Garch models without positivity constraints: exponential or log garch?

Francq, Christian and Wintenberger, Olivier and Zakoian, Jean-Michel

16 September 2012

Garch models without positivity constraints: exponential or log garch?

Christian Francq*

BP 60149, 59653 Villeneuve d'Ascq cedex, France

e-mail: christian.francq@univ-lille3.fr

Olivier Wintenberger

Place du Maréchal de Lattre de Tassigny, 75775 PARIS Cedex 16, France

e-mail: wintenberger@ceremade.dauphine.fr

and

Jean-Michel Zakoïan *

15 Boulevard Gabriel Peri, 92245 Malakoff Cedex, France

e-mail: zakoian@ensae.fr

Abstract: This paper studies the probabilistic properties and the estimation of the asymmetric log-GARCH(p, q) model. In this model, the log-volatility is written as a linear function of past values of the log-squared observations, with coefficients depending on the sign of the observations, and past log-volatility values. Conditions are obtained for the existence of solutions and finiteness of their log-moments. We also study the tail properties of the solution. Under mild assumptions, we show that the quasi-maximum likelihood estimation of the parameters is strongly consistent and asymptotically normal. Simulations illustrating the theoretical results and an application to real financial data are proposed.

Keywords and phrases: EGARCH, log-GARCH, Quasi-Maximum Likelihood, Strict stationarity, Tail index.

*Christian Francq and Jean-Michel Zakoïan gratefully acknowledges financial support of the ANR via the Project ECONOM&RISK (ANR 2010 blanc 1804 03).

1. Preliminaries

Since their introduction by Engle (1982) and Bollerslev (1986), GARCH models have attracted much attention and have been widely investigated in the literature. Many extensions have been suggested and, among them, the EGARCH (Exponential GARCH) introduced and studied by Nelson (1991) is very popular. In this model, the log-volatility is expressed as a linear combination of its past values and past values of the positive and negative parts of the innovations. Two main reasons for the success of this formulation are that (i) it allows for asymmetries in volatility (the so-called leverage effect: negative shocks tend to have more impact on volatility than positive shocks of the same magnitude), and (ii) it does not impose any positivity restrictions on the volatility coefficients.

Another class of GARCH-type models, which received less attention, seems to share the same characteristics. The log-GARCH(p,q) model has been introduced, in slightly different forms, by Geweke (1986), Pantula (1986) and Milhøj (1987). For more recent works on this class of models, the reader is referred to Sucarrat and Escribano (2010) and the references therein. The (asymmetric) log-GARCH(p, q) model takes the form

$$\begin{cases} \epsilon_t &= \sigma_t \eta_t, \\ \log \sigma_t^2 &= \omega + \sum_{i=1}^{q} \left(\alpha_{i+} 1_{\{\epsilon_{t-i}>0\}} + \alpha_{i-} 1_{\{\epsilon_{t-i}<0\}} \right) \log \epsilon_{t-i}^2 \\ & \quad + \sum_{j=1}^{p} \beta_j \log \sigma_{t-j}^2 \end{cases} \tag{1.1}$$

where $\sigma_t > 0$ and (η_t) is a sequence of independent and identically distributed (iid) variables such that $E\eta_0 = 0$ and $E\eta_0^2 = 1$. The usual symmetric log-GARCH corresponds to the case $\boldsymbol{\alpha}_+ = \boldsymbol{\alpha}_-$, with $\boldsymbol{\alpha}_+ = (\alpha_{1+}, \ldots, \alpha_{q+})$ and $\boldsymbol{\alpha}_- = (\alpha_{1-}, \ldots, \alpha_{q-})$.

Interesting features of the log-GARCH specification are the following.

(a) Absence of positivity constraints. An advantage of modeling the log-volatility rather than the volatility is that the vector $\boldsymbol{\theta} = (\omega, \boldsymbol{\alpha}_+, \boldsymbol{\alpha}_-, \boldsymbol{\beta})$

with $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)$ is not a priori subject to positivity constraints[1]. This property seems particularly appealing when exogenous variables are included in the volatility specification (see Sucarrat and Escribano, 2012).

(b) Asymmetries. Except when $\alpha_{i+} = \alpha_{i-}$ for all i, positive and negative past values of ϵ_t have different impact on the current log-volatility, hence on the current volatility. However, given that $\log \epsilon_{t-i}^2$ can be positive or negative, the usual leverage effect does not have a simple characterization, like $\alpha_{i+} < \alpha_{i-}$ say. Other asymmetries could be introduced, for instance by replacing ω by $\sum_{i=1}^q \alpha_{i+} 1_{\{\epsilon_{t-i}>0\}} + \alpha_{i-} 1_{\{\epsilon_{t-i}<0\}}$. The model would thus be stable by scaling, which is not the case of Model (1.1) except in the symmetric case.

(c) The volatility is not bounded below. Contrary to standard GARCH models and most of their extensions, there is no minimum value for the volatility. The existence of such a bound can be problematic because, for instance in a GARCH(1,1), the minimum value is determined by the intercept ω. On the other hand, the unconditional variance is proportional to ω. Log-volatility models allow to disentangle these two properties (minimum value and expected value of the volatility).

(d) Small values can have persistent effects on volatility. In usual GARCH models, a large value (in modulus) of the volatility will be followed by other large values (through the coefficient β in the GARCH(1,1), with standard notation). A sudden rise of returns (in module) will also be followed by large volatility values if the coefficient α is not too small. We thus have persistence of large returns and volatility. But small returns (in module) and small volatilities are not persistent. In a period of large volatility, a sudden drop of the return due to a small innovation, will not much alter the subsequent volatilities (because

[1]However, some desirable properties may determine the sign of coefficients. For instance, the present volatility is generally thought of as an increasing function of its past values, which entails $\beta_j > 0$. The difference with standard GARCH models is that such constraints are not required for the existence of the process and, thus, do not complicate estimation procedures.

β is close to 1 in general). By contrast, as will illustrated in the sequel, the log-GARCH provides persistence of large *and* small values.

(e) Power-invariance of the volatility specification. An interesting potential property of time series models is their stability with respect to certain transformations of the observations. Contemporaneous aggregation and temporal aggregation of GARCH models have, in particular, been studied by several authors (see Drost and Nijman (1993)). On the other hand, the choice of a power-transformation is an issue for the volatility specification. For instance, the volatility can be expressed in terms of past squared values (as in the usual GARCH) or in terms of past absolute values (as in the symmetric TGARCH) but such specifications are incompatible. On the contrary, any power transformation $|\sigma_t|^s$ (for $s \neq 0$) of a log-GARCH volatility has a log-GARCH form (with the same coefficients in $\boldsymbol{\theta}$, except the intercept ω which is multiplied by $s/2$).

The log-GARCH model has apparent similarities with the EGARCH(p, ℓ) model defined by

$$\begin{cases} \epsilon_t &= \sigma_t \eta_t, \\ \log \sigma_t^2 &= \omega + \sum_{j=1}^p \beta_j \log \sigma_{t-j}^2 + \sum_{k=1}^\ell \gamma_k \eta_{t-k}^+ + \delta_k |\eta_{t-k}|, \end{cases} \quad (1.2)$$

under the same assumptions on the sequence (η_t) as in Model (1.1). Indeed, these models have in common the above properties **(a)**, **(b)**, **(c)** and **(e)**. Concerning the property in **(d)**, and more generally the impact of shocks on the volatility dynamics, Figure 1 illustrates the differences between the two models (and also with the standard GARCH). The model coefficients have been chosen to ensure the same long-term variances when the squared innovations are equal to 1. Interestingly, a shock close to zero has a very persistent impact on the log-GARCH volatility, contrary to the EGARCH and GARCH volatilities. However, a *sequence* of shocks close to zero do impact the EGARCH volatility (but not the GARCH volatility).

This article provides a probability and statistical study of the log-GARCH, together with a comparison with the EGARCH. While the stationarity proper-

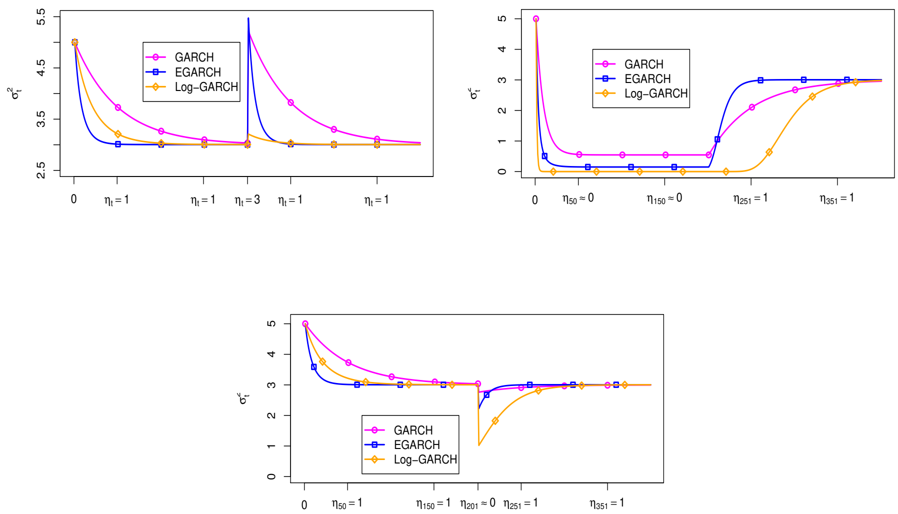

Figure 1. *Curves of the impact of shocks on volatility. The top-left graph shows that a large shock has a (relatively) small impact on the log-GARCH, a large but transitory effect on the EGARCH, and a large and very persistent effect on the classical GARCH volatility. The top-right graph shows the effect of a sequence of tiny innovations: for the log-GARCH, contrary to the GARCH and EGARCH, the effect is persistent. The bottom graph shows that even one tiny innovation causes this persistence of small volatilities for the log-GARCH.*

ties of the EGARCH are well-known, those of the asymmetric log-GARCH(p, q) model (1.1) have not yet been established, to our knowledge. As for the quasi-maximum likelihood estimator (QMLE), the consistency and asymptotic normality have only been proved in particular cases and under cumbersome assumptions for the EGARCH, and have not yet been established for the log-GARCH. Finally, it seems important to compare the two classes of models on typical financial series. The distinctive features of the two models may render each formulation more adequate for certain types of series.

The remainder of the paper is organized as follows. Section 2 studies the existence of a solution to Model (1.1). Conditions for the existence of log-moments are derived, and we characterize the leverage effect. Section 3 is devoted to the tail properties of the solution. In Section 4, the strong consistency and the asymptotic normality of the QMLE are established under mild conditions. Section 6 presents some numerical applications on simulated and real data. Proofs are collected in Section 7. Section 8 concludes.

2. Stationarity, moments and asymmetries of the log-GARCH

We start by studying the existence of solutions to Model (1.1).

2.1. Strict stationarity

Let $\mathbf{0}_k$ denote a k-dimensional vector of zeroes, and let $\boldsymbol{I}_k$ denote the k-dimensional identity matrix. Introducing the vectors

$$\begin{aligned}
\boldsymbol{\epsilon}_{t,q}^{+} &= (1_{\{\epsilon_t>0\}}\log\epsilon_t^2,\dots,1_{\{\epsilon_{t-q+1}>0\}}\log\epsilon_{t-q+1}^2)'\in\mathbb{R}^q,\\
\boldsymbol{\epsilon}_{t,q}^{-} &= (1_{\{\epsilon_t<0\}}\log\epsilon_t^2,\dots,1_{\{\epsilon_{t-q+1}<0\}}\log\epsilon_{t-q+1}^2)'\in\mathbb{R}^q,\\
\boldsymbol{z}_t &= (\boldsymbol{\epsilon}_{t,q}^{+},\boldsymbol{\epsilon}_{t,q}^{-},\log\sigma_t^2,\dots,\log\sigma_{t-p+1}^2)'\in\mathbb{R}^{2q+p},\\
\boldsymbol{b}_t &= \left((\omega+\log\eta_t^2)1_{\{\eta_t>0\}},\mathbf{0}_{q-1}',(\omega+\log\eta_t^2)1_{\{\eta_t<0\}},\mathbf{0}_{q-1}',\omega,\mathbf{0}_{p-1}'\right)'\in\mathbb{R}^{2q+p},
\end{aligned}$$

and the matrix

$$\boldsymbol{C}_t=\begin{pmatrix}
1_{\{\eta_t>0\}}\boldsymbol{\alpha}_+ & 1_{\{\eta_t>0\}}\boldsymbol{\alpha}_- & 1_{\{\eta_t>0\}}\boldsymbol{\beta}\\
\boldsymbol{I}_{q-1}\quad \mathbf{0}_{q-1} & \mathbf{0}_{(q-1)\times q} & \mathbf{0}_{(q-1)\times p}\\
1_{\{\eta_t<0\}}\boldsymbol{\alpha}_+ & 1_{\{\eta_t<0\}}\boldsymbol{\alpha}_- & 1_{\{\eta_t<0\}}\boldsymbol{\beta}\\
\mathbf{0}_{(q-1)\times q} & \boldsymbol{I}_{q-1}\quad \mathbf{0}_{q-1} & \mathbf{0}_{(q-1)\times p}\\
\boldsymbol{\alpha}_+ & \boldsymbol{\alpha}_- & \boldsymbol{\beta}\\
\mathbf{0}_{(p-1)\times q} & \mathbf{0}_{(p-1)\times q} & \boldsymbol{I}_{p-1}\quad \mathbf{0}_{p-1}
\end{pmatrix}, \tag{2.1}$$

we rewrite Model (1.1) in matrix form as

$$\boldsymbol{z}_t=\boldsymbol{C}_t\boldsymbol{z}_{t-1}+\boldsymbol{b}_t. \tag{2.2}$$

We have implicitly assumed $p>1$ and $q>1$ to write $\boldsymbol{C}_t$ and $\boldsymbol{b}_t$, but obvious changes of notation can be employed when $p\le 1$ or $q\le 1$. Let $\gamma(\mathbf{C})$ be the top

Lyapunov exponent of the sequence $\mathbf{C} = \{C_t, t \in \mathbb{Z}\}$,

$$\gamma(\mathbf{C}) = \lim_{t\to\infty} \frac{1}{t} E\left(\log \|\boldsymbol{C}_t \boldsymbol{C}_{t-1} \dots \boldsymbol{C}_1\|\right) = \inf_{t\geq 1} \frac{1}{t} E(\log \|\boldsymbol{C}_t \boldsymbol{C}_{t-1} \dots \boldsymbol{C}_1\|).$$

The choice of the norm is obviously unimportant for the value of the top Lyapunov exponent. However, in the sequel, the matrix norm will be assumed to be multiplicative. Bougerol and Picard (1992a) showed that if an equation of the form (2.2) with iid coefficients $(\boldsymbol{C}_t, \boldsymbol{b}_t)$ is irreducible[2] and if $E \log^+ \|\boldsymbol{C}_0\|$ and $E \log^+ \|\boldsymbol{b}_0\|$ are finite, $\gamma(\mathbf{C}) < 0$ is the necessary and sufficient condition for the existence of a stationary solution to (2.2). Bougerol and Picard (1992b) showed that, for the univariate GARCH(p, q) model, there exists a representation of the form (2.2) with positive coefficients, and for which the necessary and sufficient condition for the existence of a stationary GARCH model is $\gamma(\mathbf{C}) < 0$. The result can be extended to more general classes of GARCH models (see *e.g.* Francq and Zakoïan, 2010a). The problem is more delicate with the log-GARCH because the coefficients of (2.2) are not constrained to be positive. The following result and Remark 2.1 below show that $\gamma(\mathbf{C}) < 0$ is only sufficient. The condition is however necessary under the mild additional assumption that (2.2) is irreducible.

Theorem 2.1. *Assume that $E \log^+ |\log \eta_0^2| < \infty$. A sufficient condition for the existence of a strictly stationary solution to the log-GARCH model (1.1) is $\gamma(\mathbf{C}) < 0$. When $\gamma(\mathbf{C}) < 0$ there exists only one stationary solution, which is non anticipative and ergodic.*

Example 2.1 (The log-GARCH(1,1) case). In the case $p = q = 1$, omitting subscripts, we have

$$\boldsymbol{C}_t \boldsymbol{C}_{t-1} \dots \boldsymbol{C}_1 = \begin{pmatrix} 1_{\{\eta_t > 0\}} \\ 1_{\{\eta_t < 0\}} \\ 1 \end{pmatrix} \begin{pmatrix} \alpha_+ & \alpha_- & \beta \end{pmatrix} \prod_{i=1}^{t-1} \left(\alpha_+ 1_{\{\eta_i > 0\}} + \alpha_- 1_{\{\eta_i < 0\}} + \beta\right).$$

[2] See their Definition 2.3.

Assume that $E\log^+|\log\eta_t^2|<\infty$, which entails $P(\eta_0=0)=0$. Thus,

$$\gamma(\mathbf{C})=E\log\left|\alpha_+1_{\{\eta_0>0\}}+\alpha_-1_{\{\eta_0<0\}}+\beta\right|=\log|\beta+\alpha_+|^a|\beta+\alpha_-|^{1-a},$$

where $a=P(\eta_0>0)$. The condition $|\alpha_++\beta|^a|\alpha_-+\beta|^{1-a}<1$ thus guarantees the existence of a stationary solution to the log-GARCH(1,1) model.

Example 2.2 (The symmetric case). In the case $\boldsymbol{\alpha}_+=\boldsymbol{\alpha}_-=\boldsymbol{\alpha}$, one can see directly from (1.1) that $\log\sigma_t^2$ satisfies an ARMA-type equation of the form

$$\left\{1-\sum_{i=1}^{r}(\alpha_i+\beta_i)\,B^i\right\}\log\sigma_t^2=c+\sum_{i=1}^{q}\alpha_iB^iv_t$$

where B denotes the backshift operator, $v_t=\log\eta_t^2$, $r=\max\{p,q\}$, $\alpha_i=0$ for $i>q$ and $\beta_i=0$ for $i>p$. This equation is a standard ARMA(r,q) equation under the moment condition $E(\log\eta_t^2)^2<\infty$, but this assumption is not needed. It is well known that this equation admits a non degenerated and non anticipative stationary solution if and only if the roots of the AR polynomial lie outside the unit circle.

We now show that this condition is equivalent to the condition $\gamma(\mathbf{C})<0$ in the case $q=1$. Let $\boldsymbol{P}$ be the permutation matrix obtained by permuting the first and second rows of $\boldsymbol{I}_{2+p}$. Note that $\boldsymbol{C}_t=\boldsymbol{C}^+1_{\{\eta_t>0\}}+\boldsymbol{C}^-1_{\{\eta_t<0\}}$ with $\boldsymbol{C}^-=\boldsymbol{P}\boldsymbol{C}^+$. Since $\alpha_+=\alpha_-$, we have $\boldsymbol{C}^+\boldsymbol{P}=\boldsymbol{C}^+$. Thus $\boldsymbol{C}^+\boldsymbol{C}^-=\boldsymbol{C}^+\boldsymbol{P}\boldsymbol{C}^+=\boldsymbol{C}^+\boldsymbol{C}^+$ and $\|\boldsymbol{C}_t\cdots\boldsymbol{C}_1\|=\|(\boldsymbol{C}^+)^t\|$. It follows that $\gamma(\mathbf{C})=\log\rho(\boldsymbol{C}^+)$. In view of the companion form of $\boldsymbol{C}^+$, it can be seen that the condition $\rho(\boldsymbol{C}^+)<1$ is equivalent to the condition $z-\sum_{i=1}^r(\alpha_i+\beta_i)z^i=0\Rightarrow|z|>1$.

Remark 2.1 (The condition $\gamma(\mathbf{C})<0$ is not necessary). Assume for instance that $p=q=1$ and $\alpha_+=\alpha_-=\alpha$. In that case $\gamma(\mathbf{C})<0$ is equivalent to $|\alpha+\beta|<1$. In addition, assume that $\eta_0^2=1$ a.s. Then, when $\alpha+\beta\neq1$, there exists a stationary solution to (1.1) defined by $\epsilon_t=\exp(c/2)\eta_t$, with $c=\omega/(1-\alpha-\beta)$.

2.2. Existence of log-moments

It is well known that for GARCH-type models, the strict stationarity condition entails the existence of a moment of order $s > 0$ for $|\epsilon_t|$. The following Lemma shows that this is also the case for $|\log \epsilon_t^2|$ in the log-GARCH model, when the condition $E \log^+ |\log \eta_0^2| < \infty$ of Theorem 2.1 is slightly reinforced.

Proposition 2.1 (Existence of a fractional log-moment). *Assume that $\gamma(\mathbf{C}) < 0$ and that $E|\log \eta_0^2|^{s_0} < \infty$ for some $s_0 > 0$. Let ϵ_t be the strict stationary solution of (1.1). There exists $s > 0$ such that $E|\log \epsilon_t^2|^s < \infty$ and $E|\log \sigma_t^2|^s < \infty$.*

In order to give conditions for the existence of higher-order moments, we introduce some additional notation. Let $\boldsymbol{e}_i$ be the i-th column of $\boldsymbol{I}_r$, let $\boldsymbol{\sigma}_{t,r} = (\log \sigma_t^2, \dots, \log \sigma_{t-r+1}^2)'$, and let the companion matrix

$$\boldsymbol{A}_t = \begin{pmatrix} \mu_1(\eta_{t-1}) & \dots & \mu_{r-1}(\eta_{t-r+1}) & \mu_r(\eta_{t-r}) \\ & \boldsymbol{I}_{r-1} & & \mathbf{0}_{r-1} \end{pmatrix}, \tag{2.3}$$

where $\mu_i(\eta_t) = \alpha_{i+} 1_{\{\eta_t > 0\}} + \alpha_{i-} 1_{\{\eta_t < 0\}} + \beta_i$ with the convention $\alpha_{i+} = \alpha_{i-} = 0$ for $i > p$ and $\beta_i = 0$ for $i > q$. We have the Markovian representation

$$\boldsymbol{\sigma}_{t,r} = \boldsymbol{A}_t \boldsymbol{\sigma}_{t-1,r} + \boldsymbol{u}_t, \tag{2.4}$$

where $\boldsymbol{u}_t = u_t \boldsymbol{e}_1$, with

$$u_t = \omega + \sum_{i=1}^{q} \left(\alpha_{i+} 1_{\{\eta_{t-i} > 0\}} + \alpha_{i-} 1_{\{\eta_{t-i} < 0\}} \right) \log \eta_{t-i}^2.$$

The Kronecker matrix product is denoted by $\otimes$, and the spectral radius of a square matrix $\boldsymbol{M}$ is denoted by $\rho(\boldsymbol{M})$. For any (random) vector or matrix $\boldsymbol{M}$, let $\mathrm{Abs}(\boldsymbol{M})$ be the matrix, of same size as $\boldsymbol{M}$, whose elements are the absolute values of the corresponding elements of $\boldsymbol{M}$.

Proposition 2.2 (Existence of log-moments). *Let m be a positive integer. Assume that $\gamma(\mathbf{C}) < 0$ and that $E|\log \eta_0^2|^m < \infty$. Let $\boldsymbol{A}^{(m)} = E\mathrm{Abs}(\boldsymbol{A}_1)^{\otimes m}$ where $\boldsymbol{A}_t$ is defined by (2.3).*

- *If $m = 1$ or $r = 1$, then $\rho(\boldsymbol{A}^{(m)}) < 1$ implies that the strict stationary solution of (1.1) is such that $E|\log \epsilon_t^2|^m < \infty$ and $E|\log \sigma_t^2|^m < \infty$.*
- *If $\rho(\boldsymbol{C}^{(m)}) < 1$, then $E|\log \epsilon_t^2|^m < \infty$ and $E|\log \sigma_t^2|^m < \infty$.*

Remark 2.2 (**A sufficient condition for the existence of any log-moment).** Let $\boldsymbol{A}^{(\infty)} = \operatorname{ess\,sup} \operatorname{Abs}(\boldsymbol{A}_1)$ be the essential supremum of $\operatorname{Abs}(\boldsymbol{A}_1)$ term by term. Then, it follows from (7.2) in the proof that componentwise we have

$$\operatorname{Abs}(\boldsymbol{\sigma}_{t,r}) \leq \sum_{\ell=0}^{\infty} (\boldsymbol{A}^{(\infty)})^{\ell} \operatorname{Abs}(\boldsymbol{u}_{t-\ell}). \tag{2.5}$$

Therefore, the condition

$$\rho(\boldsymbol{A}^{(\infty)}) < 1 \tag{2.6}$$

ensures the existence of $E|\log \epsilon_t^2|^m$ at any order m, provided $\gamma(\mathbf{C}) < 0$ and $E|\log \eta_0^2|^m < \infty$. Now in view of the companion form of the matrix $\boldsymbol{A}^{(\infty)}$ (see *e.g.* Corollary 2.2 in Francq and Zakoïan, 2010a), (2.6) holds if and only if

$$\sum_{i=1}^{r} |\alpha_{i+} + \beta_i| \vee |\alpha_{i-} + \beta_i| < 1. \tag{2.7}$$

Example 2.3 (**Log-GARCH(1,1) continued).** *In the case $p = q = 1$, we have*

$$\boldsymbol{A}_t = \alpha_+ 1_{\{\eta_{t-1}>0\}} + \alpha_- 1_{\{\eta_{t-1}<0\}} + \beta \quad \text{and} \quad \boldsymbol{A}^{(m)} = E\left(|\boldsymbol{A}_1|\right)^m.$$

The conditions $E|\log \eta_0^2|^m < \infty$ and

$$\sum_{k=0}^{m} \begin{pmatrix} m \\ k \end{pmatrix} \left(a\,|\alpha_+|^k + (1-a)\,|\alpha_-|^k\right) |\beta|^{m-k} < 1$$

thus entail $E|\log \epsilon_t^2|^m < \infty$ for the log-GARCH(1,1) model.

Example 2.4 (**Symmetric case continued).** *When $\boldsymbol{\alpha}_+ = \boldsymbol{\alpha}_- = \boldsymbol{\alpha}$, the matrix $\boldsymbol{A}_t$ is no more random:*

$$\boldsymbol{A}^{(\infty)} = \boldsymbol{A}^{(1)} = \operatorname{Abs}(\boldsymbol{A}_1) = \begin{pmatrix} |\alpha_1 + \beta_1| & \cdots & |\alpha_r + \beta_r| \\ \boldsymbol{I}_{r-1} & & \boldsymbol{0}_{r-1} \end{pmatrix}.$$

In view of the companion form of this matrix we have $\rho\left(\boldsymbol{A}^{(1)}\right) < 1$ *if and only if*

$$\sum_{i=1}^{r} |\alpha_i + \beta_i| < 1.$$

The previous condition ensures $E|\log \epsilon_t^2|^m < \infty$ *for all* m *such that* $E|\log \eta_0^2|^m < \infty$.

2.3. Leverage effect

A well-known stylized fact of financial markets is that negative shocks on the returns impact future volatilities more importantly than positive shocks of the same magnitude. Generally, this so-called leverage effect is measured by computing the covariance between the innovation (or the return) at time $t-1$ and the current volatility. In our framework, it is more convenient to evaluate the leverage effect through the covariance between η_{t-1} and the current log-volatility. We restrict our study to the case $p = q = 1$, omitting subscripts to simplify notation.

Proposition 2.3 (**Leverage effect in the log-GARCH(1,1) model).** *Consider the log-GARCH(1,1) model under* (2.6)*. Assume that the innovations* η_t *are symmetrically distributed,* $E[|\log \eta_0|^2] < \infty$ *and* $|\beta| + \frac{1}{2}(|\alpha_+| + |\alpha_-|) < 1$. *Then*

$$cov(\eta_{t-1}, \log \sigma_t^2) = \frac{1}{2}(\alpha_+ - \alpha_-)\left\{E(|\eta_0|)\tau + E(|\eta_0| \log \eta_0^2)\right\}, \qquad (2.8)$$

where

$$\tau = E \log \sigma_t^2 = \frac{\omega + \frac{1}{2}(\alpha_+ - \alpha_-)E(\log \eta_0^2)}{1 - \beta - \frac{1}{2}(\alpha_+ + \alpha_-)}.$$

Thus, if the left hand side of (2.8) is negative the leverage effect is present: past negative innovations tend to increase the log-volatility, and hence the volatility, more than past positive innovations. However, the sign of the covariance is more complicated to determine than for other asymmetric models: it depends on all the GARCH coefficients, but also on the properties of the innovations distribution. Interestingly, the leverage effect may hold with $\alpha_+ > \alpha_-$. Simple calculation shows that for the EGARCH(1,1) model, $\mathrm{cov}(\eta_{t-1}, \log\sigma_t^2) = \gamma_1$.

3. Tail properties of the log-GARCH

In this section, we investigate differences between the EGARCH and the log-GARCH in terms of tail properties.

3.1. Existence of moments

We start by characterizing the existence of moments for the log-GARCH. The following result is an extension of Theorem 1 in Bauwens *et al.*, 2008, to the asymmetric case (see also Theorem 2 in He *et al.*, 2002 for the symmetric case with $p = q = 1$):

Proposition 3.1 (Existence of moments). *Assume that* $\gamma(\mathbf{C}) < 0$ *and that* $\rho\left(\boldsymbol{A}^{(\infty)}\right) < 1$. *Letting* $\lambda = \max_{1\le i\le q}\{|\alpha_{i+}| \vee |\alpha_{i-}|\}\sum_{\ell\ge 0}\|(\boldsymbol{A}^{(\infty)})^{\ell}\| < \infty$, *assume that for some* $s > 0$

$$E\Big[\exp\Big\{s\Big(\lambda\vee 1\Big)|\log\eta_0^2|\Big\}\Big] < \infty, \tag{3.1}$$

then the solution of the log-GARCH(p,q) model satisfies $E|\epsilon_0|^{2s} < \infty$.

Remark 3.1. In the case $p = q = 1$, condition (3.1) becomes explicit:

$$E\Big[\exp\Big\{s\Big(\frac{|\alpha_{1+}|\vee|\alpha_{1-}|}{1-|\alpha_{1+}+\beta_1|\vee|\alpha_{1-}+\beta_1|}\vee 1\Big)|\log\eta_0^2|\Big\}\Big] < \infty.$$

If α_{1+} and α_{1-} are non negative, Proposition 3.3 below shows that, if η_0 has regular variations, the conditions (2.7) and $E|\eta_0|^{2s(\alpha_{1+}\vee\alpha_{1-}\vee 1)} < \infty$ are sufficient for $E|\epsilon_0|^{2s} < \infty$. Note that condition $E|\eta_0|^{2s(\alpha_{1+}\vee\alpha_{1-}\vee 1)} < \infty$ is always weaker than condition (3.1).

The following result provides a sufficient condition for the Cramer's type condition (3.1).

Proposition 3.2. *If $E(|\eta_0|^s) < \infty$ for some $s > 0$ and η_0 admits a density f around 0 such that $f(y^{-1}) = o(|y|^{\delta})$ for $\delta < 1$ when $|y| \to \infty$ then $E\exp(s_1|\log\eta_0^2|) < \infty$ for some $s_1 > 0$.*

For an explicit expression of the unconditional moments in the case of symmetric log-GARCH(p,q) models, we refer the reader to Bauwens *et al.* (2008).

3.2. Regular variation of the log-GARCH(1,1)

Under the assumptions of Proposition 2.3 we have an explicit expression of the stationary solution. Thus it is possible to assert the regular variation properties of the log-GARCH model. Recall that L is a slowly varying function iff $L(xy)/L(x) \to 1$ as $x \to \infty$ for any $y > 0$. A random variable X is said to be regularly varying of index $r > 0$ if there exists a slowly varying function L and $p + q = 1$, $p \wedge q \geq 0$ such that

$$P(X > x) \sim px^{-r}L(x) \quad \text{and} \quad P(X \leq -x) \sim qx^{-r}L(x) \quad x \to \infty.$$

The following proposition asserts the regular variation properties of the stationary solution of the log-GARCH(1,1) model.

Proposition 3.3 (Regular variation of the log-GARCH(1,1) model)**.** *Consider the log-GARCH(1,1) model under (2.6) with $\alpha_{1+} \wedge \alpha_{1-} > 0$. If $(\beta_1 + \alpha_{1+}) \wedge (\beta_1 + \alpha_{1-}) < 0$, assume additionally that there exists $c > 0$ such that $P(1/\eta > t) \leq cP(\eta > t)$ for all $t \geq 1$.*

- *If η_0 is regularly varying of index $2r > 0$ then σ_0^2 and ϵ_0 are regularly varying of index $r/(\alpha_{1+} \vee \alpha_{1-})$ and $2r/(\alpha_{1+} \vee \alpha_{1-} \vee 1)$ respectively.*
- *If η_0 has finite moments of order $2r > 0$ then σ_0^2 and ϵ_0 have finite moments of order $r/(\alpha_{1+} \vee \alpha_{1-})$ and $2r/(\alpha_{1+} \vee \alpha_{1-} \vee 1)$ respectively.*

The square root of the volatility σ_0 can have heavier tails than the innovations when $\alpha_{1+} \vee \alpha_{1-} > 1$. Similarly, in the EGARCH(1,1) model the observations can have a much heavier tail than the innovations. Moreover, when the innovations are light tailed distributed (for instance exponentially distributed), the EGARCH can exhibit regular variation properties. It is not the case for the log-GARCH(1,1) model.

In this context of heavy tail, a natural way to deal with the dependence structure is to study the multivariate regular variation of a trajectory. As the innovations are independent, the dependence structure can only come from the volatility process. However, it is also independent in the extremes. The following is a straightforward application of Lemma 3.4 of Mikosch and Rezapour (2012).

Proposition 3.4 (Multivariate regular variation of the log-GARCH(1,1) model)**.** *Assume the conditions of Proposition 3.3 satisfied. Then the sequence* (σ_t^2) *is regularly varying with index* $r/(\alpha_{1+} \vee \alpha_{1-})$*. The limit measure of the vector* $\Sigma_d^2 = (\sigma_1^2, \dots, \sigma_d^2)'$ *is given by the following limiting relation on the Borel* σ*-field of* $(R \cup \{+\infty\})^d/\{0_d\}$

$$\frac{P(x^{-1}\Sigma_d^2 \in \cdot)}{P(\sigma^2 > x)} \to \frac{r}{\alpha_{1+} \vee \alpha_{1-}} \sum_{i=1}^{d} \int_1^\infty y^{-r/(\alpha_{1+} \vee \alpha_{1-})-1} 1_{\{ye_i \in \cdot\}} dy, \quad x \to \infty.$$

where e_i *is the* i*-th unit vector in* R^d *and the convergence holds vaguely.*

As for the innovations, the limiting measure above is concentrated on the axes. Thus it is also the case for the log-GARCH(1,1) process and its extremes values do not cluster. It is a drawback for modeling stock returns when clusters of volatilities are stylized facts. This lack of clustering is also observed for the EGARCH(1,1) model in Mikosch and Rezapour (2012), in contrast with the GARCH(1,1) model, see Mikosch and Starica (2000).

We have seen that the extremal behavior of the log-GARCH model is similar to that of the EGARCH(1,1) model, except that it has a much lighter tail for the same innovations. It is the contrary for the extremely small values. Let us consider the inverse (ϵ_t^{-1}) of the solution (ϵ_t) of a log-GARCH model. Then, if the variance of (η_t^{-1}) exists and $E(\eta_t^{-1}) = 0$, the process (ϵ_t^{-1}) satisfies a log-GARCH model. Under the assumptions of Proposition 3.3, the extremal behavior of ϵ_0^{-1} is driven by the innovation η_0^{-1} which, in the gaussian case, is inverse-gamma distributed with parameter $(1/2, 1/2)$. Thus ϵ_0^{-1} is regularly varying with index $r = 1/2$, implying that extremely small values of ϵ_0 are likely to occur. The time to return to the stationary regime from extremely small values is much longer than from extremely large values.

If now we consider the case of an EGARCH model, the invertibility condition obtained by Wintenberger and Cai (2011) imposes that the volatility σ_0^2 be bounded from below. Thus, its inverse is bounded from above and the extremal behavior of the return ϵ_0 is the same as that of the innovation η_0. As the innovations have lighter tails than the returns, extremely small values are not observed. This explains why the time to return to the stationary regime from extremely small values is much shorter than from extremely large values, see Figure 1 (the same reasoning also holds for GARCH(1,1) model).

4. Estimating the log-GARCH by QML

We now consider the statistical inference. Let $\epsilon_1, \ldots, \epsilon_n$ be observations of the stationary solution of (1.1), where $\boldsymbol{\theta}$ is equal to an unknown value $\boldsymbol{\theta}_0$ belonging to some parameter space $\Theta \subset \mathbb{R}^d$, with $d = 2q + p + 1$. A QMLE of $\boldsymbol{\theta}_0$ is defined as any measurable solution $\widehat{\boldsymbol{\theta}}_n$ of

$$\widehat{\boldsymbol{\theta}}_n = \arg\min_{\boldsymbol{\theta}\in\Theta} \widetilde{Q}_n(\boldsymbol{\theta}), \tag{4.1}$$

with

$$\widetilde{Q}_n(\boldsymbol{\theta}) = n^{-1} \sum_{t=r_0+1}^{n} \widetilde{\ell}_t(\boldsymbol{\theta}), \qquad \widetilde{\ell}_t(\boldsymbol{\theta}) = \frac{\epsilon_t^2}{\widetilde{\sigma}_t^2(\boldsymbol{\theta})} + \log \widetilde{\sigma}_t^2(\boldsymbol{\theta}),$$

where r_0 is a fixed integer and $\log \widetilde{\sigma}_t^2(\boldsymbol{\theta})$ is recursively defined, for $t = 1, 2, \ldots, n$, by

$$\log \widetilde{\sigma}_t^2(\boldsymbol{\theta}) = \omega + \sum_{i=1}^{q} \left(\alpha_{i+} 1_{\{\epsilon_{t-i}>0\}} + \alpha_{i-} 1_{\{\epsilon_{t-i}<0\}}\right) \log \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \log \widetilde{\sigma}_{t-j}^2(\boldsymbol{\theta}),$$

using positive initial values for $\epsilon_0^2, \ldots, \epsilon_{1-q}^2, \widetilde{\sigma}_0^2(\boldsymbol{\theta}), \ldots, , \widetilde{\sigma}_{1-p}^2(\boldsymbol{\theta})$.

Remark 4.1 (On the choice of the initial values). The initial values can be arbitrary positive numbers, for instance $\epsilon_0 = \cdots = \epsilon_{1-q} = \widetilde{\sigma}_0(\boldsymbol{\theta}) = \cdots = \widetilde{\sigma}_{1-p}(\boldsymbol{\theta}) = \sqrt{2}$ (for daily returns of stock market indices, in percentage, the empirical variance is often close to 2). They can also depend on the parameter, for instance $\epsilon_0 = \cdots = \epsilon_{1-q} = \widetilde{\sigma}_0(\boldsymbol{\theta}) = \cdots = \widetilde{\sigma}_{1-p}(\boldsymbol{\theta}) = \exp(\omega/2)$. It is also possible to take initial values depending on the observations, for instance $\epsilon_0 = \cdots = \epsilon_{1-q} = \widetilde{\sigma}_0(\boldsymbol{\theta}) = \cdots = \widetilde{\sigma}_{1-p}(\boldsymbol{\theta}) = \sqrt{n^{-1}\sum_{t=1}^{n} \epsilon_t^2}$. It will be shown in the sequel that the choice of r_0 and of the initial values is unimportant for the asymptotic behavior of the QMLE, provided r_0 is fixed and there exists a real random variable K independent of n such that

$$\sup_{\boldsymbol{\theta}\in\Theta} \left|\log \sigma_t^2(\boldsymbol{\theta}) - \log \widetilde{\sigma}_t^2(\boldsymbol{\theta})\right| < K, \quad \text{a.s. for } t = q-p+1, \ldots, q, \tag{4.2}$$

where $\sigma_t^2(\boldsymbol{\theta})$ is defined by (7.3) below. These conditions are supposed to hold in the sequel.

Remark 4.2 (Our choice of the initial values). Even if the initial values do not affect the asymptotic behavior of the QMLE, the finite sample value of $\widehat{\boldsymbol{\theta}}_n$ is however quite sensitive to these values. Based on simulation experiments and on illustrations on real data, for estimating log-GARCH(1,1) models we used the empirical variance of the first 5 values (*i.e.* a week for daily data) as proxi for the unknown value of σ_1^2. These initial values allow to compute $\widetilde{\sigma}_t^2$ for $t \geq 2$. In order to attenuate the influence of the initial value without loosing too much data, we chose $r_0 = 10$.

Remark 4.3 (The empirical treatment of null returns). Under the assumptions of Theorem 2.1, almost surely $\epsilon_t^2 \neq 0$. However, it may happen that some observations are equal to zero or are so close to zero that $\widehat{\boldsymbol{\theta}}_n$ cannot be computed (the computation of the $\log \epsilon_t^2$'s being required). To solve this potential problem, we imposed a lower bound for the $|\epsilon_t|$'s. We took the lower bound 10^{-8}, which is well inferior to a beep point, and we checked that nothing was changed in the numerical illustrations presented here when this lower bound was multiplied or divided by a factor of 100.

We now need to introduce some notation. For any $\boldsymbol{\theta} \in \Theta$, let the polynomials $\mathcal{A}_{\boldsymbol{\theta}}^{+}(z) = \sum_{i=1}^{q} \alpha_{i,+} z^i$, $\mathcal{A}_{\boldsymbol{\theta}}^{-}(z) = \sum_{i=1}^{q} \alpha_{i,-} z^i$ and $\mathcal{B}_{\boldsymbol{\theta}}(z) = 1 - \sum_{j=1}^{p} \beta_j z^j$. By convention, $\mathcal{A}_{\boldsymbol{\theta}}^{+}(z) = 0$ and $\mathcal{A}_{\boldsymbol{\theta}}^{-}(z) = 0$ if $q = 0$, and $\mathcal{B}_{\boldsymbol{\theta}}(z) = 1$ if $p = 0$. We also write $\mathbf{C}(\boldsymbol{\theta}_0)$ instead of $\mathbf{C}$ to emphasize that the unknown parameter is $\boldsymbol{\theta}_0$. The following assumptions are used to show the strong consistency of the QMLE.

A1: $\boldsymbol{\theta}_0 \in \Theta$ and Θ is compact.

A2: $\gamma\left\{\mathbf{C}(\boldsymbol{\theta}_0)\right\} < 0 \quad$ and $\quad \forall \boldsymbol{\theta} \in \Theta, \quad |\mathcal{B}_{\boldsymbol{\theta}}(z)| = 0 \Rightarrow |z| > 1.$

A3: The support of η_0 contains at least two positive values and two negative values, $E\eta_0^2 = 1$ and $E|\log \eta_0^2|^{s_0} < \infty$ for some $s_0 > 0$.

A4: If $p > 0$, $\mathcal{A}_{\boldsymbol{\theta}_0}^{+}(z)$ and $\mathcal{A}_{\boldsymbol{\theta}_0}^{-}(z)$ have no common root with $\mathcal{B}_{\boldsymbol{\theta}_0}(z)$. Moreover $\mathcal{A}_{\boldsymbol{\theta}_0}^{+}(1) + \mathcal{A}_{\boldsymbol{\theta}_0}^{-}(1) \neq 0$ and $|\alpha_{0q+}| + |\alpha_{0q+}| + |\beta_{0p}| \neq 0$.

A5: $E\left|\log \epsilon_t^2\right| < \infty$.

Assumptions **A1, A2** and **A4** are similar to those required for the consistency of the QMLE in standard GARCH models (see Berkes et al. 2003, Francq and Zakoian, 2004). Assumption **A3** precludes a mass at zero for the innovation, and, for identifiability reasons, imposes non degeneracy of the positive and negative parts of η_0. Note that, for other GARCH-type models, the absence of a lower bound for the volatility can entail inconsistency of the QMLE (see Francq and Zakoïan, 2010b). This is not the case for the log-GARCH under **A5**. Note that this assumption can be replaced by the sufficient conditions given in Proposition 2.2 (see also Examples 2.3 and 2.4).

Theorem 4.1 (Strong consistency of the QMLE). *Let* $(\widehat{\boldsymbol{\theta}}_n)$ *be a sequence* *of QMLE satisfying (4.1), where the ϵ_t's follow the asymmetric log-GARCH model of parameter $\boldsymbol{\theta}_0$. Under the assumptions (4.2) and* **A1**-**A5**, *almost surely $\widehat{\boldsymbol{\theta}}_n \to \boldsymbol{\theta}_0$ as $n \to \infty$.*

Let us now study the asymptotic normality of the QMLE. We need the classical additional assumption:

A6: $\boldsymbol{\theta}_0 \in \overset{\circ}{\Theta}$ and $\kappa_4 := E(\eta_0^4) < \infty$.

Because the volatility σ_t^2 is not bounded away from 0, we also need the following non classical assumption.

A7: There exists $s_1 > 0$ such that $E\exp(s_1|\log \eta_0^2|) < \infty$, and $\rho(\boldsymbol{A}^{(\infty)}) < 1$.

The Cramer condition on $|\log \eta_0^2|$ in **A7** is verified if η_t admits a density f around 0 that does not explode too fast (see Proposition 3.2).

Let $\nabla Q = (\nabla_1 Q, \ldots, \nabla_d Q)'$ and $\mathbb{H}Q = (\mathbb{H}_{1.}Q', \ldots, \mathbb{H}_{d.}Q')'$ be the vector and matrix of the first-order and second-order partial derivatives of a function $Q : \Theta \to \mathbb{R}$.

Theorem 4.2 (**Asymptotic normality of the QMLE).** *Under the assumptions of Theorem 4.1 and* **A6**-**A7**, *we have $\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, (\kappa_4 - 1)\mathbf{J^{-1}})$ as $n \to \infty$, where $\mathbf{J} = E[\nabla \log \sigma_t^2(\boldsymbol{\theta}_0)\nabla \log \sigma_t^2(\boldsymbol{\theta}_0)']$ is a positive definite matrix and $\xrightarrow{d}$ denotes convergence in distribution.*

It is worth noting that for the general EGARCH model, no similar results, establishing the consistency and the asymptotic normality, exist. See however Wintenberger and Cai (2011) for the EGARCH(1,1). The difficulty with the EGARCH is to invert the volatility, that is to write $\sigma_t^2(\boldsymbol{\theta})$ as a well-defined function of the past observables. In the log-GARCH model, invertibility reduces to the standard assumption on $\mathcal{B}_{\boldsymbol{\theta}}$ given in **A2**.

5. Asymmetric log-ACD model for duration data

The dynamics of duration between stock price changes has attracted much attention in the econometrics literature. Engle and Russel (1997) proposed the Autoregressive Conditional Duration (ACD) model, which assumes that the duration between price changes has the dynamics of the square of a GARCH. Bauwens and Giot (2000 and 2003) introduced logarithmic versions of the ACD, that do not constrain the sign of the coefficients (see also Bauwens, Giot, Grammig and Veredas (2004) and Allen, Chan, McAleer and Peiris (2008)). The asymmetric ACD of Bauwens and Giot (2003) applies to pairs of observation (x_i, y_i), where x_i is the duration between two changes of the bid-ask quotes posted by a market maker and y_i is a variable indicating the direction of change of the mid price defined as the average of the bid and ask prices ($y_i = 1$ if the mid price increased over duration x_i, and $y_i = -1$ otherwise). The asymmetric log-ACD proposed by Bauwens and Giot (2003) can be written as

$$\begin{cases} x_i &= \psi_i z_i, \\ \log \psi_i &= \omega + \sum_{k=1}^{q} \left(\alpha_{k+} 1_{\{y_{i-k}=1\}} + \alpha_{k-} 1_{\{y_{i-k}=-1\}} \right) \log x_{i-k} \\ & \quad + \sum_{j=1}^{p} \beta_j \log \psi_{i-j}, \end{cases} \tag{5.1}$$

where (z_i) is an iid sequence of positive variables with mean 1 (so that ψ_i can be interpreted as the conditional mean of the duration x_i). Note that $\epsilon_t := \sqrt{x_t} y_t$ follows the log-GARCH model (1.1), with $\eta_t = \sqrt{z_t} y_t$. Consequently, the results of the present paper also apply to log-ACD models. In particular, the parameters of (5.1) can be estimated by fitting model (1.1) on $\epsilon_t = \sqrt{x_t} y_t$.

6. Numerical Applications

6.1. An application to exchange rates

We consider returns series of the daily exchange rates of the American Dollar (USD), the Japanese Yen (JPY), the British Pound (BGP), the Swiss Franc

Table 1
Log-GARCH(1,1) and EGARCH(1,1) models fitted by QMLE on daily returns of exchange rates. The estimated standard deviation are displayed into brackets.

Log-GARCH					
	$\widehat{\omega}$	$\widehat{\alpha}_+$	$\widehat{\alpha}_-$	$\widehat{\beta}$	Log-Lik.
USD	0.024 (0.005)	0.027 (0.004)	0.016 (0.004)	0.971 (0.005)	**-0.104**
JPY	0.051 (0.007)	0.037 (0.006)	0.042 (0.006)	0.952 (0.006)	-0.354
GBP	0.032 (0.006)	0.030 (0.005)	0.029 (0.005)	0.964 (0.006)	**0.547**
CHF	0.057 (0.012)	0.046 (0.008)	0.036 (0.007)	0.954 (0.008)	1.477
CAD	0.021 (0.005)	0.025 (0.004)	0.017 (0.004)	0.969 (0.006)	-0.170
EGARCH					
	$\widehat{\omega}$	$\widehat{\gamma}$	$\widehat{\delta}$	$\widehat{\beta}$	Log-Lik.
USD	-0.202 (0.030)	-0.015 (0.014)	0.218 (0.031)	0.961 (0.010)	-0.116
JPY	-0.152 (0.021)	-0.061 (0.014)	0.171 (0.024)	0.970 (0.006)	**-0.334**
GBP	-0.447 (0.048)	-0.029 (0.021)	0.420 (0.041)	0.913 (0.017)	0.503
CHF	-0.246 (0.046)	-0.071 (0.022)	0.195 (0.035)	0.962 (0.009)	**1.568**
CAD	-0.091 (0.017)	-0.008 (0.010)	0.103 (0.019)	0.986 (0.005)	**-0.161**

(CHF) and Canadian Dollar (CAD) with respect to the Euro. The observations cover the period from January 5, 1999 to January 18, 2012, which corresponds to 3344 observations. The data were obtained from the web site
http://www.ecb.int/stats/exchange/eurofxref/html/index.en.html.

Table 1 displays the estimated log-GARCH(1,1) and EGARCH(1,1) models for each series. For all series, except the CHF, condition (2.7) ensuring the existence of any log-moment is satisfied. Most of the estimated models present asymmetries. However, the leverage effect is more visible in the EGARCH than in the log-GARCH models. This is particularly apparent for the JPY model for which $\widehat{\gamma}$ is clearly negative. For all models, the persistence parameter β is very high. The last column shows that for the USD and the GBP, the log-GARCH has a higher (quasi) log-likelihood than the EGARCH. The converse is true for the three other assets. A study of the residuals, not reported here, is in accordance with the better fit of one particular model for each series. This study confirms that the models do not capture exactly the same empirical properties, and are thus not perfectly substitutable.

Table 2
Log-GARCH(1,1) models fitted on 5 simulations of a log-GARCH(1,1) model.

Iter	$\widehat{\omega}$	$\widehat{\alpha}_+$	$\widehat{\alpha}_-$	$\widehat{\beta}$
1	0.025 (0.004)	0.028 (0.004)	0.018 (0.004)	0.968 (0.005)
2	0.021 (0.003)	0.023 (0.003)	0.013 (0.003)	0.976 (0.004)
3	0.026 (0.003)	0.028 (0.004)	0.017 (0.003)	0.969 (0.004)
4	0.022 (0.003)	0.024 (0.004)	0.018 (0.003)	0.972 (0.004)
5	0.024 (0.003)	0.028 (0.004)	0.014 (0.003)	0.974 (0.003)

Table 3
EGARCH(1,1) models fitted on 5 simulations of a log-GARCH(1,1) model.

Iter	$\widehat{\omega}$	$\widehat{\gamma}$	$\widehat{\delta}$	$\widehat{\beta}$
1	-0.095 (0.016)	-0.014 (0.009)	0.104 (0.017)	0.976 (0.006)
2	-0.127 (0.018)	0.009 (0.010)	0.148 (0.021)	0.976 (0.007)
3	-0.147 (0.018)	0.001 (0.010)	0.177 (0.022)	0.971 (0.007)
4	-0.136 (0.019)	-0.012 (0.010)	0.155 (0.022)	0.976 (0.007)
5	-0.146 (0.019)	-0.009 (0.010)	0.177 (0.023)	0.971 (0.007)

6.2. A Monte Carlo experiment

To evaluate the finite sample performance of the QML for the two models we made the following numerical experiments. We first simulated the log-GARCH(1,1) model, with $n = 3344$, $\eta_t \sim \mathcal{N}(0,1)$, and a parameter close to those of Table 1, that is $\boldsymbol{\theta}_0 = (0.024, 0.027, 0.016, 0.971)$. Tables 2 and 3 display the log-GARCH(1,1) and EGARCH(1,1) models fitted on these simulations. The first table shows that the log-GARCH(1,1) is accurately estimated. In a second time, we repeated the same experiments for simulations of an EGARCH(1,1) model of parameter $(\omega_0, \gamma_0, \delta_0, \beta_0) = (-0.204, -0.012, 0.227, 0.963)$. Tables 4 and 5 are the analogs of Tables 2 and 3 for the simulations of this EGARCH model instead of the log-GARCH. Tables 5 indicates that the EGARCH are satisfactorily estimated.

Table 4
Log-GARCH(1,1) models fitted on 5 simulations of an EGARCH(1,1) model.

Iter	$\widehat{\omega}$	$\widehat{\alpha}_+$	$\widehat{\alpha}_-$	$\widehat{\beta}$
1	0.039 (0.008)	0.071 (0.008)	0.052 (0.007)	0.874 (0.015)
2	0.055 (0.006)	0.058 (0.007)	0.052 (0.006)	0.913 (0.010)
3	0.052 (0.008)	0.070 (0.008)	0.060 (0.007)	0.873 (0.015)
4	0.051 (0.008)	0.076 (0.008)	0.056 (0.007)	0.878 (0.014)
5	0.056 (0.007)	0.061 (0.007)	0.060 (0.007)	0.896 (0.012)

Table 5
EGARCH(1,1) models fitted on 5 simulations of an EGARCH(1,1) model.

Iter	$\widehat{\omega}$	$\widehat{\gamma}$	$\widehat{\delta}$	$\widehat{\beta}$
1	-0.220 (0.022)	-0.024 (0.013)	0.235 (0.023)	0.950 (0.010)
2	-0.196 (0.020)	-0.029 (0.012)	0.219 (0.022)	0.961 (0.008)
3	-0.222 (0.022)	-0.005 (0.013)	0.241 (0.024)	0.947 (0.010)
4	-0.227 (0.022)	-0.025 (0.012)	0.248 (0.023)	0.950 (0.010)
5	-0.209 (0.021)	-0.003 (0.012)	0.234 (0.023)	0.955 (0.009)

7. Proofs

7.1. Proof of Theorem 2.1

Since the random variable $\|\boldsymbol{C}_0\|$ is bounded, we have $E\log^+\|\boldsymbol{C}_0\| < \infty$. The moment condition on η_t entails that we also have $E\log^+\|\boldsymbol{b}_0\| < \infty$. When $\gamma(\mathbf{C}) < 0$, Cauchy's root test shows that, almost surely (a.s.), the series

$$\boldsymbol{z}_t = \boldsymbol{b}_t + \sum_{n=0}^{\infty} \boldsymbol{C}_t \boldsymbol{C}_{t-1} \cdots \boldsymbol{C}_{t-n} \boldsymbol{b}_{t-n-1} \tag{7.1}$$

converges absolutely for all t and satisfies (2.2). A strictly stationary solution to model (1.1) is then obtained as $\epsilon_t = \exp\left\{\frac{1}{2}\boldsymbol{z}_{2q+1,t}\right\}\eta_t$, where $\boldsymbol{z}_{i,t}$ denotes the i-th element of $\boldsymbol{z}_t$. This solution is non anticipative and ergodic, as a measurable function of $\{\eta_u, u \leq t\}$.

We now prove that (7.1) is the unique nonanticipative solution of (2.2) when $\gamma(\mathbf{C}) < 0$. Let $(\boldsymbol{z}_t^*)$ be a strictly stationary process satisfying $\boldsymbol{z}_t^* = \boldsymbol{C}_t \boldsymbol{z}_{t-1}^* + \boldsymbol{b}_t$.

For all $N \geq 0$,

$$\boldsymbol{z}_t^* = \boldsymbol{z}_t(N) + \boldsymbol{C}_t \ldots \boldsymbol{C}_{t-N}\boldsymbol{z}_{t-N-1}^*, \quad \boldsymbol{z}_t(N) = \boldsymbol{b}_t + \sum_{n=0}^{N} \boldsymbol{C}_t\boldsymbol{C}_{t-1}\cdots\boldsymbol{C}_{t-n}\boldsymbol{b}_{t-n-1}.$$

We then have

$$\|\boldsymbol{z}_t - \boldsymbol{z}_t^*\| \leq \left\| \sum_{n=N+1}^{\infty} \boldsymbol{C}_t\boldsymbol{C}_{t-1}\cdots\boldsymbol{C}_{t-n}\boldsymbol{b}_{t-n-1} \right\| + \|\boldsymbol{C}_t \ldots \boldsymbol{C}_{t-N}\|\|\boldsymbol{z}_{t-N-1}^*\|.$$

The first term in the right-hand side tends to 0 a.s. when $N \to \infty$. The second term tends to 0 in probability because $\gamma(\mathbf{C}) < 0$ entails that $\|\boldsymbol{C}_t \ldots \boldsymbol{C}_{t-N}\| \to 0$ a.s. and the distribution of $\|\boldsymbol{z}_{t-N-1}^*\|$ is independent of N by stationarity. We have shown that $\boldsymbol{z}_t - \boldsymbol{z}_t^* \to 0$ in probability when $N \to \infty$. This quantity being independent of N we have $\boldsymbol{z}_t = \boldsymbol{z}_t^*$ a.s. for any t. □

7.2. *Proof of Proposition 2.1*

Let X be a random variable such that $X > 0$ a.s. and $EX^r < \infty$ for some $r > 0$. If $E \log X < 0$, then there exists $s > 0$ such that $EX^s < 1$ (see *e.g.* Lemma 2.3 in Berkes, Horváth and Kokoszka, 2003). Noting that $E\|\boldsymbol{C}_t \cdots \boldsymbol{C}_1\| \leq (E\|\boldsymbol{C}_1\|)^t < \infty$ for all t, the previous result shows that when $\gamma(\mathbf{C}) < 0$ we have $E\|\boldsymbol{C}_{k_0} \cdots \boldsymbol{C}_1\|^s < 1$ for some $s > 0$ and some $k_0 \geq 1$. One can always assume that $s < 1$. In view of (7.1), the c_r-inequality and standard arguments (see *e.g.* Corollary 2.3 in Francq and Zakoïan, 2010a) entail that $E\|\boldsymbol{z}_t\|^s < \infty$, provided $E\|\boldsymbol{b}_t\|^s < \infty$, which holds true when $s \leq s_0$. The conclusion follows. □

7.3. *Proof of Proposition 2.2*

By (2.4), componentwise we have

$$\mathrm{Abs}(\boldsymbol{\sigma}_{t,r}) \leq \mathrm{Abs}(\boldsymbol{u}_t) + \sum_{\ell=0}^{\infty} \boldsymbol{A}_{t,\ell}\mathrm{Abs}(\boldsymbol{u}_{t-\ell-1}), \quad \boldsymbol{A}_{t,\ell} := \prod_{j=0}^{\ell} \mathrm{Abs}(\boldsymbol{A}_{t-j}), \quad (7.2)$$

where each element of the series is defined a priori in $[0, \infty]$. In view of the form (2.3) of the matrices $\boldsymbol{A}_t$, each element of

$$\boldsymbol{A}_{t,\ell}\mathrm{Abs}(\boldsymbol{u}_{t-\ell-1}) = |u_{t-\ell-1}| \prod_{j=0}^{\ell} \mathrm{Abs}(\boldsymbol{A}_{t-j})\boldsymbol{e}_1$$

is a sum of products of the form $|u_{t-\ell-1}| \prod_{j=0}^{k} |\mu_{\ell_j}(\eta_{t-i_j})|$ with $0 \le k \le \ell$ and $0 \le i_0 < \cdots < i_k \le \ell + 1$. To give more detail, consider for instance the case $r = 3$. We then have

$$\boldsymbol{A}_{t,1}\mathrm{Abs}(\boldsymbol{u}_{t-2}) = \begin{pmatrix} |\mu_1(\eta_{t-1})||\mu_1(\eta_{t-2})||u_{t-2}| + |\mu_2(\eta_{t-2})||u_{t-2}| \\ |\mu_1(\eta_{t-2})||u_{t-2}| \\ |u_{t-2}| \end{pmatrix}.$$

Noting that $|u_{t-\ell-1}|$ is a function of $\eta_{t-\ell-2}$ and its past values, we obtain $E\boldsymbol{A}_{t,1}\mathrm{Abs}(\boldsymbol{u}_{t-2}) = E\mathrm{Abs}(\boldsymbol{A}_t)E\mathrm{Abs}(\boldsymbol{A}_{t-1})E\mathrm{Abs}(\boldsymbol{u}_{t-2})$. More generally, it can be shown by induction on ℓ that the i-th element of the vector $\boldsymbol{A}_{t-1,\ell-1}\mathrm{Abs}(\boldsymbol{u}_{t-\ell-1})$ is independent of the i-th element of the first row of $\mathrm{Abs}(\boldsymbol{A}_t)$. It follows that $E\boldsymbol{A}_{t,\ell}\mathrm{Abs}(\boldsymbol{u}_{t-\ell-1}) = E\mathrm{Abs}(\boldsymbol{A}_t)E\boldsymbol{A}_{t-1,\ell-1}\mathrm{Abs}(\boldsymbol{u}_{t-\ell-1})$. The property extends to $r \neq 3$. Therefore, although the matrices involved in the product $\boldsymbol{A}_{t,\ell}\mathrm{Abs}(\boldsymbol{u}_{t-\ell-1})$ are not independent (in the case $r > 1$), we have

$$\begin{aligned} E\boldsymbol{A}_{t,\ell}\mathrm{Abs}(\boldsymbol{u}_{t-\ell-1}) &= \prod_{j=0}^{\ell} E\mathrm{Abs}(\boldsymbol{A}_{t-j})E\mathrm{Abs}(\boldsymbol{u}_{t-\ell-1}) \\ &= \left(\boldsymbol{A}^{(1)}\right)^{\ell+1} E\mathrm{Abs}(\boldsymbol{u}_1). \end{aligned}$$

In view of (7.2), the condition $\rho(\boldsymbol{A}^{(1)}) < 1$ then entails that $E\mathrm{Abs}(\boldsymbol{\sigma}_{t,r})$ is finite.

The case $r = 1$ is treated by noting that $\boldsymbol{A}_{t,\ell}\mathrm{Abs}(\boldsymbol{u}_{t-\ell-1})$ is a product of independent random variables.

To deal with the cases $r \neq 1$ and $m \neq 1$, we work with (2.2) instead of (2.4). This Markovian representation has an higher dimension but involves independent coefficients $\boldsymbol{C}_t$. Define $\boldsymbol{C}_{t,\ell}$ by replacing $\boldsymbol{A}_{t-j}$ by $\boldsymbol{C}_{t-j}$ in $\boldsymbol{A}_{t,\ell}$. We then have

$$EC_{t,\ell}^{\otimes m}\mathrm{Abs}(\boldsymbol{b}_{t-\ell-1})^{\otimes m} = \left(\boldsymbol{C}^{(m)}\right)^{\ell+1} E\mathrm{Abs}(\boldsymbol{b}_1)^{\otimes m}.$$

For all $m \geq 1$, let $\|\boldsymbol{M}\|_m = (E\|\boldsymbol{M}\|^m)^{1/m}$ where $\|\boldsymbol{M}\|$ is the sum of the absolute values of the elements of the matrix $\boldsymbol{M}$. Using the elementary relations $\|\boldsymbol{M}\|\|\boldsymbol{N}\| = \|\boldsymbol{M}\otimes\boldsymbol{N}\|$ and $E\|\mathrm{Abs}(\boldsymbol{M})\| = \|E\mathrm{Abs}(\boldsymbol{M})\|$ for any matrices $\boldsymbol{M}$ and $\boldsymbol{N}$, the condition $\rho(\boldsymbol{C}^{(m)}) < 1$ entails $E\left\|\boldsymbol{C}_{t,\ell}\mathrm{Abs}(\boldsymbol{b}_{t-\ell-1})\right\|^m = \|E\boldsymbol{C}_{t,\ell}^{\otimes m}\mathrm{Abs}(\boldsymbol{b}_{t-\ell-1})^{\otimes m}\| \to 0$ at the exponential rate as $\ell\to\infty$, and thus

$$\|\mathrm{Abs}(\boldsymbol{z}_t)\|_m \leq \|\mathrm{Abs}(\boldsymbol{b}_t)\|_m + \sum_{\ell=0}^{\infty}\|\boldsymbol{C}_{t,\ell}\mathrm{Abs}(\boldsymbol{b}_{t-\ell-1})\|_m < \infty,$$

which allows to conclude. □

7.4. Proof of Proposition 2.3

By the concavity of the logarithm function, the condition $|\alpha_+ + \beta||\alpha_- + \beta| < 1$ is satisfied. By Example 2.1 and the symmetry of the distribution of η_0, the existence of a strictly stationary solution process (ϵ_t) is thus guaranteed. By 2.3, this solution satisfies $E|\log\epsilon_t^2| < \infty$. Let

$$a_t = (\alpha_+ 1_{\{\eta_t>0\}} + \alpha_- 1_{\{\eta_t<0\}})\eta_t, \qquad b_t = (\alpha_+ 1_{\{\eta_t>0\}} + \alpha_- 1_{\{\eta_t<0\}})\eta_t \log\eta_t^2.$$

We have

$$Ea_t = (\alpha_+ - \alpha_-)E(\eta_0 1_{\{\eta_0>0\}}), \qquad Eb_t = (\alpha_+ - \alpha_-)E(\eta_0\log\eta_0^2 1_{\{\eta_0>0\}}),$$

using the symmetry assumption for the second equality. Thus

$$\mathrm{cov}(\eta_{t-1}, \log(\sigma_t^2)) \quad = \quad E[a_{t-1}\log(\sigma_{t-1}^2) + b_{t-1}],$$

and the conclusion follows. □

7.5. Proof of Proposition 3.1

By definition, $|\log(\sigma_t^2)| \leq \|\boldsymbol{\sigma}_{t,r}\| = \|\mathrm{Abs}(\boldsymbol{\sigma}_{t,r})\|$. Then, we have

$$\begin{aligned} E|\sigma_t^2|^s &\leq E\left\{\exp(s\|\mathrm{Abs}(\boldsymbol{\sigma}_{t,r})\|)\right\} = \sum_{k=0}^{\infty} \frac{s^k \|\mathrm{Abs}(\boldsymbol{\sigma}_{t,r})\|_k^k}{k!} \\ &\leq \sum_{k=0}^{\infty} \frac{s^k \|\mathrm{Abs}(\boldsymbol{u}_0)\|_k^k \left\{\sum_{\ell=0}^{\infty} \|(\boldsymbol{A}^{(\infty)})^\ell\|\right\}^k}{k!} \\ &= E\exp\left\{s\|\mathrm{Abs}(\boldsymbol{u}_0)\| \sum_{\ell=0}^{\infty} \|(\boldsymbol{A}^{(\infty)})^\ell\|\right\}, \end{aligned}$$

where the last inequality comes from (2.5). By definition $\boldsymbol{u}_0 = (u_0, 0'_{r-1})'$ with

$$u_0 = \omega + \sum_{i=1}^{q} (\alpha_{i+} 1_{\eta_{-i}>0} + \alpha_{i-} 1_{\eta_{-i}<0}) \log \eta_{-i}^2.$$

Thus $\|\mathrm{Abs}(\boldsymbol{u}_0)\| \leq |u_0| \leq |\omega| + \max_{1\leq i\leq q} |\alpha_{i+}| \vee |\alpha_{i-}| \sum_{j=1}^{q} |\log \eta_{-j}^2|$ and it follows that

$$E|\sigma_t^2|^s \leq \exp\left\{s|\omega| \sum_{\ell=0}^{\infty} \|(\boldsymbol{A}^{(\infty)})^\ell\|\right\} \left\{E\exp\left(s\lambda|\log\eta_0^2|\right)\right\}^q < \infty$$

under (3.1). □

7.6. Proof of Proposition 3.2

Without loss of generality assume that f exists on $[-1, 1]$. Then there exists $M > 0$ such that $f(1/y) \leq M|y|^\delta$ for all $y \geq 1$ and we obtain

$$\begin{aligned} E\exp(s_1|\log\eta_0^2|) &\leq \int_{|x|<1} \exp(2s_1\log(1/x)) f(x)dx + \int \exp(s_1 \log(x^2)) dP_\eta(x) \\ &\leq 2M \int_1^\infty y^{2(s_1-1)+\delta} dy + E(|\eta_0|^{2s_1}). \end{aligned}$$

The upper bound is finite for sufficiently small s_1 and the result is proved. □

7.7. Proof of Proposition 3.3

To prove the first assertion, note that if η_0 is regularly varying of index $2r$ then η_0^2 is regularly varying of index r. Thus $u_1 = \omega + (\alpha_{1+}1_{\{\eta_0>0\}} + \alpha_{1-}1_{\{\eta_0<0\}})\log(\eta_0^2)$ is such that

$$\begin{aligned} P(e^{u_0} > x) &= P(\eta_0 > 0)P\left(\eta_0^{2\alpha_{1+}} > xe^{-\omega} \mid \eta_0 > 0\right) \\ &\quad + P(\eta_0 < 0)P\left(\eta_0^{2\alpha_{1-}} > xe^{-\omega} \mid \eta_0 < 0\right). \end{aligned}$$

Then e^{u_0} is also regularly varying with index $r/(\alpha_{1+}) \wedge r/(\alpha_{1-})$. An application Lemma 3.3 in Mikosch and Rezapour (2012) yields the first assertion. The second assertion follows easily by independence of η_0 and σ_0 with respective regularly variation indexes r and $r/(\alpha_{1+} \vee \alpha_{1-})$. □

7.8. Proof of Theorem 4.1

We will use the following standard result (see *e.g.* Exercise 2.11 in Francq and Zakoian, 2010a).

Lemma 7.1. *Let (X_n) be a sequence of random variables. If $\sup_n E|X_n| < \infty$, then almost surely $n^{-1}X_n \to 0$ as $n \to \infty$. The almost sure convergence may fail when $\sup_n E|X_n| = \infty$. If the sequence (X_n) is bounded in probability, then $n^{-1}X_n \to 0$ in probability.*

Turning to the proof of Theorem 4.1, first note that **A2**, **A3** and Proposition 2.1 ensure the a.s. absolute convergence of the series

$$\log \sigma_t^2(\boldsymbol{\theta}) := \mathcal{B}_{\boldsymbol{\theta}}^{-1}(B)\left\{\omega + \sum_{i=1}^{q}\left(\alpha_{i+}1_{\{\epsilon_{t-i}>0\}} + \alpha_{i-}1_{\{\epsilon_{t-i}<0\}}\right)\log \epsilon_{t-i}^2\right\}. \quad (7.3)$$

Let

$$Q_n(\boldsymbol{\theta}) = n^{-1}\sum_{t=r_0+1}^{n}\ell_t(\boldsymbol{\theta}), \qquad \ell_t(\boldsymbol{\theta}) = \frac{\epsilon_t^2}{\sigma_t^2(\boldsymbol{\theta})} + \log \sigma_t^2(\boldsymbol{\theta}).$$

Using standard arguments, as in the proof of Theorem 2.1 in Francq and Zakoian (2004) (hereafter FZ), the consistency is obtained by showing the following intermediate results

$$i)\ \lim_{n\to\infty}\sup_{\boldsymbol{\theta}\in\Theta}|Q_n(\boldsymbol{\theta})-\widetilde{Q}_n(\boldsymbol{\theta})|=0 \text{ a.s.};$$

$$ii)\ \text{if } \sigma_1^2(\boldsymbol{\theta})=\sigma_1^2(\boldsymbol{\theta}_0) \text{ a.s. then } \boldsymbol{\theta}=\boldsymbol{\theta}_0;$$

$$iii)\ \text{if } \boldsymbol{\theta}\neq\boldsymbol{\theta}_0\ ,\quad E\ell_t(\boldsymbol{\theta})>E\ell_t(\boldsymbol{\theta}_0);$$

$iv)$ any $\boldsymbol{\theta}\neq\boldsymbol{\theta}_0$ has a neighborhood $V(\boldsymbol{\theta})$ such that

$$\liminf_{n\to\infty}\inf_{\boldsymbol{\theta}^*\in V(\boldsymbol{\theta})}\widetilde{Q}_n(\boldsymbol{\theta}^*)>E\ell_t(\boldsymbol{\theta}_0)\quad\text{a.s.}$$

Because of the multiplicative form of the volatility, the step *i)* is more delicate than in the standard GARCH case. In the case $p=q=1$, we have

$$\log\sigma_t^2(\boldsymbol{\theta})-\log\widetilde{\sigma}_t^2(\boldsymbol{\theta})=\beta^{t-1}\left\{\log\sigma_1^2(\boldsymbol{\theta})-\log\widetilde{\sigma}_1^2(\boldsymbol{\theta})\right\},\quad\forall t\geq 1.$$

In the general case, as in FZ, using (4.2) one can show that for almost all trajectories,

$$\sup_{\boldsymbol{\theta}\in\Theta}\left|\log\sigma_t^2(\boldsymbol{\theta})-\log\widetilde{\sigma}_t^2(\boldsymbol{\theta})\right|\leq K\rho^t,\tag{7.4}$$

where $\rho\in(0,1)$ and $K>0$. When the initial values are chosen as suggested by Remark 4.2, K is the realization of a random variable which is measurable with respect to $\sigma(\{\eta_u,u\leq 5\})$. First complete the proof of $i)$ in the case $p=q=1$ and $\alpha_+=\alpha_-$, for which the notation is more explicit. In view of the multiplicative form of the volatility

$$\sigma_t^2(\boldsymbol{\theta})=e^{\beta^{t-1}\log\sigma_1^2(\boldsymbol{\theta})}\prod_{i=0}^{t-2}e^{\beta^i\left\{\omega+\alpha\log\epsilon_{t-1-i}^2\right\}},\tag{7.5}$$

we have

$$\begin{aligned}\frac{1}{t}\log\left|\frac{1}{\sigma_t^2(\boldsymbol{\theta})}-\frac{1}{\widetilde{\sigma}_t^2(\boldsymbol{\theta})}\right| &= \frac{-1}{t}\sum_{i=0}^{t-2}\beta^i\left\{\omega+\alpha\log\epsilon_{t-1-i}^2\right\}\\ &\quad+\frac{1}{t}\log\left|e^{-\beta^{t-1}\log\sigma_1^2(\boldsymbol{\theta})}-e^{-\beta^{t-1}\log\widetilde{\sigma}_1^2(\boldsymbol{\theta})}\right|.\end{aligned}$$

Applying Lemma 7.1, the first term of the right-hand side of the equality tends almost surely to zero because it is bounded by a variable of the form $|X_t|/t$, with $E|X_t| < \infty$, under **A5**. The second term is equal to

$$\frac{1}{t} \log \left| \left\{ \log \sigma_1^2(\boldsymbol{\theta}) - \log \widetilde{\sigma}_1^2(\boldsymbol{\theta}) \right\} \beta^{t-1} e^{-\beta^{t-1} x^*} \right|,$$

where x^* is between $\log \sigma_1^2(\boldsymbol{\theta})$ and $\log \widetilde{\sigma}_1^2(\boldsymbol{\theta})$. This second term thus tends to $\log|\beta| < 0$ when $t \to \infty$. It follows that

$$\sup_{\boldsymbol{\theta}\in\Theta} \left| \frac{1}{\sigma_t^2(\boldsymbol{\theta})} - \frac{1}{\widetilde{\sigma}_t^2(\boldsymbol{\theta})} \right| \leq K\rho^t, \tag{7.6}$$

where K and ρ are as in (7.4). Now consider the general case. Iterating (1.1), using the compactness of Θ and the second part of **A2**, we have

$$\begin{aligned} \log \sigma_t^2(\boldsymbol{\theta}) &= \sum_{i=1}^{t-1} c_i(\boldsymbol{\theta}) + c_{i+}(\boldsymbol{\theta}) 1_{\{\epsilon_{t-i}>0\}} \log \epsilon_{t-i}^2 + c_{i-}(\boldsymbol{\theta}) 1_{\{\epsilon_{t-i}<0\}} \log \epsilon_{t-i}^2 \\ &\quad + \sum_{j=1}^{p} c_{t,j}(\boldsymbol{\theta}) \log \sigma_{q+1-j}^2(\boldsymbol{\theta}) \end{aligned}$$

with

$$\sup_{\boldsymbol{\theta}\in\Theta} \max\{|c_i(\boldsymbol{\theta})|, |c_{i+}(\boldsymbol{\theta})|, |c_{i-}(\boldsymbol{\theta})|, |c_{i,1}(\boldsymbol{\theta})|, \dots, |c_{i,p}(\boldsymbol{\theta})|\} \leq K\rho^i, \quad \rho \in (0,1). \tag{7.7}$$

We then obtain a multiplicative form for $\sigma_t^2(\boldsymbol{\theta})$ which generalizes (7.5), and deduce that

$$\frac{1}{t} \log \left| \frac{1}{\sigma_t^2(\boldsymbol{\theta})} - \frac{1}{\widetilde{\sigma}_t^2(\boldsymbol{\theta})} \right| = a_1 + a_2,$$

where

$$a_1 = \frac{-1}{t} \sum_{i=1}^{t-1} c_i(\boldsymbol{\theta}) + c_{i+}(\boldsymbol{\theta}) 1_{\{\epsilon_{t-i}>0\}} \log \epsilon_{t-i}^2 + c_{i-}(\boldsymbol{\theta}) 1_{\{\epsilon_{t-i}<0\}} \log \epsilon_{t-i}^2 \to 0 \quad \text{a.s.}$$

in view of (7.7) and Lemma 7.1, and for x_j^*'s between $\log \sigma_{q+1-j}^2(\boldsymbol{\theta})$ and $\log \widetilde{\sigma}_{q+1-j}^2(\boldsymbol{\theta})$,

$$\begin{aligned} a_2 &= \frac{1}{t} \log \left| \exp\left\{ -\sum_{j=1}^{p} c_{t,j}(\boldsymbol{\theta}) \log \sigma_{q+1-j}^2(\boldsymbol{\theta}) \right\} - \exp\left\{ -\sum_{j=1}^{p} c_{t,j}(\boldsymbol{\theta}) \log \widetilde{\sigma}_{q+1-j}^2(\boldsymbol{\theta}) \right\} \right| \\ &= \frac{1}{t} \log \left| -\sum_{j=1}^{p} c_{t,j}(\boldsymbol{\theta}) \left\{ \log \sigma_{q+1-j}^2(\boldsymbol{\theta}) - \log \sigma_{q+1-j}^2(\boldsymbol{\theta}) \right\} \exp\left\{ -\sum_{k=1}^{p} c_{t,k}(\boldsymbol{\theta}) \log x_k^* \right\} \right| \\ &= \frac{1}{t} \log \left| -\sum_{j=1}^{p} c_{t,j}(\boldsymbol{\theta}) \right| + o(1) \quad \text{a.s.} \end{aligned}$$

using (4.2) and (7.7). Using again (4.2), it follows that $\limsup_{n\to\infty} a_2 \le \log\rho < 0$. We conclude that (7.6) holds true in the general case. The proof of *i)* then follows from (7.4)-(7.6), as in FZ.

To show *ii)*, note that we have

$$\mathcal{B}_{\boldsymbol{\theta}}(B)\log\sigma_t^2(\boldsymbol{\theta}) = \omega + \mathcal{A}^+_{\boldsymbol{\theta}}(B)1_{\{\epsilon_t>0\}}\log\epsilon_t^2 + \mathcal{A}^-_{\boldsymbol{\theta}}(B)1_{\{\epsilon_t<0\}}\log\epsilon_t^2. \tag{7.8}$$

If $\log\sigma_1^2(\boldsymbol{\theta}) = \log\sigma_1^2(\boldsymbol{\theta}_0)$ a.s., by stationarity we have $\log\sigma_t^2(\boldsymbol{\theta}) = \log\sigma_t^2(\boldsymbol{\theta}_0)$ for all t, and thus we have almost surely

$$\begin{aligned}&\left\{\frac{\mathcal{A}^+_{\boldsymbol{\theta}}(B)}{\mathcal{B}_{\boldsymbol{\theta}}(B)} - \frac{\mathcal{A}^+_{\boldsymbol{\theta}_0}(B)}{\mathcal{B}_{\boldsymbol{\theta}_0}(B)}\right\}1_{\{\epsilon_t>0\}}\log\epsilon_t^2 + \left\{\frac{\mathcal{A}^-_{\boldsymbol{\theta}}(B)}{\mathcal{B}_{\boldsymbol{\theta}}(B)} - \frac{\mathcal{A}^-_{\boldsymbol{\theta}_0}(B)}{\mathcal{B}_{\boldsymbol{\theta}_0}(B)}\right\}1_{\{\epsilon_t<0\}}\log\epsilon_t^2\\ = &\quad \frac{\omega_0}{\mathcal{B}_{\boldsymbol{\theta}_0}(1)} - \frac{\omega}{\mathcal{B}_{\boldsymbol{\theta}}(1)}.\end{aligned}$$

Denote by R_t any random variable which is measurable with respect to $\sigma\left(\{\eta_u, u\le t\}\right)$. If

$$\frac{\mathcal{A}^+_{\boldsymbol{\theta}}(B)}{\mathcal{B}_{\boldsymbol{\theta}}(B)} \neq \frac{\mathcal{A}^+_{\boldsymbol{\theta}_0}(B)}{\mathcal{B}_{\boldsymbol{\theta}_0}(B)} \quad \text{or} \quad \frac{\mathcal{A}^-_{\boldsymbol{\theta}}(B)}{\mathcal{B}_{\boldsymbol{\theta}}(B)} \neq \frac{\mathcal{A}^-_{\boldsymbol{\theta}_0}(B)}{\mathcal{B}_{\boldsymbol{\theta}_0}(B)}, \tag{7.9}$$

there exists a non null $(c_+, c_-)' \in \mathbb{R}^2$, such that

$$c_+1_{\{\eta_t>0\}}\log\epsilon_t^2 + c_-1_{\{\eta_t<0\}}\log\epsilon_t^2 + R_{t-1} = 0 \text{ a.s.}$$

This is equivalent to the two equations

$$\left(c_+\log\eta_t^2 + c_+\log\sigma_t^2 + R_{t-1}\right)1_{\{\eta_t>0\}} = 0$$

and

$$\left(c_-\log\eta_t^2 + c_-\log\sigma_t^2 + R_{t-1}\right)1_{\{\eta_t<0\}} = 0.$$

Note that if an equation of the form $a\log x^2 1_{\{x>0\}} + b1_{\{x>0\}} = 0$ admits two positive solutions then $a = 0$. This result, **A3**, and the independence between η_t and (σ_t^2, R_{t-1}) imply that $c_+ = 0$. Similarly we obtain $c_- = 0$, which leads to a contradiction. We conclude that (7.9) cannot hold true, and the conclusion follows from **A4**.

Since $\sigma_t^2(\boldsymbol{\theta})$ is not bounded away from zero, the beginning of the proof of *iii)* slightly differs from that given by FZ in the standard GARCH case. In view of (7.8), the second part of **A2** and **A5** entail that $E|\log \sigma_t^2(\boldsymbol{\theta})| < \infty$ for all $\boldsymbol{\theta} \in \Theta$. It follows that $E\ell_t^-(\boldsymbol{\theta}) < \infty$ and $E|\ell_t(\boldsymbol{\theta}_0)| < \infty$.

The rest of the proof of *iii)*, as well as that of *iv)*, are identical to those given in FZ. □

7.9. Proof of Theorem 4.2

A Taylor expansion gives

$$\nabla_i Q_n(\widehat{\boldsymbol{\theta}}_n) - \nabla_i Q_n(\boldsymbol{\theta}_0) = \mathbb{H}_{i.} Q_n(\widetilde{\boldsymbol{\theta}}_{n,i})(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \quad \text{for all } 1 \leq i \leq d,$$

where the $\widetilde{\boldsymbol{\theta}}_{n,i}$'s are such that $\|\widetilde{\boldsymbol{\theta}}_{n,i} - \boldsymbol{\theta}_0\| \leq \|\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\|$. As in Section 5 of Bardet and Wintenberger (2009), the asymptotic normality is obtained by showing:

1. $n^{1/2}\nabla Q_n(\boldsymbol{\theta}_0) \to \mathcal{N}(\mathbf{0}, (\kappa_4 - 1)\mathbf{J})$,
2. $\|\mathbb{H} Q_n(\widetilde{\boldsymbol{\theta}}_n) - \mathbf{J}\|$ converges a.s. to 0 for any sequence $(\widetilde{\boldsymbol{\theta}}_n)$ converging a.s. to $\boldsymbol{\theta}_0$ and $\mathbf{J}$ is invertible,
3. $n^{1/2}\|\nabla \widetilde{Q}_n(\widehat{\boldsymbol{\theta}}_n) - \nabla Q_n(\widehat{\boldsymbol{\theta}}_n)\|$ converges a.s. to 0.

In order to prove the points 1-3 we will use the following Lemma

Lemma 7.2. *Under the assumptions of Theorem 4.1 and* **A7**, *for any $m > 0$ there exists a neighborhood $\mathcal{V}$ of $\boldsymbol{\theta}_0$ such that $E[\sup_{\mathcal{V}}(\sigma_t^2/\sigma_t^2(\boldsymbol{\theta}))^m] < \infty$ and $E[\sup_{\mathcal{V}} |\log \sigma_t^2(\boldsymbol{\theta})|^m] < \infty$.*

Proof. We have

$$\begin{aligned}\log \sigma_t^2(\boldsymbol{\theta}_0) - \log \sigma_t^2(\boldsymbol{\theta}) \quad &= \quad \omega_0 - \omega + \sum_{j=1}^{p} \beta_j \{\log \sigma_{t-j}^2(\boldsymbol{\theta}_0) - \log \sigma_{t-j}^2(\boldsymbol{\theta})\} \\ &+ \mathbf{V}_{\boldsymbol{\theta}_0 - \boldsymbol{\theta}} \boldsymbol{\sigma}_{t-1,r} + \mathcal{A}^+_{\boldsymbol{\theta}_0 - \boldsymbol{\theta}}(B) 1_{\eta_t > 0} \log \eta_t^2 + \mathcal{A}^-_{\boldsymbol{\theta}_0 - \boldsymbol{\theta}}(B) 1_{\eta_t < 0} \log \eta_t^2\end{aligned}$$

with $\boldsymbol{\sigma}_{t,r} = (\log \sigma_t^2(\boldsymbol{\theta}_0), \ldots, \log \sigma_{t-r+1}^2(\boldsymbol{\theta}_0))'$,

$\mathbf{V}_{\boldsymbol{\theta}} = (\alpha_{1+} 1_{\{\eta_{t-1} > 0\}} + \alpha_{1-} 1_{\{\eta_{t-1} < 0\}} + \beta_1, \ldots, \alpha_{r+} 1_{\{\eta_{t-r} > 0\}} + \alpha_{r-} 1_{\{\eta_{t-r} < 0\}} + \beta_r)$.

Under **A2**, we then have

$$\log\sigma_t^2(\boldsymbol{\theta}_0)-\log\sigma_t^2(\boldsymbol{\theta})=\mathcal{B}_{\boldsymbol{\theta}}^{-1}(B)\left\{\omega_0-\omega+\mathbf{V}_{\boldsymbol{\theta}_0-\boldsymbol{\theta}}\boldsymbol{\sigma}_{t-1,r}\right.$$
$$\left.+(\mathcal{A}^+_{\boldsymbol{\theta}_0-\boldsymbol{\theta}}(B)1_{\eta_t>0}\log\eta_t^2+\mathcal{A}^-_{\boldsymbol{\theta}_0-\boldsymbol{\theta}}(B)1_{\eta_t<0}\log\eta_t^2\right\}.$$

Under **A7** the assumptions of Proposition 3.1 hold. From the proof of that proposition, we thus have that $E\exp(\delta\|\mathrm{Abs}(\boldsymbol{\sigma}_{t,r})\|)$ is finite for some $\delta>0$.

Now, note that $\mathbf{V}_{\boldsymbol{\theta}}$, $\mathcal{A}^+_{\boldsymbol{\theta}}(1)$ and $\mathcal{A}^+_{\boldsymbol{\theta}}(1)$ are continuous functions of $\boldsymbol{\theta}$. Choosing a sufficiently small neighborhood $\mathcal{V}$ of $\boldsymbol{\theta}_0$, one can make $\sup_{\mathcal{V}}\|\mathbf{V}_{\boldsymbol{\theta}_0-\boldsymbol{\theta}}\|$, $\sup_{\mathcal{V}}|\mathcal{A}^+_{\boldsymbol{\theta}_0-\boldsymbol{\theta}}(1)|$ and $\sup_{\mathcal{V}}|\mathcal{A}^+_{\boldsymbol{\theta}_0-\boldsymbol{\theta}}(1)|$ arbitrarily small. Thus $E[\exp(m\sup_{\mathcal{V}}\|\mathbf{V}_{\boldsymbol{\theta}_0-\boldsymbol{\theta}}\boldsymbol{\sigma}_{t,r}\|)]$ and $E[\exp(m\sup_{\mathcal{V}}\|(\mathcal{A}^+_{\boldsymbol{\theta}_0-\boldsymbol{\theta}}(B)1_{\eta_{t-1}>0}+\mathcal{A}^-_{\boldsymbol{\theta}_0-\boldsymbol{\theta}}(B)1_{\eta_{t-1}<0})\log(\eta_{t-1}^2)\|)]$ are finite for an appropriate choice of $\mathcal{V}$ depending on m. We conclude that $E\left[\exp\left(m\sup_{\mathcal{V}}\left|\log\left\{\sigma_t^2(\boldsymbol{\theta}_0)/\sigma_t^2(\boldsymbol{\theta})\right\}\right|\right)\right]<\infty$ and the first assertion of the lemma is proved.

Consider now the second assertion. We have

$$\sup_{\mathcal{V}}|\log\sigma_t^2(\boldsymbol{\theta})|\le|\log\sigma_t^2|+\sup_{\mathcal{V}}|\log(\sigma_t^2(\boldsymbol{\theta}_0)/\sigma_t^2(\boldsymbol{\theta}))|.$$

We have already shown that the second term admits a finite moment of order m. So does the first term, under **A7**, by Remark 2.2. □

Now let us prove that the point 1. follows from the fact that $\nabla Q_n(\boldsymbol{\theta}_0)$ is a martingale in L^2. Indeed

$$\nabla Q_n(\boldsymbol{\theta}_0)=n^{-1}\sum_{t=r_0+1}^{n}(1-\eta_t^2)\nabla\log\sigma_t^2(\boldsymbol{\theta}_0).$$

As η_t is independent of $\log\sigma_t^2(\boldsymbol{\theta}_0)$ and $E\eta_t^2=1$ the Central Limit Theorem for martingale differences applies whenever $\boldsymbol{Q}=(\kappa_4-1)E(\nabla\log\sigma_t^2(\boldsymbol{\theta}_0)\nabla\log\sigma_t^2(\boldsymbol{\theta}_0)')$ exists. For any $\boldsymbol{\theta}\in\overset{\circ}{\Theta}$, the random vector

$\nabla \log \sigma_t^2(\boldsymbol{\theta})$ is the stationary solution of the equation

$$\nabla \log \sigma_t^2(\boldsymbol{\theta}) = \sum_{j=1}^{p} \beta_j \nabla \log \sigma_{t-j}^2(\boldsymbol{\theta}) + \begin{pmatrix} 1 \\ \boldsymbol{\epsilon}_{t-1,q}^{+} \\ \boldsymbol{\epsilon}_{t-1,q}^{-} \\ \boldsymbol{\sigma}_{t-1,p}^2(\boldsymbol{\theta}) \end{pmatrix}, \tag{7.10}$$

where $\boldsymbol{\sigma}_{t,p}^2(\boldsymbol{\theta}) = (\log \sigma_t^2(\boldsymbol{\theta}), \dots, \log \sigma_{t-p+1}^2(\boldsymbol{\theta}))'$.

Assumption **A2** entails that $\nabla \log \sigma_t^2(\boldsymbol{\theta})$ is a linear combination of $\boldsymbol{\epsilon}_{t-i,q}^{+}$, $\boldsymbol{\epsilon}_{t-i,q}^{-}$ and $\log \sigma_{t-i}^2(\boldsymbol{\theta})$ for $i \geq 1$. Lemma 7.2 ensures that, for any $m > 0$, there exists a neighborhood $\mathcal{V}$ of $\boldsymbol{\theta}_0$ such that $E[\sup_{\mathcal{V}} |\log \sigma_{t-i}^2(\boldsymbol{\theta})|^m] < \infty$. By Remark 2.2, $\boldsymbol{\epsilon}_{t-i,q}^{+}$ and $\boldsymbol{\epsilon}_{t-i,q}^{-}$ admit moments of any order. Thus, for any $m > 0$ there exists $\mathcal{V}$ such that $E[\sup_{\mathcal{V}} \|\nabla \log \sigma_t^2(\boldsymbol{\theta})\|^m] < \infty$. In particular, $\nabla \log \sigma_t^2(\boldsymbol{\theta}_0)$ admits moments of any order. Thus point 1. is proved.

Turning to point 2., we have

$$\mathbb{H}Q_n(\boldsymbol{\theta}) = n^{-1} \sum_{t=r_0+1}^{n} \mathbb{H}\ell_t(\boldsymbol{\theta}),$$

where

$$\mathbb{H}\ell_t(\boldsymbol{\theta}) = \left(1 - \frac{\eta_t^2 \sigma_t^2(\boldsymbol{\theta}_0)}{\sigma_t^2(\boldsymbol{\theta})}\right) \mathbb{H} \log \sigma_t^2(\boldsymbol{\theta}) + \frac{\eta_t^2 \sigma_t^2(\boldsymbol{\theta}_0)}{\sigma_t^2(\boldsymbol{\theta})} \nabla \log \sigma_t^2(\boldsymbol{\theta}) \nabla \log \sigma_t^2(\boldsymbol{\theta})'. \tag{7.11}$$

By Lemma 7.2, the term $\sigma_t^2(\boldsymbol{\theta}_0)/\sigma_t^2(\boldsymbol{\theta})$ admits moments of order as large as we need uniformly on a well chosen neighborhood $\mathcal{V}$ of $\boldsymbol{\theta}_0$. Let us prove that it is also the case for $\mathbb{H} \log \sigma_t^2(\boldsymbol{\theta})$. Computation gives

$$\mathbb{H} \log \sigma_t^2(\boldsymbol{\theta}) = \sum_{j=1}^{p} \beta_j \mathbb{H} \log \sigma_{t-j}^2(\boldsymbol{\theta}) + \begin{pmatrix} \mathbf{0}_{(2q+1)\times d} \\ \nabla' \boldsymbol{\sigma}_{t-1,p}^2(\boldsymbol{\theta}) \end{pmatrix} + \begin{pmatrix} \mathbf{0}_{(2q+1)\times d} \\ \nabla' \boldsymbol{\sigma}_{t-1,p}^2(\boldsymbol{\theta}) \end{pmatrix}'.$$

From this relation and **A2** we obtain

$$\mathbb{H} \log \sigma_t^2(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{0}_{(2q+1)\times d} \\ \mathcal{B}_{\boldsymbol{\theta}}(B)^{-1} \nabla' \boldsymbol{\sigma}_{t-1,p}^2(\boldsymbol{\theta}) \end{pmatrix} + \begin{pmatrix} \mathbf{0}_{(2q+1)\times d} \\ \mathcal{B}_{\boldsymbol{\theta}}(B)^{-1} \nabla' \boldsymbol{\sigma}_{t-1,p}^2(\boldsymbol{\theta}) \end{pmatrix}'.$$

Thus $\mathbb{H} \log \sigma_t^2(\boldsymbol{\theta})$ belongs to $\mathcal{C}(\mathcal{V})$ and is integrable because we can always choose $\mathcal{V}$ such that $\sup_{\mathcal{V}} \|\nabla' \boldsymbol{\sigma}_{t-1,p}^2(\boldsymbol{\theta})\| \in L^m$ (see the proof of point 1. above).

An application of the Cauchy-Schwarz inequality in the RHS term of (7.11) yields the integrability of $\sup_{\mathcal{V}} \mathbb{H}\ell_t(\boldsymbol{\theta})$. The first assertion of point 2. is proved by an application of the ergodic theorem on $(\mathbb{H}\ell_t(\boldsymbol{\theta}))$ in the Banach space $\mathcal{C}(\mathcal{V})$ equipped with the supremum norm:

$$\sup_{\mathcal{V}} \|\mathbb{H}Q_n(\boldsymbol{\theta}) - E[\mathbb{H}\ell_0(\boldsymbol{\theta})]\| \to 0 \qquad a.s.$$

An application of Theorem 4.1 ensures that $\hat{\boldsymbol{\theta}}_n$ belongs a.s. to $\mathcal{V}$ for sufficiently large n. Thus

$$\|\mathbb{H}Q_n(\hat{\boldsymbol{\theta}}_n)-E[\mathbb{H}\ell_0(\boldsymbol{\theta}_0)]\| \leq \sup_{\mathcal{V}} \|\mathbb{H}Q_n(\boldsymbol{\theta})-E[\mathbb{H}\ell_0(\boldsymbol{\theta})]\|+\|E[\mathbb{H}\ell_0(\hat{\boldsymbol{\theta}}_n)]-E[\mathbb{H}\ell_0(\boldsymbol{\theta}_0)]\|$$

converges a.s. to 0 by continuity of $\boldsymbol{\theta} \to E[\mathbb{H}\ell_0(\boldsymbol{\theta})]$ at $\boldsymbol{\theta}_0$ as a consequence of a dominating argument on $\mathcal{V}$. The first assertion of point 2. is proved.

The matrix $\boldsymbol{J}$ is non invertible iff there exists a non null deterministic vector $\boldsymbol{\lambda} = (\lambda_0, \lambda_1, \dots, \lambda_{p+2q})'$ such that $\boldsymbol{\lambda}' \nabla \log \sigma_t^2(\boldsymbol{\theta}_0) = 0$ a.s. If the latter equality holds, (7.10) entails $(1, \boldsymbol{\epsilon}_{t,q}^+, \boldsymbol{\epsilon}_{t,q}^-, \boldsymbol{\sigma}_{t,p}^2(\boldsymbol{\theta}_0))\,\boldsymbol{\lambda} = 0$ a.s. By **A3** and arguments used for proving the point (ii) in the proof of Theorem 4.1, we deduce $\lambda_{p+1} = \lambda_{p+q+1} = 0$. For the same reason if $\lambda_1 = \cdots = \lambda_i = 0$ then $\lambda_{p+2} = \cdots = \lambda_{p+1+i} = 0$ and $\lambda_{p+q+2} = \cdots = \lambda_{p+q+1+i} = 0$ for $1 \leq i \leq p \wedge q$. Thus as $\boldsymbol{\lambda} \neq 0$ there exists $\lambda_k \neq 0$ for $1 \leq k \leq q$ such that $\lambda_j = 0$ for $j < k$. Then, $\log(\sigma_{t-k}^2)$ is a linear combination of $(\log(\sigma_{t-j}^2))_{k<j\leq q}$, $(1_{\{\epsilon_{t-j}>0\}} \log(\epsilon_{t-j}^2))_{k<j\leq p}$ and $(1_{\{\epsilon_{t-j}<0\}} \log(\epsilon_{t-j}^2))_{k<j\leq p}$. We thus find a log-GARCH(p', q') representation with $p' < p$ and $q' < q$, in contradiction with **A4**. Thus $\boldsymbol{J}$ is invertible.

From (7.10) and an equivalent representation for $\nabla \log \widetilde{\sigma}_t^2(\boldsymbol{\theta})$, we have

$$\begin{aligned}\nabla \log \sigma_t^2(\boldsymbol{\theta}) - \nabla \log \widetilde{\sigma}_t^2(\boldsymbol{\theta}) &= \sum_{j=1}^{p} \beta_j (\nabla \log \sigma_{t-j}^2(\boldsymbol{\theta}) - \nabla \log \widetilde{\sigma}_{t-j}^2(\boldsymbol{\theta})) \\ &\quad + \begin{pmatrix} \mathbf{0}_{2q+1} \\ \boldsymbol{\sigma}_{t-1,p}^2(\boldsymbol{\theta}) - \widetilde{\boldsymbol{\sigma}}_{t-1,p}^2(\boldsymbol{\theta}) \end{pmatrix}\end{aligned}$$

where $\widetilde{\boldsymbol{\sigma}}^2_{t,p}$ is defined as $\boldsymbol{\sigma}^2_{t,p}$. Thus, there exist continuous functions d_i and $d_{t,i}$ defined on Θ such that

$$\begin{aligned}\nabla \log \sigma_t^2(\boldsymbol{\theta}) - \nabla \log \widetilde{\sigma}_t^2(\boldsymbol{\theta}) &= \sum_{i=1}^{t-1} d_i(\boldsymbol{\theta})(\log \sigma_{t-i}^2(\boldsymbol{\theta}) - \log \widetilde{\sigma}_{t-i}^2(\boldsymbol{\theta})) \\ &\quad + \sum_{j=1}^{p} d_{t,j}(\boldsymbol{\theta}) \nabla \log \sigma_{p+1-j}^2(\boldsymbol{\theta}).\end{aligned}$$

The sequences of functions (d_i), $(d_{i,j})$, $1 \le j \le p$, satisfy the same uniform rate of convergence as the functions c_i, c_{i+}, c_{1-} and $c_{i,j}$ in (7.7). An application of (7.4) yields to the existence of $K > 0$ and $\rho \in (0,1)$ such that $\sup_\Theta \|\nabla \log \sigma_t^2(\boldsymbol{\theta}) - \nabla \log \widetilde{\sigma}_t^2(\boldsymbol{\theta})\| \le K\rho^t$, for almost all trajectories. Point 3. follows easily and the asymptotic normality is proved. □

8. Conclusion

In this paper, we investigated the probabilistic properties of the log-GARCH(p,q) model. We found sufficient conditions for the existence of moments and log-moments of the strictly stationary solutions. We analyzed the dependence structure through the leverage effect and the regular variation properties, and we compared this structure with that of the EGARCH model.

As far as the estimation is concerned, it should be emphasized that the log-GARCH model appears to be much more tractable than the EGARCH. Indeed, we established the strong consistency and the asymptotic normality of the QMLE under mild assumptions. For EGARCH models, such properties have only been established for the first-order model and with strong invertibility constraints (see Wintenberger and Cai, 2011). By comparison with standard GARCH, the log-GARCH model is not more difficult to handle: on the one hand, the fact that the volatility is not bounded below requires an additional log-moment assumption, but on the other hand the parameters are nor positively constrained.

A natural extension of this work, aiming at pursuing the comparison between the two classes of models, would rely on statistical tests. By embedding the log-GARCH model in a more general framework including the log-GARCH, it should be possible to consider a LM test of the log-GARCH null assumption. Another problem of interest would be to check validity of the estimated models. We leave these issues for further investigation, viewing the results of this paper as a first step in these directions.

References

Allen, D. , Chan, F., McAleer, M., and Peiris, S. (2008) Finite sample properties of the QMLE for the log-ACD model: Application to Australian stocks. *Journal of Econometrics* 147, 163–183.

Bardet, J.-M. and Wintenberger, O. (2009) Asymptotic normality of the Quasi Maximum Likelihood Estimator for multidimensional causal processes. *Ann. Statist.* 37, 2730–2759.

Bauwens, L. and Giot, P. (2000) The logarithmic ACD model: An application to the bidask quote process of three NYSE stocks. *Annales D'Economie et de Statistique* 60, 117–145.

Bauwens, L. and Giot, P. (2003) Asymmetric ACD models: introducing price information in ACD models. *Empirical Economics* 28, 709–731.

Bauwens, L., Galli, F. and Giot, P. (2008) The moments of log-ACD models. *QASS* 2, 1–28.

Bauwens, L., Giot, P., Grammig, J. and Veredas, D. (2004) A comparison of financial duration models via density forecast. *International Journal of Forecasting* 20, 589–604.

Berkes, I., Horváth, L. and Kokoszka, P. (2003) GARCH processes: structure and estimation. *Bernoulli* 9, 201–227.

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

Bougerol, P. and Picard, N. (1992a) Strict stationarity of generalized autoregressive processes. *Annals of Probability* 20, 1714–1729.

Bougerol, P. and Picard, N. (1992b) Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* 52, 115–127.

Drost, F.C. and Nijman, T.E. (1993) Temporal Aggregation of GARCH Processes. *Econometrica* 61, 909–927.

Engle, R.F. (1982) Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica* 50, 987–1008.

Engle, R.F. and Russell, J.R. (1998) Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. *Econometrica* 66, 1127–1162.

Francq, C. and Zakoïan, J-M. (2004) Maximum Likelihood Estimation of Pure GARCH and ARMA-GARCH. *Bernoulli* 10, 605–637.

Francq, C. and Zakoïan, J-M. (2010a) *GARCH Models : structure, statistical inference and financial applications*. John Wiley.

Francq, C. and Zakoïan, J-M. (2010b) Inconsistency of the MLE and inference based on weighted LS for LARCH models. *Journal of Econometrics* 159, 151–165.

Geweke, J. (1986) Modeling the Persistence of Conditional Variances: A Comment. *Econometric Review* 5, 57–61.

He, C., Teräsvirta, T. and Malmsten, H. (2002) Moment structure of a family of first-order exponential GARCH models. *Econometric Theory* 18, 868–885.

Mikosch, T. and Rezapour, M. (2012) Stochastic volatility models with possible extremal clustering. *Bernoulli* To appear.

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

Milhøj, A. (1987) A Multiplicative Parameterization of ARCH Models. *Working paper*, Department of Statistics, University of Copenhagen.

Nelson D.B. (1991) Conditional Heteroskedasticity in Asset Returns : a New Approach. *Econometrica* 59, 347–370.

Pantula, S.G. (1986) Modeling the Persistence of Conditional Variances: A Comment. *Econometric Review* 5, 71–74.

Sucarrat, G. and Escribano, A. (2010) The Power Log-GARCH Model. *Working document*, Economic Series 10-13, University Carlos III, Madrid.

Sucarrat, G. and Escribano, A. (2012) Automated Model Selection in Finance: General-to-Specific Modelling of the Mean and Volatility Specifications. *Oxford Bulletin Of Economics And Statistics* 74, 716–735.

Wintenberger, O. and Cai, S. (2011) Parametric inference and forecasting in continuously invertible volatility models. Preprint arXiv:1106.4983v3.