + \beta_2 < \alpha_1 + \beta_1 < 1, \quad \alpha_1 < \beta_2, \quad 0 < \alpha_1, \quad 0 < \alpha_2, \quad 0 < \omega_1 \quad \text{and} \quad 0 < \omega_2,$$

which defines, in addition to the conditions of stationary and existence of moments, the finder parameter space of a QML estimation.

Now, for $N = 3$, one derives the GARCH(3,3) specification arising from the CGARCH(3) model, that is

$$\sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + a_2 \varepsilon_{t-2}^2 + a_3 \varepsilon_{t-3}^2 + b_1 \sigma_{t-1}^2 + b_2 \sigma_{t-2}^2 + b_3 \sigma_{t-3}^2,$$

where $a_0 = \omega_1(1-\beta_2-\beta_3-\beta_2\beta_3)+\omega_2(1-\beta_1-\beta_3-\beta_1\beta_3)+\omega_3(1-\beta_1-\beta_2-\beta_1\beta_2)$, $a_1 = (\alpha_1+\alpha_2+\alpha_3)$, $a_2 = -\alpha_1(\beta_2+\beta_3)-\alpha_2(\beta_1+\beta_3)-\alpha_3(\beta_1+\beta_2)$, $a_3 = -\alpha_1\beta_2\beta_3-\alpha_2\beta_1\beta_3-\alpha_3\beta_1\beta_2$, $b_1 = \beta_1+\beta_2+\beta_3$, $b_2 = -\beta_1\beta_2-\beta_2\beta_3-\beta_1\beta_3$, and $b_3 = -\beta_1\beta_2\beta_3$.

Contrary to the CGARCH(2), the relaxation of constraints on the CGARCH(N) parameters resulting from the GARCH(N, N) specification for $N \geqslant 3$ is not obvious even in the sense of Nelson and Cao [10] which does not provide an explicit formulation of these constraints for GARCHs of orders $p, q \geqslant 3$. Although the general form of the GARCH model associated with a given CGARCH model had been established by Karanasos [9], the relaxed formulation of the conditions of non-negativity of σ_t^2 for any $N \geqslant 1$ remains an open question. Consequently, the determination of a suitable parameter space for the QML estimation, is only possible ex-post via its GARCH representation. Thus, the pre-estimation of σ_t^2 is only ex-post feasible in order to construct the quasi-likelihood.

Settar et al. [12] put forward a new approach to deal with the non-negativity of the conditional variance generated by the GARCH(1,1) model without any prior constraints on the parameters except for those of weak stationary and the existence of moments. To this aim, a constrained Kalman filter via a state space representation is implemented to predict the conditional variance which is used to estimate the quasi-likelihood function. GARCH parameters are subsequently estimated by the quasi-maximum likelihood using the simultaneous perturbation stochastic approximation (SPSA) [13–15]. This method is called robust estimation of the conditional variance and the parameter estimation is denoted QCKSA. Besides, the use of such kind of optimization algorithm is due to the randomness of the quasi-likelihood construction. But his relevance goes beyond that as its performance does not depend on the number of parameters. In what therefore context does the method proposed in Settar et al. [12] overlap with the estimation of the CGARCH model? Indeed, it is interesting to note that

(i) The CGARCH belongs foremost to the GARCH family being a restricted GARCH(N, N) whose parameters are given according to the parameters of the N GARCH(1,1) components [4, 9]. Hence the QCKSA method is as valid for the estimation of the CGARCH as the GARCH.

(ii) As claimed previously, the CGARCH is not only subject to the non-negativity constraints of its parameters ensuring the positivity of its conditional variance [2], but these constraints can be ex-post relaxed in the sense of Nelson and Cao [10] after the derivation of its GARCH representation. This shows the usefulness of the robust estimation of the conditional variance proposed by Settar et al. [12], by imposing directly the non-negativity of the conditional variance without constraining the model parameters a priori.

(iii) It is well known that the more the conditional variance of the CGARCH is decomposed into more volatilities, the more significant the model is by better capturing the volatility dynamics. This consequently leads to a higher number of parameters. Hence the interest of the SPSA algorithm in the estimation of the CGARCH independently of its dimension.

Thus, within the framework of the GARCH estimation (i), given on the one hand the eventual high order of the CGARCH resulting from (iii) and the requirement to identify the constraints on the parameters relaxing the CGARCH as mentioned in (ii) on the other hand, the prior determination of the parameter space ensuring the convergence of the SPSA algorithm applied to the CGARCH may be flexible so as to avoid the ex-post identification of the parameter space resulting from the GARCH specification of the CGARCH model. In this vein, we show in this paper that the construction of such a space is always feasible as the product of the parameter spaces of each GARCH(1,1) component.

This paper is organized as follows. The next section reviews the QCKSA approach steps. Section 3 provides results of convergence analyses of the QCKSA algorithm applied to the CGARCH model (5). Section 4 examines the performance of our method via Monte Carlo experiment. Further, empirical application to the S&P500 return series is presented in Section 5. Finally, Section 6 concludes.

2. Preliminary

Consider the i-th GARCH(1,1) specification associated with the conditional variance component $\sigma_{i,t}^2$, $i = 1, \ldots, N$. Thus, we outline succinctly the steps of the QML estimation of the i-th GARCH(1,1) parameter vector $\theta_i = {}^t(\omega_i, \alpha_i, \beta_i) \in \Theta_i \subset \mathbb{R}^{+*} \times \mathbb{R}^{+2}$ based on the the robust estimation of the conditional variance and the SPSA algorithm [12]. Henceforth, we respectively denote $\widehat{\sigma}_{i,t}^2 : \Theta_i \to \mathbb{R}$

and $\widetilde{\sigma}^2_{i,t} : \Theta_i \to \mathbb{R}^{+*}$ the conditional variance predicted by the Kalman filter and its robust pre-estimation, of unknown parameter θ_i.

Conditions of the weak stationary and the existence of the fourth moment that define Θ_i are respectively given by

$$\alpha_i + \beta_i < 1, \tag{6}$$

$$\beta_i^2 + 2\alpha_i\beta_i + 3\alpha_i^2 < 1. \tag{7}$$

Under these assumptions we pre-estimate $\sigma^2_{i,t}$ using the Kalman filter via the following state space representation:

$$\begin{aligned}\sigma^2_{i,t} &= \omega_i + (\alpha_i + \beta_i)\sigma^2_{i,t-1} + \alpha_i\nu_{i,t-1},\\ \varepsilon_t^2 &= \sigma^2_{i,t} + \nu_{i,t},\end{aligned}$$

where $\nu_{i,t} = \varepsilon_t^2 - \sigma^2_{i,t} = \varepsilon_t^2 - \mathbb{E}(\varepsilon_t^2|\varepsilon_{t-1}^2, \ldots, \varepsilon_1^2)$ is by definition the linear innovation of the i-th GARCH(1,1) since from the specification (5), $\sigma^2_{i,t}$ is the forecast of ε_t^2 based on its own lagged values.

Then, under the initial conditions:

$$\widehat{\sigma}^2_{i,0|0} = \mathbb{E}(\sigma^2_{i,0}) = \frac{\omega_i}{1 - \alpha_i - \beta_i},$$

and

$$P_{0|0} = \mathrm{Var}(\sigma^2_{i,0}) = \frac{2\omega_i^2\alpha_i^2}{(1 - \alpha_i - \beta_i)^2(1 - \beta_i^2 - 2\alpha_i\beta_i - 3\alpha_i^2)}.$$

We get the Kalman filter equations allowing to predict in a first step $\sigma^2_{i,t}$ by $\widehat{\sigma}^2_{i,t|t-1}$, that are

$$\widehat{\sigma}^2_{i,t|t-1}(\theta_i) = \omega_i + (\alpha_i + \beta_i)\widehat{\sigma}^2_{i,t-1|t-1}(\theta_i), \tag{8}$$

$$P_{t|t-1}(\theta_i) = (\alpha_i + \beta_i)^2 P_{t-1|t-1}(\theta_i) + \alpha_i^2\nu(\theta_i), \tag{9}$$

$$K_t(\theta_i) = P_{t|t-1}(\theta_i)(P_{t|t-1}(\theta_i) + \nu(\theta_i))^{-1}, \tag{10}$$

$$\widehat{\sigma}^2_{i,t|t}(\theta_i) = \widehat{\sigma}^2_{i,t|t-1}(\theta_i) + K_t(\theta_i)(\varepsilon_t^2 - \widehat{\sigma}^2_{i,t|t-1}(\theta_i)), \tag{11}$$

$$P_{t|t}(\theta_i) = (1 - K_t(\theta_i))P_{t|t-1}(\theta_i). \tag{12}$$

In a second step, we apply the density truncation algorithm 1 under the non-negativity constraint $\frac{1}{N_{1-\tau}} \leqslant \sigma_t^2 \leqslant N_{1-\tau}$ which depends on a threshold $N_{1-\tau}$ given for a confidence level $1 - \tau$ and for any $t = 1, \ldots, n$, as

$$1 - \tau = \mathbb{P}\left\{\frac{1}{N_{1-\tau}} \leqslant \sigma_t^2 \leqslant N_{1-\tau}\right\} \tag{13}$$

One obtains a non-negative robust estimate of $\sigma^2_{i,t}(\theta_i)$, namely $\widetilde{\sigma}^2_{i,t}(\theta_i)$ given by

$$\widetilde{\sigma}^2_{i,t} = \sqrt{P_{t|t-1}}\,\mu_\Sigma + \widehat{\sigma}^2_{i,t|t-1}. \tag{14}$$

Then, using the SPSA algorithm 2 for $p = q = 1$ (see also Allal and Benmoumen [13] for the standard GARCH(1,1)), a QCKSA estimate of $\theta_i \in \Theta_i$ is given by

$$\widehat{\theta}_i = \arg\min_{\theta_i\in\Theta_i} \widehat{l}_{i,n}(\theta_i),$$

where

$$\widehat{l}_{i,n}(\theta_i) = \frac{1}{n}\sum_{t=1}^{n}\frac{\varepsilon_t^2}{\widetilde{\sigma}^2_{i,t}(\theta_i)} + \log\left(\widetilde{\sigma}^2_{i,t}(\theta_i)\right) \tag{15}$$

and

$$\Theta_i = \left\{\theta_i \in \mathbb{R}^3 /\ \omega_i > 0,\ \alpha_i + \beta_i < 1,\ \beta_i^2 + 2\alpha_i\beta_i + 3\alpha_i^2 < 1\right\}.$$

Numerically, $\widehat{\theta}_i$ is obtained as the solution of the recursion equation given in (2) by

$$\widehat{\theta}_{k+1} = \widehat{\theta}_k - a_k\widehat{g}(\widehat{\theta}_k). \tag{16}$$

Algorithm 1 Constrained Kalman filtering by Pdf truncation

Require: $\tau = 0.5\%$, $Z_{0.995} = 2.575$: The $(1-\tau)^{th}$ quantile of the standard Gaussian distribution;
1: **for** $t = 1, \dots, n$
2: $\sigma^2_{t|t-1} \sim N(\widehat{\sigma}^2_{t|t-1}, P_{t|t-1})$;
3: $\Sigma_t = \dfrac{\sigma^2_{t|t-1} - \widehat{\sigma}^2_{t|t-1}}{\sqrt{P_{t|t-1}}} \sim N(0,1)$;
4: $N_{1-\tau} = \sqrt{P_{t|t-1}}\, Z_{0.995} + \widehat{\sigma}^2_{t|t-1}$;
5: $l_t = \dfrac{1 - N_{1-\tau}\widehat{\sigma}^2_{t|t-1}}{N_{1-\tau}\sqrt{P_{t|t-1}}}$ and $u_t = \dfrac{N_{1-\tau} - \widehat{\sigma}^2_{t|t-1}}{\sqrt{P_{t|t-1}}}$;
6: $\widetilde{f}(x) = \dfrac{\sqrt{2}}{\sqrt{\pi}\left[\mathrm{erf}(u_t/\sqrt{2}) - \mathrm{erf}(l_t/\sqrt{2})\right]} \exp\left(-x^2/2\right) 1_{[l_t,u_t]}(x)$, $\widetilde{f}$: The Gaussian truncated density of Σ_t normalized between l_t and u_t. erf(.): the truncation error function
7: $\mu_\Sigma = \mathbb{E}_{\widetilde{f}}(\Sigma_t)$;
8: $\widetilde{\sigma}^2_{t|t-1} = \sqrt{P_{t|t-1}}\, \mu_\Sigma + \widehat{\sigma}^2_{t|t-1}$.

Algorithm 2 Simultaneous Perturbation Stochastic Approximation

Require: $(a, c, \lambda) = (0.16, 0.2, 0.602)$ and $A = I/10$, I: number of iterations;
1: initialization: $\theta_0 \in \Theta_1$;
2: **for** $k = 1, \dots, I$
3: $a_k = a(A + k + 1)^{-\lambda}$;
4: $\{\Delta_{k,l}\}_l \sim \text{iidBer}\left(\pm 1, \dfrac{1}{2}\right)$, $\Delta_k = {}^t(\Delta_{k,1}, \dots, \Delta_{k,p+q+1}) \in \mathbb{R}^{p+q+1}$;
5: **if** $\theta_k = {}^t(\omega_{1,k}, \alpha_{1,k}, \beta_{1,k}) \in \Theta_1$ **then**
6: $\{\delta_k^-, \delta_k^+\} \sim \text{iidU}\,[0,1]$;
7: $y_k^+(\widehat{\theta}_k) = \widehat{l}_{i,n}(\widehat{\theta}_k + c\Delta_k) + \delta_k^+$ and $y_k^-(\widehat{\theta}_k) = \widehat{l}_{i,n}(\widehat{\theta}_k - c\Delta_k) + \delta_k^-$;
8: $\widehat{g}(\widehat{\theta}_k) = \dfrac{y_k^+(\widehat{\theta}_k) - y_k^-(\widehat{\theta}_k)}{2c}\ {}^t\left(\Delta^{-1}_{k,1}, \dots, \Delta^{-1}_{k,p+q+1}\right)$;
9: $\widehat{\theta}_{k+1} = \widehat{\theta}_k - a_k \widehat{g}(\widehat{\theta}_k)$;
10: **if** $\|\widehat{\theta}_{k+1} - \widehat{\theta}_k\| > \varepsilon$ **then**
11: $k = k + 1$;
12: **else**
13: Return $\widehat{\theta}_k$;
14: **else**
15: $\widehat{\theta}_k = \widehat{\theta}_{k-1}$;

3. Sketch of convergence

The purpose of this section is to prove the existence of a QML estimate of the CGARCH(N) parameters by specifying a parameter space providing the convergence of the SPSA algorithm. Before that, we demonstrate such convergence for the i-th GARCH(1,1), $i = 1, \dots, N$.

3.1. GARCH(1,1) model

In order to analyse the convergence of the SPSA algorithm applied to the GARCH(1,1), we refer to the assumptions **(5.1)**–**(5.6)** [16].

First, we note that the assumptions **(5.2)**–**(5.4)** are satisfied by the choices of a_k, Δ_k and δ_k made respectively at steps 3, 4 and 6 ([16], p. 45). Assumption **(5.5)** can be practically satisfied through ex-post checks made in step 5 to avoid the divergence of the sequence of iterations (θ_k) in cases of non-stationary or non-existence of moments. Returning to the assumption **(5.1)**, an idea that allows to satisfy it is to extract from Θ_i the smallest compact (bounded and closed subset of Θ_i) over which $\widehat{l}_{i,n}$ is smooth in the sense of **(5.1)**. Thus, we need to assume that ω_i is bounded, i.e. $\eta \leqslant \omega_i \leqslant \overline{\omega}_i$ for some $\eta > 0$ and $\overline{\omega}_i > 0$.

Therefore, we consider the subset $\Theta_{i,\eta} \subset \Theta_i$ defined by

$$\Theta_{i,\eta} = \left\{\theta_i \in \mathbb{R}^3 / \ \eta \leqslant \omega_i \leqslant \overline{\omega}_i, \ \alpha_i + \beta_i \leqslant 1 - \eta, \ \beta_i^2 + 2\alpha_i\beta_i + 3\alpha_i^2 \leqslant 1 - \eta\right\}.$$

It follows that $\Theta_{i,\eta}$ is compact since it is both closed as a union of closed intervals of $\mathbb{R}^3$, and is bounded since for any $\theta_i \in \Theta_{i,\eta}$, $0 < \omega_i \leqslant \overline{\omega}_i$, $0 < \alpha_i < 1$ et $0 < \beta_i < 1$. Moreover, $\Theta_{i,\eta} \to \Theta_i$ as $\eta \to 0^+$.

Lemma 1. *For all $\eta > 0$, $\widehat{l}_{i,n}$ is smooth over $\Theta_{i,\eta}$ in the sense of **(5.1)**.*

Proof. Let $\eta > 0$ and let's respectively denote by ϕ_t, ψ_t and ν the functions defined over $\Theta_{i,\eta}$, and for any $t \in \mathbb{N}^*$ by

$$\phi_t(\theta_i) = \widehat{\sigma}^2_{i,t|t-1}(\theta_i), \quad \psi_t(\theta_i) = P_{t|t-1}(\theta_i) \quad \text{and} \quad \nu(\theta_i) = \mathbb{E}\nu_{i,t},$$

where [2]

$$\nu(\theta_i) = \frac{2\omega_i^2\left(1 + \alpha_i + \beta_i\right)}{\left(1 - \alpha_i - \beta_i\right)\left(1 - 3\alpha_i^2 - \beta_i^2 - 2\alpha_i\beta_i\right)}. \tag{17}$$

From (8), (10) and (11), one easily obtains that

$$\phi_t(\theta_i) = \omega_i + \frac{\left(\alpha_i + \beta_i\right)\left(\nu(\theta_i)\phi_{t-1}(\theta_i) + \psi_{t-1}(\theta_i)\varepsilon^2_{t-1}\right)}{\psi_{t-1}(\theta_i) + \nu(\theta_i)}. \tag{18}$$

Likewise, it is deduced from (9), (10) and (12) that

$$\psi_t(\theta_i) = \left[(\alpha_i + \beta_i)^2 \frac{\psi_{t-1}(\theta_i)}{\psi_{t-1}(\theta_i) + \nu(\theta_i)} + \alpha_i^2\right]\nu(\theta_i). \tag{19}$$

First, we prove by mathematical induction that

$$\psi_t \in \mathcal{C}^\infty(\Theta_{i,\eta}), \ \ \forall\, t \in \mathbb{N}^*. \tag{20}$$

Indeed, for $t = 1$, $\psi_1 \in \mathcal{C}^\infty(\Theta_{i,\eta})$ is a function of $P_{1|0}$ being, according to (9), a function of $P_{0|0} \in \mathcal{C}^\infty(\Theta_{i,\eta})$ given (6) and (7).

Now assume that $\psi_t \in \mathcal{C}^\infty(\Theta_{i,\eta})$. Then, under (6) and (7),

$$\nu \in \mathcal{C}^\infty(\Theta_{i,\eta}). \tag{21}$$

Since

$$\psi_t(\theta_i) + \nu(\theta_i) > 0, \ \ \forall\, \theta_i \in \Theta_{i,\eta}. \tag{22}$$

Then

$$\psi_{t+1} \in \mathcal{C}^\infty(\Theta_{i,\eta}).$$

Hence the result (20).

Now, let's show by mathematical induction that

$$\phi_t \in \mathcal{C}^\infty(\Theta_{i,\eta}) \ \ \forall \ \ t \in \mathbb{N}^*. \tag{23}$$

Indeed, for $t = 1$, $\phi_1 \in \mathcal{C}^\infty(\Theta_{i,\eta})$ is as a function of $\widehat{\sigma}^2_{i,1|0}$ being according to (8), a function of $\widehat{\sigma}^2_{i,0|0} \in \mathcal{C}^\infty(\Theta_{i,\eta})$ given (6).

Assuming that $\phi_t \in \mathcal{C}^\infty(\Theta_\eta)$, given (20), (21) and (22), one can deduce that $\phi_{t+1} \in \mathcal{C}^\infty(\Theta_{i,\eta})$. From which results (23).

Therefore, according to (14), it follows the next

$$\widetilde{\sigma}^2_{i,t|t-1} \in \mathcal{C}^\infty(\Theta_{i,\eta}). \tag{24}$$

By deduction, given (15) the next is true

$$\widehat{l}_{i,n} \in \mathcal{C}^\infty(\Theta_{i,\eta}).$$

Since $\Theta_{i,\eta}$ is a compact, then $\widehat{l}_{i,n}$ is quite smooth over $\Theta_{i,\eta}$ in the sense of **(5.1)**. ■

Based on the above, the following proposition can be formulated ensuring the almost sure convergence of the recursive equation (16).

Proposition 1. Let H be the set of local minima of $\widehat{l}_{i,n}$ satisfying **(5.6)**. Then for all $\eta > 0$, there is $\overline{c} > 0$ such as for all $c \in]0, \overline{c}]$,

$$\theta_i \to H_\eta \quad \text{a.s.},$$

where H_η is a H-neighbour given by

$$H_\eta = \{\theta_i \in \Theta_{i,\eta} / \ \|\theta_i - \theta_i^*\| < \eta,\ \theta_i^* \in H\}.$$

Proof. Theorem 5.3 in [16] is applied to $\widehat{l}_{i,n}$ by checking the assumptions **(5.1)–(5.5)**. Indeed, $\widehat{l}_{i,n}$ satisfies the assumption **(5.1)** according to the lemma (1). Further, the assumptions **(5.2)–(5.5)** are verified by the choice of the sequences a_k, Δ_k and $\delta_k^{\pm}$ made previously ([16], p. 45). Thus, assuming **(5.6)**, the claim follows. ■

3.2. CGARCH(N) model, $N \geqslant 1$.

Now, the existence of a QCKSA estimate of each GARCH(1,1) parameter composing the CGARCH(N) is proved over $\Theta_{i,\eta}$ for $\eta > 0$ and $i = 1, \ldots, N$. In the sequel, the quasi-log likelihood of the CGARCH(N) is maximized according to the parameter space $\widetilde{\Theta}_\eta = \prod_{i=1}^N \Theta_{i,\eta},\ \eta > 0$.

Since the quasi-log likelihood is given for all $\theta = {}^t(\omega_1, \alpha_1, \beta_1, \ldots, \omega_i, \alpha_i, \beta_i, \ldots, \omega_N, \alpha_N, \beta_N) \in \widetilde{\Theta}_\eta$ by

$$\widehat{\mathcal{L}}_n(\theta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\left(\frac{1}{n}\sum_{t=1}^n \frac{\varepsilon_t^2}{\widetilde{\sigma}^2_{t|t-1}(\theta)} + \log\left(\widetilde{\sigma}^2_{t|t-1}(\theta)\right)\right).$$

Then, maximizing $\widehat{\mathcal{L}}_n$ means in other words minimizing the criterion $\widehat{l}_n$ given by

$$\widehat{l}_n(\theta) = \frac{1}{n}\sum_{t=1}^n \frac{\varepsilon_t^2}{\widehat{\sigma}_t^2(\theta)} + \log(\widehat{\sigma}_t^2(\theta)), \tag{25}$$

where $\widehat{\sigma}_t^2$ is the conditional variance estimated for any $\theta \in \widetilde{\Theta}_\eta$ by

$$\widehat{\sigma}_t^2(\theta) = \sum_{i=1}^N \widetilde{\sigma}^2_{i,t}(\theta_i). \tag{26}$$

By construction $\widehat{\sigma}_t^2$ is non-negative one so that $\widehat{l}_n$ is well defined $\widetilde{\Theta}_\eta$.

Remark 1. (26) is in line with our approach based on the estimation of CGARCH model without the need to its GARCH representation, because otherwise the pre-estimation of $\sigma^2_{i,t}$ for the GARCH(1,1) [12] remains applicable to such a representation to pre-estimate $\widehat{\sigma}_t^2$, which does not correspond to the aim of our approach. Hence, we don't have to apply the algorithm 1 but we exploit directly the algorithm 2 to minimize $\widehat{l}_n$ with respect to $\theta = {}^t(\omega_1, \alpha_1, \beta_1, \ldots, \omega_i, \alpha_i, \beta_i, \ldots, \omega_N, \alpha_N, \beta_N)$, by substituting Θ_1 by $\widetilde{\Theta}$ for some $\eta > 0$ and setting $p = q = N$ as described by the following algorithm.

Algorithm 3 QCKSA

1: **for** $i = 1, \ldots, N$
2: Apply the algorithm 1 to compute the constrained conditional variances $\widetilde{\sigma}^2_{i,t}$;
3: Apply the algorithm 2 to $\widehat{l}_{i,n}$ to obtain the SPSA-estimation $\widehat{\theta_i} = \arg\min_{\theta_i \in \Theta_{i,\eta}} \widehat{l}_{i,n}(\theta_i)$;
4: **for** $t = 1, \ldots, n$
5: $\widehat{\sigma}^2_t(\theta) = \sum_{i=1}^{N} \widetilde{\sigma}^2_{i,t}(\widehat{\theta}_i)$;
6: Apply the algorithm 2 to $\widehat{l}_n$ to obtain the estimate $\widehat{\theta} = \arg\min_{\theta \in \widetilde{\Theta}_\eta} \widehat{l}_n(\theta)$.

The following Lemma provides the smoothness of $\widehat{l}_n$ over $\widetilde{\Theta}_\eta$ which verifies **(5.1)**.

Lemma 2. *For all $\eta > 0$, $\widehat{l}_n$ is smooth over $\widetilde{\Theta}_\eta$ in the sense of **(5.1)**.*

Proof. Let $\eta > 0$. It follows from (24) and (26) that $\widehat{\sigma}^2_t \in \mathcal{C}^\infty(\widetilde{\Theta}_\eta)$ as a sum of functions that belong to $\mathcal{C}^\infty$ class over the portions of $\widetilde{\Theta}_\eta$.

Therefore, given (25), one can obtain that $\widehat{l}_n \in \mathcal{C}^\infty(\widetilde{\Theta}_\eta)$. Moreover, $\widetilde{\Theta}_\eta$ is a compact as a product of the $(\Theta_{i,\eta})_{i=1,\ldots,N}$ compacts. Hence the result. ■

The following proposal leads to minimize locally $\widehat{l}_n$ over $\widetilde{\Theta}_\eta$.

Proposition 2. Let L be the set of local minima of $\widehat{l}_n$ verifying **(5.6)**. Then, for any $\eta > 0$, there is $\overline{c}' > 0$ such that for any $c \in \,]0, \overline{c}']$,

$$\theta \to L_\eta \quad \text{a.s.},$$

where L_η is a η-neighbour of L given by

$$L_\eta = \left\{ \theta \in \widetilde{\Theta}_\eta / \ \|\theta - \theta^*\| < \eta, \quad \theta^* \in L \right\}.$$

Proof. Assuming **(5.6)**, it is reasonable to apply this time Theorem 5.3 in [16] to $\widehat{l}_n$ since the latter verifies **(5.1)** according to the lemma (2) and **(5.2)**–**(5.5)** are satisfied by the choice of the sequences a_k, Δ_k and $\delta^{\pm}_k$ made previously. Hence the result. ■

Remark 2. θ^* still possesses the asymptotic properties of a quasi-maximum likelihood estimate under the assumptions **A1–A6** given in Francq and Zakoian ([17], Chapter 7).

4. Simulation experiment

Our aim in this section is to check through a Monte Carlo experiment that the proposed algorithms improve the estimations obtained by quasi-maximum likelihood [4] for a large size of observations.

Let's consider the CGARCH(2) model:

$$\varepsilon_t = \left(\sigma^2_{1,t} + \sigma^2_{2,t}\right)^{\frac{1}{2}} \eta_t = \sigma_t \eta_t, \quad \eta_t \sim \text{iidN}(0,1), \tag{27}$$

$$\sigma^2_{1,t} = 0.005 + 0.04\varepsilon^2_{t-1} + 0.9\sigma^2_{1,t-1}, \tag{28}$$

$$\sigma^2_{2,t} = 0.5 + 0.4\varepsilon^2_{t-1} + 0.3\sigma^2_{2,t-1}. \tag{29}$$

Note that the GARCH(1,1) specifications (28) and (29) are both weak stationary and of finite 4-th moments according respectively to conditions (6) and (7).

Then, there are generated 150 replications of sample size $n = 10000$. First, we estimate the CGARCH parameters using the QML method [4]. The obtained parameter estimates denoted $\widehat{\theta}_{QML}$ are used as initial values of θ to apply the QCKSA algorithm 3 to the generated data whereby algorithms 1 and 2 are applied to pre-estimate $\sigma^2_{1,t}$ and $\sigma^2_{2,t}$ from which we obtain a pre-estimation of σ^2_t as given by (26). Afterwards, the likelihood criterion $\widehat{l}_n$ (25) is minimized using algorithm 2. As claimed earlier, the identification of the GARCH(2,2) specification resulting from the CGARCH(2) (27) is not required, nor the resulting relaxed restrictions on the parameters. The obtained parameter estimator is denoted $\widehat{\theta}_{QCKSA}$. The mean squared error (MSE) is used to compare the accuracy of the parameter estimations. As can be clearly seen from Table 1, the QCKSA method has outperformed the the QML method by recording the smallest MSE values.

Table 1. Estimation accuracy of CGARCH(2) parameters by QCKSA and QML.

θ	ω_1	α_1	β_1	ω_2	α_2	β_2
θ_0	0.005	0.04	0.9	0.5	0.4	0.3
$\widehat{\theta}_{QCKSA}$	0.0047	0.0381	0.8905	0.4991	0.4003	0.2921
MSE	< 0.0001	< 0.0001	0.0001	< 0.0001	< 0.0001	0.0001
$\widehat{\theta}_{QML}$	0.0179	0.0301	0.7418	0.6810	0.3702	0.3856
MSE	0.1650	0.0031	0.0409	0.0398	0.0010	0.0074

5. Empirical illustration

There is presented an empirical application of the QCKSA method for the estimation of CGARCH parameters according to the specification (5). The application concerns the series of daily returns of the S&P500 index over the period from 28/10/2010 to 27/11/2020 with a sample size $n = 2539$ observations (figure 1). The data used for the analysis can be freely downloaded from the website `http://finance.yahoo.com`.

Table 2. Descriptive statistics of S&P500 returns. $Q_r(10)$ and $Q_{r^2}(10)$ are Ljung-Box's Q-statistics at lag 10.

Min	Max	Mean	Variance	kurtosis	$Q_r(10)$	$Q_{r^2}(10)$
-5.5438	3.8948	0.0192	0.2289	20.8808	66.817	602.29

The statistics reported in Table 2 reveal the presence of the standard features of excess kurtosis and significant autocorrelation of the series of returns and squares of returns as shown by the Q-statistics of Ljung–Box (e.g., at order 10). Thus, the series of returns denoted r_t is fitted by an AR(1) with a CGARCH(2) noise, i.e.,

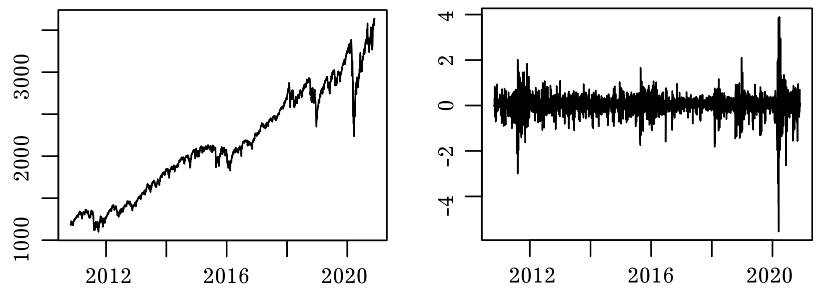

Fig. 1. S&P500 index (left) and the corresponding series of returns (right) from October 28, 2010 to November 27, 2020.

$$r_t = \phi_0 + \phi_1 r_{t-1} + \varepsilon_t, \quad (\varepsilon_t|\varepsilon_{t-1}, \varepsilon_{t-2}, \dots \varepsilon_1) \sim \text{iidN}(0, \sigma_t^2) \tag{30}$$

$$\varepsilon_t = (\sigma_{1,t}^2 + \sigma_{2,t}^2)^{\frac{1}{2}} \eta_t = \sigma_t \eta_t, \quad \eta_t \sim \text{iidN}(0,1)$$

$$\sigma_{i,t}^2 = \omega_i + \alpha_i \varepsilon_{t-1}^2 + \beta_i \sigma_{i,t-1}^2, \quad i = 1, 2 \tag{31}$$

Firstly, there are estimated by QML the parameters of the CGARCH given by (30), namely ω_1, α_1, β_1, ω_2, α_2 and β_2. Then, there are estimated by QCKSA the CGARCH(2) parameters as given by the algorithm 3 using the QML parameter estimates as initial values. Table 3 represents the estimates obtained in addition to the corresponding log-likelihood values.

The estimation results presented in Table 3 show that the $\sigma_{1,t}^2$ component captures long term volatility fluctuations while the second component $\sigma_{2,t}^2$ captures the short-term volatility fluctuations. Indeed, using the QML (resp. QCKSA) method, $\sigma_{1,t}^2$ starts small in amplitude (0.0079 (resp. 0.0062)),

Table 3. Estimation of CGARCH(2) parameters by QML and QCKSA. Log-lik represents the log-lik likelihood.

	AR(1)		CGARCH(2)						
	$\widehat{\phi}_0$	$\widehat{\phi}_1$	$\widehat{\omega}_1$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\omega}_2$	$\widehat{\alpha}_2$	$\widehat{\beta}_2$	log-Lik
QML	0.0191	−0.1621	0.0079	0.0007	0.9944	0.0877	0.2118	0.7355	−992.6334
QCKSA	0.0191	−0.1621	0.0062	0.0004	0.9913	0.0833	0.1821	0.7351	−985.7104

but decreases very slowly with a decrease rate of 0.9944 (resp. 0.9913), while $\sigma^2_{2,t}$ starts strongly in amplitude (0.0877 (resp. 0.0833)), but decreases very fast with a decrease rate of 0.7355 (resp. 0.7351) (Figures 2 and 3). This is also in line with the persistence estimates for each component, given by $\widehat{\alpha}_i + \widehat{\beta}_i$, $i = 1, 2$, whereby the $\sigma^2_{1,t}$ component captures the persistence effect of σ^2_t recording 0.9973 (resp. 0.9917) exceeding that of $\sigma^2_{2,t}$ which records 0.9473 (resp. 0.9172). Moreover, $\sigma^2_{1,t}$ is much less sensitive to shocks ε^2_{t-1} than $\sigma^2_{2,t}$ ($0.0007 < 0.2118$ and $0.0004 < 0.1821$).

Note in particular that the QCKSA estimation distinguishes better the behaviour of σ^2_t in terms of volatility persistence for each component since it reflects better the low persistence of $\sigma^2_{2,t}$ compared to that obtained by QML ($0.9172 < 0.9473$). Thus, it provided results in accordance with the structure of the CGARCH capturing both the short and long term behaviour of the volatility, while also recording a relatively higher likelihood value ($-985.7104 > -992.6334$).

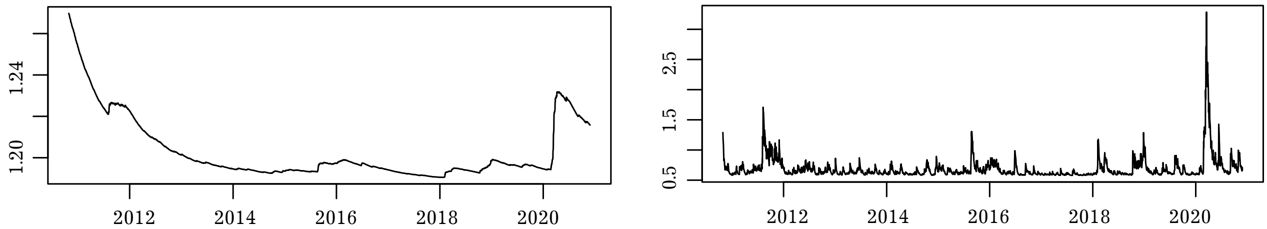

Fig. 2. Volatility of the long-term component $\sigma^2_{1,t}$ (left) and the short-term component $\sigma^2_{2,t}$ (right) estimated by QML.

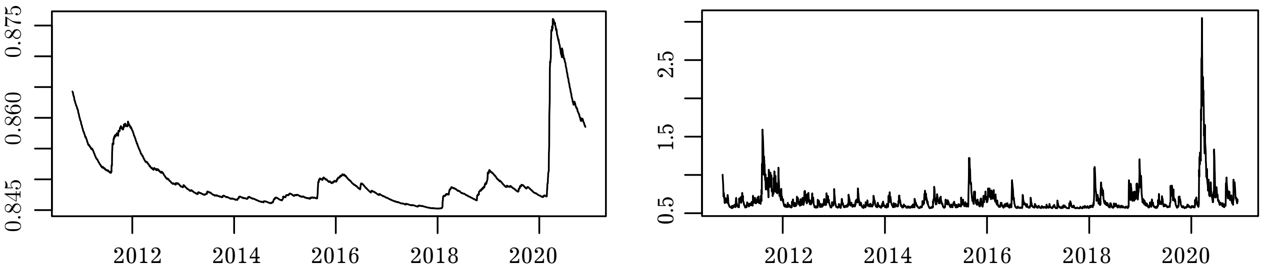

Fig. 3. Volatility of the long-term component $\sigma^2_{1,t}$ (left) and the short-term component $\sigma^2_{2,t}$ (right) estimated by QCKSA.

6. Conclusion

In this paper, the convergence of the SPSA algorithm applied to the CGRACH has been proven in a parameter space defined a priori from the parameter spaces of its GARCH(1,1) components allowing the pre-estimation of its conditional variance and satisfying the relaxed constraints of Nelson and Cao. The simulation experiment show that our estimation method outperforms the existing QML approach in term of the accuracy of the parameter estimation. Further, the empirical study of the S&P500 return series showed the performance of the QCKSA estimation to capture the short and long term dynamics of the CGARCH volatility, by maximizing its likelihood compared to the QML estimation.

[1] Engle R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica: Journal of the econometric society. **50** (4), 987–1007 (1982).

[2] Bollerslev T. Generalized autoregressive conditional heteroskedasticity. Journal of econometrics. **31** (3), 307–327 (1986).

[3] Bollerslev T. On the correlation structure for the generalized autoregressive conditional heteroskedastic process. Journal of Time Series Analysis. **9**, 121–131 (1988).

[4] Ding Z., Granger C. W. Modelling volatility persistence of speculative returns: a new approach. Journal of econometrics. **73** (1), 185–215 (1996).

[5] Ding Z., Granger C. W., Engle R. F. A long memory property of stock market returns and a new model. Journal of empirical finance. **1** (1), 83–106 (1993).

[6] Andersen T. G., Bollerslev T. Heterogeneous information arrivals and return volatility dynamics: Uncovering the longrun in high frequency returns. The journal of Finance. **52**, 975–1005 (1997).

[7] Andersen T. G., Bollerslev T., Diebold F. X. The distribution of realized exchange rate volatility. Journal of the American statistical association. **96** (453), 42–55 (2001).

[8] Bollerslev T., Wright J. H. Semiparametric estimation of long-memory volatility dependencies: The role of highfrequency data. Journal of econometrics. **98** (1), 81–106 (2000).

[9] Karanasos M. The second moment and the autocovariance function of the squared errors of the GARCH model. Journal of Econometrics. **90** (1), 63–76 (1999).

[10] Maheu J. Can GARCH models capture long-range dependence? Studies in Nonlinear Dynamics & Econometrics. **9** (4) (2005).

[11] Engle R. F., Lee G. A long-run and shortrun component model of stock return volatility. Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive WJ Granger. 475 (1999).

[12] Settar A., Idrissi N. F., Badaoui M. New approach in dealing with the non-negativity of the conditional variance in the estimation of GARCH model. Central European Journal of Economic Modelling and Econometrics. **13**, 55 (2021).

[13] Allal J., Benmoumen M. Parameter Estimation for GARCH(1,1) Models Based on Kalman Filter. Advances and Applications in Statistics. **25**, 15 (2011).

[14] Spall J. C. Implementation of the simultaneous perturbation algorithm for stochastic optimization. IEEE Transactions on aerospace and electronic systems. **34** (3), 817–823 (1998).

[15] Spall J. C. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. IEEE transactions on automatic control. **37** (3), 332–341 (1992).

[16] Bhatnagar S., Prashanth H. L., Prashanth L. A. Lecture Notes in Control and Information Sciences. Series Advisory. **434** (2013).

[17] Francq C., Zakoian J. M. GARCH models: structure, statistical inference and financial applications. John Wiley & Sons (2019).

Квазімаксимальна оцінка правдоподібності моделі Component-GARCH за допомогою алгоритму стохастичного наближення із застосуванням до S&P500

Сеттар А.[1], Фатмі Н. І.[1], Бадауї М.[1,2]

[1] *IPIM, Національні школи прикладних наук Хорібги, Університет Султана Мулея Слімана, Марокко*
[2] *LaMSD, Вища технологічна школа Удjди Перший університет Мохаммеда, Марокко*

Компонент GARCH (CGARCH) підходить для кращого відображення короткострокової та довгострокової динаміки волатильності. Тим не менше, простір параметрів, що складається з обмежень невід'ємності умовної дисперсії, нерухомості та існування моментів, є лише попередньо визначеним через представлення GARCH CGARCH. Це пов'язано з відсутністю загального методу визначення апріорі слабких обмежень невід'ємності умовної дисперсії CGARCH(N) для будь-якого $N \geqslant 1$. У цій роботі простір параметрів CGARCH, побудований із просторів параметрів компонента GARCH(1,1), апріорі надається для ідентифікації його форми GARCH. Такий простір виконує слабкі обмеження невід'ємності умовної дисперсії CGARCH, що попередньо оцінюється, забезпечуючи існування оцінки QML у значенні алгоритму стохастичного наближення. Представлено імітаційний експеримент, а також емпіричне застосування до індексу S&P500, і обидва вони показують ефективність запропонованого методу.

Ключові слова: *компонент GARCH, умовна дисперсія, стохастичне наближення, фільтр Кальмана, квазімаксимальна ймовірність.*