Asymptotic behavior of the Laplacian quasi-maximum likelihood estimator of affine causal processes
Jean-Marc Bardet, Yakoub Boularouk, Khedidja Djaballah

To cite this version:

HAL Id: hal-01249503
https://hal.archives-ouvertes.fr/hal-01249503v2
Submitted on 20 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License
Asymptotic behavior of the Laplacian quasi-maximum likelihood estimator of affine causal processes

Jean-Marc Bardet1,† Yakoub Boularouk2,*** and Khedija Djaballah2,†

1 S.A.M.M., Université Panthéon-Sorbonne, 90, rue de Tolbiac, 75634, Paris, France. e-mail: bardet@univ-paris1.fr
2 Université des Sciences et de la Technologie Houari Boumédiène, Alger, Algérie. e-mail: y.boularouk@centre-univ-mila.dz, khdjeddour@hotmail.com

Abstract: We prove the consistency and asymptotic normality of the Laplacian Quasi-Maximum Likelihood Estimator (QMLE) for a general class of causal time series including ARMA, AR(∞), GARCH, ARCH(∞), ARMA-GARCH, APARCH, ARMA-APARCH,..., processes. We notably exhibit the advantages (moment order and robustness) of this estimator compared to the classical Gaussian QMLE. Numerical simulations confirms the accuracy of this estimator.


Keywords and phrases: Laplacian Quasi-Maximum Likelihood Estimator, Strong consistency, Asymptotic normality, ARMA-ARCH processes.

1. Introduction

This paper is devoted to establish the consistency and the asymptotic normality of a parametric estimator for a general class of time series. This class was already defined and studied in Doukhan and Wintenberger (2007), Bardet and Wintenberger (2009) and Bardet et al. (2012). Hence, we will consider an observed sample \((X_1, \ldots, X_n)\) where \((X_t)_{t \in \mathbb{Z}}\) is a solution of the following equation:

\[
X_t = M_0(X_{t-1}, X_{t-2}, \ldots) \zeta_t + f_0(X_{t-1}, X_{t-2}, \ldots), \quad t \in \mathbb{Z},
\]  

(1.1)

where

- \(\theta_0 \in \Theta \subset \mathbb{R}^d, d \in \mathbb{N}^*,\) is an unknown vector of parameters, also called the ”true” parameters;
- \((\zeta_t)_{t \in \mathbb{Z}}\) is a sequence of centred independent identically distributed random variables (i.i.d.r.v.) with symmetric probability distribution, i.e. \(\zeta_0 \overset{\mathcal{L}}{=} -\zeta_0,\) satisfying \(E[|\zeta_0|^r] < \infty\) with \(r \geq 1\) and \(E[|\zeta_0|] = 1.\) If \(r \geq 2,\) denote \(\sigma^2_\zeta = \text{Var}(\zeta_0);\)
- \((\theta, (x_n)_{n \in \mathbb{N}}) \rightarrow M_0((x_n)_{n \in \mathbb{N}}) \in (0, \infty)\) and \((\theta, (x_n)_{n \in \mathbb{N}}) \rightarrow f_0((x_n)_{n \in \mathbb{N}}) \in \mathbb{R}\) are two known applications.

For instance, if \(M_0(X_{t-1}, X_{t-2}, \ldots) = 1\) and \(f_0(X_{t-1}, X_{t-2}, \ldots) = \alpha_0 X_{t-1} \) with \(|\alpha_0| < 1\) then \((X_t)\) is a causal AR(1) process. In Doukhan and Wintenberger (2007) and Bardet and Wintenberger (2009), it was proved that all the most famous stationary time series used in econometrics, such as ARMA, AR(∞), GARCH, ARCH(∞), TARCH, ARMA-GARCH processes can be written as a causal stationary solution of (1.1).

In Bardet and Wintenberger (2009), it was also established that under several conditions on \(M_\theta, f_\theta\) and if \(E[|\zeta_0|^r]\) with \(r \geq 2,\) the usual Gaussian Quasi-Maximum Likelihood Estimator (QMLE) of \(\theta\) is strongly consistent and when \(r \geq 4\) it is asymptotically normal. This estimator was first defined by Weiss (1986) for AR(1) processes, and the asymptotic study of this estimator was first obtained by Lumsdaine (1996) for GARCH(1,1) processes, Berkes et al. (2003) for GARCH\((p, q)\) processes, Francq and Zakoian (2004) for ARMA-GARCH processes, Straumann and Mikosch (2006) for general heteroskedastic models, and Robinson and Zaffaroni (2006) for ARCH(∞) processes. The results of Bardet and Wintenberger (2009) devoted to processes satisfying almost everywhere (1.1) as well as its multivariate generalisation, provide a general and unified framework for studying the asymptotic properties of the Gaussian QMLE.

However, the definition of the Gaussian QMLE is explicitly obtained with the assumption that \((\zeta_t)\) is a Gaussian sequence and even if it could be applied when the probability distribution of \((\zeta_t)\) is non-Gaussian, it keeps some drawbacks of this initial assumption. Indeed, the computation of this estimators requires the minimization of a least squares contrast (typically \(\sum_{t=1}^n M^{-2}_\theta(X_t - f_\theta)\)) and this is usually done by the least squares method for the asymptotic normality (and therefore confidence intervals or tests). For numerous real data such requirement...
ARMA-ARCH($\infty$) and ARMA-APARCH processes.
Numerical Monte-Carlo experiments were realized to illustrate the theoretical results. And the results of these simulations are convincing, especially when the accuracy of Laplacian-QMLE is compared with the one of Gaussian-QMLE: except for Gaussian distribution of ($\xi_t$), the Laplacian-QMLE provides a sharper estimation than the Gaussian-QMLE for all the other probability distributions we considered. This is notably the case, and this is not a surprise, for a Gaussian mixing which mimics the presence of outliers. This provides an effective advantage of the Laplacian QMLE compared to the Gaussian QMLE.

The following Section 2 will be devoted to provide the definitions and assumptions. In Section 3 the main results are stated with numerous examples of application, while Section 4 presents the results of Monte-Carlo experiments and Section 5 contains the proofs.

2. Definition and assumptions

2.1. Definition of the estimator

Let ($X_1, \cdots, X_n$) be an observed trajectory of $X$ which is an a.s. solution of (1.1) where $\theta \in \Theta \subset \mathbb{R}^d$ is unknown. For estimating $\theta$ we consider the log-likelihood of ($X_1, \cdots, X_n$) conditionally to ($X_0, X_{-1}, \cdots$). If $h$ is the probability density (with respect to Lebesgue measure) of $\zeta_0$, then, from the affine causal definition of $X$, this log-likelihood can be written:

$$
\log \left( L_0(X_1, \cdots, X_n) \right) = \sum_{t=1}^{n} \log \left( \frac{1}{M_t^\theta} h \left( \frac{X_t - f_t^\theta}{M_t^\theta} \right) \right)
$$

where $M_t^\theta := M_\theta(X_{t-1}, X_{t-2}, \cdots)$ and $f_t^\theta := f_\theta(X_{t-1}, X_{t-2}, \cdots)$, with the assumption that $M_t^\theta > 0$. However, $M_t^\theta$ and $f_t^\theta$ are generally not computable since $X_0, X_{-1}, \cdots$ are unknown. Thus, a quasi-log-likelihood is considered instead of the log-likelihood and it is defined by:

$$
\log \left( QL_\theta(X_1, \cdots, X_n) \right) = \sum_{t=1}^{n} \log \left( \frac{1}{\hat{M}_t^\theta} h \left( \frac{X_t - \hat{f}_t^\theta}{\hat{M}_t^\theta} \right) \right),
$$

with $\hat{f}_t^\theta := f_\theta(X_{t-1}, \cdots, X_1, u)$ and $\hat{M}_t^\theta := M_\theta(X_{t-1}, \cdots, X_1, u)$, where $u = (u_n)_{n \in \mathbb{N}}$ is a finitely non-zero sequence. The choice of $(u_n)_{n \in \mathbb{N}}$ does not have any consequences on the asymptotic behaviour of $L_n$, and $(u_n)$ could typically be chosen as a sequence of zeros. Finally, if it exists, a Quasi-Maximum Likelihood Estimator (QMLE) is defined by:

$$
\tilde{\theta}_n := \text{Argmax}_{\theta \in \Theta} \log \left( QL_\theta(X_1, \cdots, X_n) \right).
$$

Usually, the "instrumental" probability density $h$ is the Gaussian density, i.e.

$$
h(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \text{ for } x \in \mathbb{R}
$$

and this provides the Gaussian-QMLE of $\theta$.

Here, we chose as instrumental probability density the Laplacian density, i.e.,

$$
h(x) = \frac{1}{2} e^{-|x|} \text{ for } x \in \mathbb{R}, \tag{2.1}
$$

and this implies $E[|\xi_0|] = 1$.

Therefore, we respectively define the Laplacian-likelihood and Laplacian-quasi-likelihood by:

$$
\begin{align*}
L_n(\theta) & = -\sum_{t=1}^{n} q_t(\theta) \quad \text{with} \quad q_t(\theta) = \log |M_t^\theta| + |M_t^\theta|^{-1} |X_t - f_t^\theta| \tag{2.2} \\
\hat{L}_n(\theta) & = -\sum_{t=1}^{n} \hat{q}_t(\theta) \quad \text{with} \quad \hat{q}_t(\theta) := \log |\hat{M}_t^\theta| + |\hat{M}_t^\theta|^{-1} |X_t - \hat{f}_t^\theta|, \tag{2.3}
\end{align*}
$$
Hence, if it exists, a Laplacian-QMLE \( \hat{\theta}_n \) is a maximizer of \( \hat{L}_n \):
\[
\hat{\theta}_n := \arg \max_{\theta \in \Theta} \hat{L}_n(\theta).
\]
We restrict the set \( \Theta \) in such a way that a stationary solution \( (X_t) \) of order 1 or 2 of (1.1) exists. Additional conditions are also required for insuring the consistency and the asymptotic normality of \( \hat{\theta}_n \). More details are given now.

2.2. Existence and stationarity

As it was already done in Doukhan and Wintenberger (2007) and Bardet and Wintenberger (2009), several Lipschitz-type inequalities on \( f_\theta \) and \( M_\theta \) are required for obtaining the existence and \( r \)-order stationary ergodic causal solution of (1.1).

First, denote \( \|g_\theta\|_\Theta = \sup_{\theta \in \Theta} \|g_\theta\| \) with \( m \in \mathbb{N}^* \) and \( \| \cdot \| \) the usual Euclidean norm (for vectors or matrix). Now, let us introduce the generic symbol \( K \) for any of the functions \( f \) or \( M \). For \( k = 0, 1, 2 \) and some subset \( \Theta \) of \( \mathbb{R}^d \), define a Lipschitz assumption on function \( K_{\theta} \):

**Assumption (A\( k(K, \Theta) \))** \( \forall x \in \mathbb{R}^\infty, \theta \in \Theta \Rightarrow K_{\theta}(x) \in C^k(\Theta) \) and \( \partial^K_{\theta} K_{\theta} \) satisfies \( \| \partial^K_{\theta} K_{\theta}(0) \|_\Theta < \infty \) and there exists a sequence \( (\alpha_j^{(k)}(K, \Theta))_j \) of nonnegative numbers such that \( \forall x, y \in \mathbb{R}^N \)
\[
\| \partial^K_{\theta} K_{\theta}(x) - \partial^K_{\theta} K_{\theta}(y) \|_\Theta \leq \sum_{j=1}^{\infty} \alpha_j^{(k)}(K, \Theta) |x_j - y_j|, \text{ with } \sum_{j=1}^{\infty} \alpha_j^{(k)}(K, \Theta) < \infty.
\]

For ensuring a stationary \( r \)-order solution of (1.1), for \( r \geq 1 \), define the set
\[
\Theta(r) := \left\{ \theta \in \mathbb{R}^d, (A_0(f, \{\theta\})) \text{ and } (A_0(M, \{\theta\})) \right\},
\]
\[
\sum_{j=1}^{\infty} \alpha_j^{(0)}(f, \{\theta\}) + (E[|\zeta_0|^r])^{1/r} \sum_{j=1}^{\infty} \alpha_j^{(0)}(M, \{\theta\}) < 1.
\]

Then, from Doukhan and Wintenberger (2007), we obtain:

**Proposition 2.1.** If \( \theta_0 \in \Theta(r) \) for some \( r \geq 1 \), then there exists a unique causal \( (X_t) \) is independent of \( (\zeta_i)_{i \geq 1} \) for \( t \in \mathbb{Z} \) solution \( X \) of (1.1), which is stationary, ergodic and satisfies \( E[|X_0|^r] < \infty \).

The following lemma insures that if a process \( X \) satisfies Proposition 2.1, a causal predictable ARMA process with \( X \) as innovation also satisfies Proposition 2.1. We first provide the classical following notion for a sequence \( (u_n)_{n \in \mathbb{N}} \) of real numbers:

\( (u_n)_{n \in \mathbb{N}} \) is an exponentially decreasing sequence (EDS)
\[ \iff \exists \rho \in [0, 1[ \text{ such that } u_n = O(\rho^n) \text{ when } n \to \infty. \]

**Lemma 2.1.** Let \( X \) be a.s. a causal stationary solution of (1.1) for \( \theta_0 \in \mathbb{R}^d \). Let \( \tilde{X} \) be such that \( \tilde{X}_t = \Lambda_{\beta}(L) X_t \) for \( t \in \mathbb{Z} \) with \( \Lambda_{\beta}(L) = P^{-1}_\beta(L) Q_{\beta_0}(L) \) where \( P_{\beta_0}, Q_{\beta_0} \) are the coprime polynomials of a causal invertible ARMA\( (p, q) \) processes with a vector of parameters \( \beta_0 \in \mathbb{R}^{p+q} \). Denote \( \Lambda_{\beta_0}^{-1}(x) = Q_{\beta_0}^{-1}(x) P_{\beta_0}(x) = 1 + \sum_{j=1}^{\infty} \psi_j(\beta_0)x^j \). Then \( \tilde{X} \) is a.s. a causal stationary solution of the equation
\[
\tilde{X}_t = \tilde{M}_{\theta_0}((\tilde{X}_{t-i})_{i \geq 1}) \zeta_t + \tilde{f}_{\theta_0}((\tilde{X}_{t-i})_{i \geq 1}) \text{ for } t \in \mathbb{Z},
\]
where \( \tilde{f}_{\theta_0} \) and \( \tilde{M}_{\theta_0} \) are given in (5.1) and \( \theta_0 = (\theta_0, \beta_0) \). Moreover, for \( i = 0, 1, 2 \) and with \( K = f \) or \( M \) and \( \tilde{K} = \tilde{f} \) or \( \tilde{M} \),
- if \( \alpha_j^{(i)}(K, \theta_0) = O(j^{-\beta}) \) and \( \beta > 1 \), then \( \alpha_j^{(i)}(\tilde{K}, \{\tilde{\theta}_0\}) = O(j^{-\beta}) \);
- if \( \alpha_j^{(i)}(K, \{\theta_0\}) \) is EDS, then \( \alpha_j^{(i)}(\tilde{K}, \{\tilde{\theta}_0\}) \) is EDS.
2.3. Assumptions required for the convergence of the Laplacian-QMLE

The Laplacian-QMLE could converge and be asymptotically Gaussian but this requires some additional assumptions on Θ and functions $f_\theta$ and $M_\theta$:

- **Condition C1** (Compactness) Θ is a compact set.
- **Condition C2** (Lower bound of the conditional variance) There exists a deterministic constant $M > 0$ such that for all $\theta \in \Theta$ and $x \in \mathbb{R}^N$, then $M_\theta(x) > M$.
- **Condition C3** (Identifiability) The functions $M_\theta$ and $f_\theta$ are such that: for all $\theta_1$, $\theta_2 \in \Theta$, then $M_{\theta_1} = M_{\theta_2}$ and $f_{\theta_1} = f_{\theta_2}$ implies that $\theta_1 = \theta_2$.

3. Consistency and asymptotic normality of the estimator

3.1. Consistency and asymptotic normality

First we prove the strong consistency of a sequence of Laplacian-QMLE for a solution of (1.1). The proof of this theorem, is postponed in Section 5, as well as the other proofs.

**Theorem 3.1.** Assume Conditions C1, C2 and C3 hold and $\theta_0 \in \Theta(r) \cap \Theta$ with $r \geq 1$. Let $X$ be the stationary solution of (1.1). If $(A_0(f, \Theta))$ and $(A_0(M, \Theta))$ hold with

$$
\alpha_j^{(0)}(f, \Theta) + \alpha_j^{(0)}(M, \Theta) = O(j^{-\ell}) \text{ for some } \ell > \frac{2}{\min(r, 2)}
$$

then a sequence of Laplacian-QMLE $(\hat{\theta}_n)_n$ strongly converges, that is $\hat{\theta}_n \overset{a.s.}{\to} \theta_0$.

Of course, the conditions required for this strong consistency of a sequence of Laplacian-QMLE are almost the same than the ones required for the strong consistency of a sequence of Gaussian-QMLE except that $r \in [1, 2)$ is proved to be possible in Theorem 3.1 and not in case of Gaussian-QMLE (see Bardet and Wintenberger (2009)). Moreover, if $r = 2$, the condition (3.1) on Lipschitzian coefficients is weaker for Laplacian-QMLE than for Gaussian-QMLE. As we will see below, many usual time series can satisfy the assumptions of Theorem 3.1: for example, an AR($\infty$) process can be defined for satisfying the strong consistency of Laplacian-QMLE while the conditions given in Bardet and Wintenberger (2009) do not ensure the strong consistency of Gaussian-QMLE.

Now we state an extension of Theorem 1 established in Davis and Dunsmuir (1997) which will be an essential step of the proof of the asymptotic normality of the estimator.

**Theorem 3.2.** Let $(Z_t)_{t \in \mathbb{Z}}$ be a sequence of i.i.d.r.v such as $\text{Var}(Z_0) = \sigma^2 < \infty$, with common distribution function which is symmetric ($F(-x) = 1 - F(x)$ for $x \in \mathbb{R}$) and is continuously differentiable in a neighborhood of 0 with derivative $f(0)$ in 0. Denote $\mathcal{F}_t = \sigma(Z_t, Z_{t-1}, \cdots)$ for $t \in \mathbb{Z}$ and let $(Y_t)_{t \in \mathbb{Z}}$ and $(V_t)_{t \in \mathbb{Z}}$ two stationary processes adapted to $(\mathcal{F}_t)_t$ and such as $E[Y_0^2V_0^2] < \infty$. Then

$$
\sum_{t=1}^n V_{t-1}(|Z_t - n^{-1/2}Y_{t-1}| - |Z_t|) \overset{D}{\to} \mathcal{N}_d\left(\mathcal{N}\left(f(0) E[V_0^2], E[V_0^2V_0^2]\right)\right)
$$

Then, the asymptotic normality of the Laplacian-QMLE can be established using additional assumptions:

**Theorem 3.3.** Assume that $\theta_0 \in \Theta(r)$ where $r \geq 2$ and $\Theta$ denotes the interior of $\Theta$. Let $X$ be the stationary solution of the equation (1.1). Assume that the conditions of Theorem 3.1 hold and for $i = 1, 2$, assume $(A_i(f, \Theta))$ and $(A_i(M, \Theta))$ hold. Then, if the cumulative probability function of $\tilde{\theta}_0$ is continuously differentiable in a neighborhood of 0 with derivative $g(0)$ in 0 and if matrix $\Gamma_F$ or $\Gamma_M$, defined in (5.21), are definite positive symmetric matrix, then

$$
\sqrt{n}(\hat{\theta}_n - \theta_0) \overset{D}{\to} \mathcal{N}_d\left(\mathcal{N}(0, (\Gamma_M + 2g(0)\Gamma_F)^{-1}(\sigma^2 - 1)\Gamma_M + \Gamma_F)((\Gamma_M + 2g(0)\Gamma_F)^{-1})\right).
$$
As it was already proved for the median estimator (see van der Vaart (2000)) or for least absolute deviations estimator of ARMA process (see Davis and Dunsmuir (1997)), it is not surprising that the probability density function $g$ of the white noise $(\xi_i)_i$ impacts the asymptotic covariance of (3.3). However, when $f_\theta = 0$, this is not such the case and this is what happens for GARCH processes see Francq et al. (2011) where the probability density $g$ does not appear in the asymptotic covariance.

### 3.2. Comments on these limit theorems

Essentially, these limit theorems could appear close or even very close to the results of 3 other references we chronologically list below but also from which we highlight the differences:

- The first related paper is Davis and Dunsmuir (1997) which is cited many times. The framework of this paper is restricted to the LAD (similar to the Laplacian-QMLE) of the parameters of ARMA$[p, q]$ process or residuals of least-square estimation with ARMA$[p, q]$ errors. If the framework (1.1) is clearly more general since it includes for instance GARCH, ARMA-GARCH or APARCH process, the proof we used for establishing the asymptotic normality of the Laplacian estimator is clearly inspired by the one of Davis and Dunsmuir (1997). Thus our results could appear as extensions of this paper.

- The second and certainly closest paper, Bardet and Wintenberger (2009). The considered framework is exactly the same, i.e. general causal affine models and the estimation method is the same, i.e. the quasi-maximum likelihood estimation (QMLE). However in Bardet and Wintenberger (2009) the QMLE is based on an ”instrumental” Gaussian density instead of a Laplacian one. As it is such the case for instance by comparing quantile with least square regression, this implies three main differences:

1. The moment conditions $r$ of both the limit theorems (strong consistency and asymptotic normality) are weaker with Laplacian QMLE than with the Gaussian one. Indeed, the absolute value of conditional log-density $q_\theta(\theta)$ is bounded by an affine function of $|X_t|$ in the Laplacian case while it is bounded by a quadratic polynomial of $X_t$ in the Gaussian case. As a consequence, $r = 1$ (respectively $r = 2$) could be required for the strong consistency (resp. asymptotic normality) of the Laplacian QMLE while $r = 2$ (resp. $r = 4$) is required for the Gaussian QMLE. This gain on moment condition can be crucial for instance in an econometric framework where the Kurtosis of data is sometimes infinite.

2. The proof of Theorem 3.1 is simpler and sharper than the proof of strong consistency in Bardet and Wintenberger (2009). Indeed, in our new proof, we use a condition of almost sure uniform consistency based on a general and powerful result established in Kounias and Weng (1969) while a Feller-type condition was ”only” used in Bardet and Wintenberger (2009). This difference leads to a very sharp condition on the decreasing rate of the Lipschitzian coefficients $(\alpha^0_k)$ for Laplacian QMLE, $\ell > 1$ in (3.1), while $\ell > 3/2$ is required for Gaussian QMLE.

3. The proof of Theorem 3.3 is totally different to the one for Gaussian QMLE since the conditional log-density is no more differentiable with respect to the parameters. A kind of proof similar to the one used for establishing the asymptotic normality of the median is required. Hence, in a first step we had to prove an extension of a central limit for adapted processes established in Davis and Dunsmuir (1997), i.e. our Theorem 3.2, and we used it in a second step for establishing the asymptotic normality of the Laplacian QMLE. Note also that the conditions on the derivatives of functions $f_\theta$ and $M_\theta$ are clearly weaker with Laplacian than with Gaussian QMLE.

- The third related paper is Francq et al. (2011). The framework of this paper is restricted to linear causal models $(X_t = \sigma_t(\theta)\xi_t)$ in contrast with the affine causal models $(X_t = M^d_t\xi_t + f^d_t)$ considered in (1.1). Hence ARMA but also ARMA-GARCH or ARMA-APARCH processes are not considered in this framework. Moreover the required moment is $r = 4$ (instead of $r = 2$ in our conditions) and the condition on the approximation of $\sigma_t(\theta)$, i.e. $\sup_{\theta} |\sigma_t(\theta) - \hat{\sigma}_t(\theta)| \leq C_1 \rho^t$ is clearly weaker than our Lipschitzian condition (for instance ARCH$(\infty)$ processes with Riemannian decay of the coefficients could satisfy our conditions but not their conditions). In
Francoq et al. (2011), a large family of instrumental probability densities, *i.e.* generalized Gaussian densities, including Laplace density, but their proof of asymptotic normality mimics the proof using derivatives of Gaussian QMLE since the "shift" component $f^d_\theta$ typically present for ARMA processes, is not considered in their models. Note also that Francoq and Zakoian (2015) also studies non-Gaussian QMLE but their assumption A9 implies that the Laplace density is not considered in their asymptotic normality of the QMLE.

Finally it appears that our results provide an original extension or counterpart of these three related references.

### 3.3. Examples

In this section, several examples of time series satisfying the conditions of previous results are considered. Like it could be boring to state the results for all sufficiently famous processes, we refer, *mutatis mutandis*, to Bardet and Wintenberger (2009) and Bardet et al. (2012) for ARCH($\infty$) and TARCH($\infty$).

1/ **APARCH processes.** APARCH($\delta, p, q$) model has been introduced (see Ding et al. (1993)) as the solution of equations

\[
\begin{align*}
X_t &= \sigma_t \zeta_t, \\
\sigma^2_t &= \omega + \sum_{i=1}^p \alpha_i (|X_{t-i}| - \gamma_i X_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma^\delta_{t-j},
\end{align*}
\]  

(3.4)

where $\delta \geq 1$, $\omega > 0$, $-1 < \gamma_i < 1$ and $\alpha_i \geq 0$ for $i = 1, \ldots, p$, $\beta_j \geq 0$ for $j = 1, \ldots, q$ with $\alpha_p, \beta_q$ strictly positive and $\sum_{j=1}^q \beta_j < 1$. Hence, we denote here $\theta = (\delta, \omega, \alpha_1, \ldots, \alpha_p, \gamma_1, \ldots, \gamma_q, \beta_1, \ldots, \beta_q)$.

Using $L$ the usual backward operator such as $LX_t = X_{t-1}$, $(1 - \sum_{j=1}^q \beta_j L^j)^{-1}$ exists and simple computations imply for $t \in \mathbb{Z}$:

\[
\sigma_t^\delta = \left(1 - \sum_{j=1}^q \beta_j L^j\right)^{-1}\left[\omega + \sum_{i=1}^p \alpha_i (1 - \gamma_i)\delta (\max(X_{t-i}, 0))^\delta + \alpha_i (1 + \gamma_i)\delta (\min(X_{t-i}, 0))^\delta\right]
\]

\[
= b_0 + \sum_{i \geq 1} b^+_i (\max(X_{t-i}, 0))^\delta + \sum_{i \geq 1} b^-_i (\min(-X_{t-i}, 0))^\delta.
\]

where $b_0 = w(1 - \sum_{j=1}^q \beta_j)^{-1}$ and the coefficients $(b^+_i, b^-_i)_{i \geq 1}$ are defined by the recursion relations

\[
\begin{align*}
b^+_i &= \sum_{k=1}^q \beta_k b^+_{i-k} + \alpha_i (1 - \gamma_i)^\delta \quad \text{with} \quad \alpha_i (1 - \gamma_i) = 0 \quad \text{for} \quad i > p \\
b^-_i &= \sum_{k=1}^q \beta_k b^-_{i-k} + \alpha_i (1 + \gamma_i)^\delta \quad \text{with} \quad \alpha_i (1 + \gamma_i) = 0 \quad \text{for} \quad i > p
\end{align*}
\]

(3.5)

with $b^+_i = b^-_i = 0$ for $i \leq 0$. As a consequence, for APARCH model, $f^d_\theta \equiv 0$ and $M^d_\theta = \sigma_t$. It is clear that $\alpha^{(0)}_j(f, \Theta) = 0$ and simple computations imply $\alpha^{(0)}_j(M, \Theta) = \sup_{\theta \in \Theta} \max (|b^+_j(\theta)|^{1/\delta}, |b^-_j(\theta)|^{1/\delta})$. Therefore $A_0(f, \Theta)$ holds and $\sum_{j=1}^q \beta_j < 1$ implies that a sequence defined by $u_n = \sum_{k=1}^q \beta_k u_{n-k}$ for $n$ large enough is such as $(u_n)_{n \in \mathbb{N}}$ is an exponentially decreasing sequence and therefore $A_0(M, \Theta)$ holds. Thus for $r \geq 1$, the stationarity set $\Theta(r)$ is defined by

\[
\Theta(r) = \left\{ \theta \in \mathbb{R}^{2p+q+2} \left/ \left( \mathbb{E} [ |\zeta_0|^r ] \right)^{1/r} \sum_{j=1}^\infty \max (|b^+_j|^{1/\delta}, |b^-_j|^{1/\delta}) < 1 \right. \right\}
\]

(3.6)

Now the strong consistency and asymptotic normality of the Laplacian-QMLE for APARCH models can be established (see the proof in Section 5):

**Proposition 3.1.** Assume that $X$ is a stationary solution of (3.4) with $\theta_0 \in \Theta$ where $\Theta$ is a compact subset of $\Theta(r)$ defined in (3.6). Then,

1. If $r = 1$, then $\hat{\theta}_n \overset{a.s.}{\to} \theta_0$.
2. If $r = 2$, and if $\Gamma_M$ defined in (5.21) is a definite positive symmetric matrix, then

\[
\sqrt{n}(\hat{\theta}_n - \theta_0) \overset{D}{\to} N_{2p+q+2}(0, (\sigma^2_\zeta - 1) \Gamma_M^{-1}).
\]
To our knowledge, this is the first statement the asymptotic properties of Laplacian-QMLE for APARCH processes.

2/ ARMA-GARCH processes. ARMA\((p, q)\)-GARCH\((p', q')\) processes have been introduced by Ding et al. (1993) and Ling and McAleer (2003) as the solution of the system of equations

\[
P_{0}(L) X_{t} = Q_{0}(L) \varepsilon_{t},
\]

\[
\varepsilon_{t} = \sigma_{t} \xi_{t}, \quad \text{with} \quad \sigma_{t}^{2} = c_{0} + \sum_{i=1}^{p'} c_{i} \varepsilon_{t-i}^{2} + \sum_{i=1}^{q'} d_{i} \sigma_{t-i}^{2}
\]

(3.7)

where

- \(c_{0} > 0, c_{i} \geq 0\) for \(i = 1, \ldots, p'\), \(d_{i} \geq 0\) for \(i = 1, \ldots, q'\), \(\sum_{i=1}^{q'} d_{i} < 1\) and \(c_{p'}, d_{q'}\) positive;
- \(P_{0}(x) = 1 - a_{1} x - \cdots - a_{p} x^{p}\) and \(Q_{0}(x) = 1 - b_{1} x - \cdots - b_{q} x^{q}\) are coprime polynomials with \(\sum_{i=1}^{p} |a_{i}| < 1\) and \(\sum_{i=1}^{q} |b_{i}| < 1\).

Let \(\theta = (c_{0}, c_{1}, \ldots, c_{p'}, d_{1}, \ldots, d_{q'}, a_{1}, \ldots, a_{p}, b_{1}, \ldots, b_{q})\). We are going to use Lemma 2.1. Since \((\varepsilon_{t})\) is supposed to be a GARCH\((p', q')\), then \(f_{\theta} = 0\) and \(M_{\theta} = ((1 - \sum_{j=1}^{q'} d_{j} L^{j})^{-1} (c_{0} + c_{1} \varepsilon_{t-1}^{2} + \ldots + c_{p'} \varepsilon_{t-p}^{2})^{1/2}\) and direct computations imply that the Lipschitz coefficients of \((\varepsilon_{t})\) are such as \(a_{0}^{(0)} f_{\theta} = 0\) and \(a_{0}^{(0)} M_{\theta} = |\beta_{1}|\) with \((1 + \sum_{j=1}^{\infty} \beta_{j} x^{j})(1 - \sum_{j=1}^{q'} d_{j} x^{j}) = \sum_{j=0}^{q'} c_{j} x^{j}\).

Therefore \((\alpha_{j}^{(0)} f_{\theta})_{j}\) and \((\alpha_{j}^{(0)} M_{\theta})_{j}\) are EDS (see for instance Berkes and Horváth (2004)). Thus \((A_{0}(f_{\theta}))_{j}\) and \((A_{0}(M_{\theta}))_{j}\) hold.

Considering the ARMA part and denoting \((\psi_{j})\) such as \((1 + \sum_{j=1}^{\infty} \psi_{j} x^{j}) (1 - \sum_{j=1}^{\infty} a_{j} x^{j}) = (1 - \sum_{j=1}^{\infty} b_{j} x^{j})\), then from Lemma 2.1 we deduce that:

\[
\begin{align*}
\alpha_{j}^{(0)} &= |\psi_{j}| \\
\alpha_{j}^{(0)} M_{\theta} &= \sum_{k=1}^{j} |\psi_{k} \times |\beta_{j-k}|.
\end{align*}
\]

Then we deduce that \((\alpha_{j}^{(0)} f_{\theta})_{j}\) and \((\alpha_{j}^{(0)} M_{\theta})_{j}\) are EDS, \((A_{0}(f_{\theta}))_{j}\) and \((A_{0}(M_{\theta}))_{j}\) hold, and \(X\) is a.s. a solution of \((1.1)\) for \(\theta\) included in the \(r\)-order stationarity set \(\Theta(r)\) defined by

\[
\Theta(r) = \left\{ \theta \in \mathbb{R}^{p+q+p'+q'+1} \left| \sum_{i=1}^{\infty} |\psi_{i}(\theta)| + (E[|\xi_{0}|^{r}])^{1/2} \sum_{j=1}^{\infty} \sum_{k=1}^{j} |\psi_{k} \times |\beta_{j-k}| < 1 \right. \right\}.
\]

(3.8)

Now the strong consistency and asymptotic normality of the Laplacian-QMLE for ARMA-GARCH processes can be established:

**Proposition 3.2.** Assume that \(X\) is a stationary solution of \((3.7)\) with \(\theta_{0} \in \Theta\) where \(\Theta\) is a compact subset of \(\Theta(r)\) defined in \((3.8)\). Then,

1. If \(r = 1\), then \(\hat{\theta}_{n} \xrightarrow{a.s.} \theta_{0}\).
2. If \(r = 2\), and if \(\Gamma_{f}\) and \(\Gamma_{M}\) defined in \((5.21)\) are definite positive symmetric matrix, then the asymptotic normality \((3.3)\) of \(\hat{\theta}_{n}\) holds.

This result is a new one and extends the previous results already obtained with Gaussian-QMLE for such processes (see for instance, Ling and McAleer (2003) and Bardet and Wintenberger (2009)).

3/ ARMA-ARCH\((\infty)\) processes. ARMA\((p, q)\)-ARCH\((\infty)\) processes are a natural extension of ARMA-GARCH processes. They are the solution of the system of equations

\[
P_{0}(L) X_{t} = Q_{0}(L) \varepsilon_{t},
\]

\[
\varepsilon_{t} = \sigma_{t} \xi_{t}, \quad \text{with} \quad \sigma_{t}^{2} = c_{0} + \sum_{i=1}^{\infty} c_{i} \varepsilon_{t-i}^{2}
\]

(3.9)

where

- \(c_{0} > 0, c_{i} \geq 0\) for \(i \geq 1\);
• $P_b(x) = 1 - a_1 x - \cdots - a_p x^p$ and $Q_b(x) = 1 - b_1 x - \cdots - b_q x^q$ are coprime polynomials with $\sum_{i=1}^p |a_i| < 1$ and $\sum_{i=1}^q |b_i| < 1$.

ARCH($\infty$) processes were introduced by Robinson (1991) and the asymptotic properties of Gaussian-QMLE were studied in Robinson and Zaffaroni (2006), Straumann and Mikosh (2006) or Bardet and Wintenberger (2009). Hence, we assume that there exists $\beta = (\beta_1, \ldots, \beta_m)$ such as for all $i \in \mathbb{N}$, $c_i = c(i, \beta)$, with $c(\cdot)$ a known function. Let $\theta = (\beta_1, \ldots, a_p, b_1, \ldots, b_q)$. We are going to use Lemma 2.1. Since $(\varepsilon_t)$ is supposed to be an ARCH($\infty$), then $f_\theta^\infty = 0$ and $M_\theta^\infty = (\varepsilon_0, \beta + \sum_{i=1}^\infty c(i, \beta) \varepsilon_i^2)^{1/2}$ and direct computations imply that the Lipschitz coefficients of $(\varepsilon_t)$ are such as $\alpha_j^{(0)}(f, \{\theta_0\}) = 0$ and $\alpha_j^{(0)}(M, \{\theta_0\}) = (j, \beta_0)$. Therefore we assume that there exists $\ell > 1$ such as

$$c(j, \beta_0) = O(j^{-\ell}) \text{ when } j \to \infty.$$  \hfill (3.10)

Thus $(A_0(f, \{\theta_0\}))$ and $(A_0(M, \{\theta_0\}))$ hold.

Then we deduce that $(\alpha_j^{(0)}(f, \{\theta_0\}))_j$ is EDS and $(\alpha_j^{(0)}(M, \{\theta_0\}))_j = O(j^{-\ell})$. Then $(A_0(f, \{\theta_0\}))$ and $(A_0(M, \{\theta_0\}))$ hold, and $\Theta$ is a.s. a solution of (1.1) for $\theta$ included in the $r$-order stationarity set $\Theta(r)$ defined by

$$\Theta(r) = \left\{ \theta \in \mathbb{R}^{p+q+m} / \sum_{i=1}^\infty |\psi_i(\theta)| + (E |\zeta_0|^r)^{1/r} \sum_{j=1}^\infty \sum_{k=1}^\infty |\psi_k| \times c(j, \beta_0) < 1 \right\}.$$  \hfill (3.11)

Now the strong consistency and asymptotic normality of the Laplacian-QMLE for ARMA-ARCH($\infty$) processes can be established:

**Proposition 3.3.** Assume that $X$ is a stationary solution of (3.9) where (3.10) holds and with $\theta_0 \in \Theta$ where $\Theta$ is a compact subset of $\Theta(r)$ defined in (3.11). Then,

1. If $r \geq 1$ and $\ell > 2 / \min(r,2)$, then $\hat{\theta}_n \stackrel{a.s.}{\longrightarrow}_{n \to \infty} \theta_0$.
2. If $r = 2$, $\ell > 1$ and if $\partial_j^2 c(j, \beta) = O(j^{-\ell})$ for $i = 1, 2$, and if $\Gamma_f$ and $\Gamma_M$ defined in (5.21) are definite positive symmetric matrix, then the asymptotic normality (3.3) of $\hat{\theta}_n$ holds.

This result is a new one. Note that $\ell > 1$ and $r = 2$ is required for the asymptotic normality of Laplacian-QMLE while $r = 4$ and $\ell > 2$ is required for Gaussian-QMLE for such processes (see for instance Bardet and Wintenberger (2009)). This confers a clear advantage to Laplacian-QMLE.

**4/ ARMA-APARCH processes.** The ARMA($p,q$)-APARCH($p', q'$) processes have been also introduced by Ding et al. (1993) as the solutions of the equations

$$\begin{cases}
P_b(L) X_t = Q_b(L) \varepsilon_t, \\
\varepsilon_t = \sigma_t \zeta_t, \text{ with } \sigma_t^\delta = \omega + \sum_{i=1}^{p'} \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^{q'} \beta_j \sigma_{t-j}^\delta
\end{cases}$$ \hfill (3.12)

where:

- $\delta \geq 1$, $\omega > 0$, $-1 < \gamma_i < 1$ and $\alpha_i \geq 0$ for $i = 1, \ldots, p' - 1$, $\beta_j \geq 0$ for $j = 1, \ldots, q' - 1$, $\alpha_{p'}, \beta_{q'}$ positive real numbers and $\sum_{i=1}^{p'} \alpha_i \leq 1$;
- $P_b(x) = 1 - a_1 x - \cdots - a_p x^p$ and $Q_b(x) = 1 - b_1 x - \cdots - b_q x^q$ are coprime polynomials with $\sum_{i=1}^p |a_i| < 1$ and $\sum_{i=1}^q |b_i| < 1$.

Let $\theta = (\delta, \omega, \alpha_1, \ldots, \alpha_{p'}, \gamma_1, \ldots, \gamma_{p'}, \beta_1, \ldots, \beta_{q'}, a_1, \ldots, a_p, b_1, \ldots, b_q)$. Then, as for ARMA-GARCH processes, we are going to use Lemma 2.1. Thanks to the computations realized for APARCH processes, we obtain $\alpha_j^{(0)}(f, \{\theta_0\}) = 0$.
and \( \alpha_j^{(0)}(M^\varepsilon, \{\theta_0\}) = \max(|b_1^+|^{1/\delta}, |b_1^-|^{1/\delta}) \) with \((b_1^+, b_1^-)_{i \geq 1}\) defined in (3.5).

Then, we have
\[
\begin{align*}
\alpha_j^{(0)}(f, \{\theta_0\}) & \leq |\psi_j| \sum_{k=1}^{j} |\psi_k| \times \max \left( |b_{j-k}^+|^{1/\delta}, |b_{j-k}^-|^{1/\delta} \right), \\
\alpha_j^{(0)}(M, \{\theta_0\}) & \leq |\psi_j| \sum_{k=1}^{j} |\psi_k| \times \max \left( |b_{j-k}^+|^{1/\delta}, |b_{j-k}^-|^{1/\delta} \right).
\end{align*}
\]

\( (\psi_j) \) such as \((1 + \sum_{j=1}^{\infty} \psi_j x^j) \left(1 - \sum_{j=1}^{\infty} a_j x^j \right) = (1 - \sum_{j=1}^{\infty} b_j x^j) \). From Lemma 2.1, \((A_0(f, \Theta)) \) and \((A_0(M, \Theta)) \) hold since \((\alpha_j^{(0)}(f^\varepsilon; \{\theta_0\}))_j = 0 \) and \((\alpha_j^{(0)}(M^\varepsilon, \{\theta_0\}))_j \) are EDS. As a consequence, for \( r \geq 1 \), the stationarity set \( \Theta(r) \) is defined by
\[
\Theta(r) = \left\{ \theta \in \mathbb{R}^{d+q+f'+q'+2} : \sqrt{\frac{\sum_{j=1}^{\infty} |\psi_j| + (E|\zeta_0|^r)}{r}} \sum_{j=1}^{\infty} \sum_{k=1}^{j} |\psi_k| \times \max \left( |b_{j-k}^+|^{1/\delta}, |b_{j-k}^-|^{1/\delta} \right) < 1 \right\}.
\]

Now, we are able to provide the asymptotic properties of QMLE for ARMA-APARCH models.

**Proposition 3.4.** Assume that \( X \) is a stationary solution of (3.12) with \( \theta_0 \in \Theta \) where \( \Theta \) is a compact subset of \( \Theta(r) \) defined in (3.8). Then,

1. If \( r = 1 \), then \( \hat{\theta}_n \xrightarrow{a.s.} \theta_0 \).
2. If \( r = 2 \), and if \( \Gamma_f \) and \( \Gamma_M \) defined in (5.21) are definite positive symmetric matrix, then the asymptotic normality (3.3) of \( \theta_n \) holds.

This result is stated for the first time for Laplacian-QMLE. The case of Gaussian-QMLE for ARMA-APARCH could be also obtained following the previous decomposition and the paper Bardet and Wintenberger (2009). Once again, the asymptotic normality of Laplacian-QMLE only requires \( r = 2 \) while this requires \( r = 4 \) for Gaussian-QMLE.

**4. Numerical Results**

To illustrate the asymptotic results stated previously, we realized Monte-Carlo experiments on the beavior of Laplacian-QMLE (denoted \( \widehat{\theta}_n^{QL} \)) for several time series models, sample sizes and probability distributions. A comparison with the results obtained by Gaussian QMLE (denoted \( \widehat{\theta}_n^{GQL} \)) is also proposed.

More precisely, the considered probability distributions of \((\zeta_t)\) are:

- Centred Gaussian distribution denoted \( \mathcal{N} \);
- Centred Laplacian distribution denoted \( \mathcal{L} \);
- Centred Uniform distribution denoted \( \mathcal{U} \);
- Centred Student distribution with 3 freedom degrees, denoted \( t_3 \);
- Normalized centred Gaussian mixture with probability distribution \( 0.05 \times \mathcal{N}(-2, 0.16) + \mathcal{N}(0, 1) + 0.05 \times \mathcal{N}(2, 0.16) \) and denoted \( \mathcal{M} \).

All these probability distributions are normalized such as \( E[|\zeta_0|] = 1 \), required for Laplacian-QMLE. For using Gaussian-QMLE requiring \( \sigma_\zeta^2 = 1 \), it is necessary to consider the model with \( M^\gamma = \frac{E[|\zeta_0|]}{\sigma_\zeta} M_\theta \) instead of \( M_\theta \).

Several models of time series satisfying (1.1) and the assumptions of Theorem 3.1 and 3.3 are considered:

- a ARMA(1, 1) process defined by \( X_t = \phi X_{t-1} + \zeta_t + \theta \zeta_{t-1} \) with \( \phi = 0.4 \) and \( \theta = 0.6 \);
- a ARCH(1) process defined by \( X_t = \zeta_t \sqrt{\omega + \alpha X_{t-1}^2} \) with \( \omega = 0.4 \) and \( \alpha = 0.2 \);
- a GARCH(1, 1) process defined by \( X_t = \zeta_t \sigma_t \) where \( \sigma_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta \sigma_{t-1}^2 \) with \( \alpha_0 = 0.2 \), \( \alpha_1 = 0.4 \) and \( \beta = 0.2 \);
- a ARMA(1,1)-GARCH(1,1) process defined by \( X_t = \phi X_{t-1} + \varepsilon_t + \varepsilon_{t-1} \) where \( \varepsilon_t = \zeta_t \sigma_t \) and \( \sigma_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \) with \( \phi = 0.4 \), \( \theta = 0.6 \), \( \alpha_0 = 0.2 \), \( \alpha_1 = 0.4 \) and \( \beta = 0.1 \);
- a ARMA(1,1)-APARCH(1,1) process defined by \( X_t = \phi X_{t-1} + \varepsilon_t + \varepsilon_{t-1} \) where \( \varepsilon_t = \zeta_t \sigma_t \) and \( \sigma_t = \alpha_0 + \alpha_1 (|\varepsilon_{t-1}| - \gamma \varepsilon_{t-1}) \) with \( \phi = 0.4 \), \( \theta = 0.6 \), \( \alpha_0 = 0.2 \), \( \alpha_1 = 0.4 \), \( \gamma = 0.5 \), \( \beta = 0.1 \) and \( \delta = 1.2 \).
Hence we computed the root-mean-square error (RMSE) from 1000 independent replications of $\hat{\theta}_n^{LQL}$ and $\bar{\theta}_n^{LQL}$ for those processes and the results are presented in Table 1 on page 10 and 2 on page 11.

**Conclusion of the numerical results:** On the one hand, it is clear that the RMSE decreases as the sample size increases, which validates the theoretical results (consistency of the estimators). On the other hand, Table 1 and 2 show that the Laplacian-QMLE provides more accurate estimation than the Gaussian-QMLE for several types of noise, except of course in the case of a Gaussian distribution (even in this case the RSME of both the estimators are almost the same).

5. Proofs

**Proof of Lemma 2.1.** First, as $X$ is a stationary process and the ARMA($p,q$) process is causal invertible then $\bar{X}$ is also a stationary process (the coefficients of $\Lambda_{\beta_0}$ are EDS). Moreover, it is well known that $(\psi_j(\beta_0))_{j\in\mathbb{N}}$ is EDS. Then we have:

$$\bar{X}_t = \Lambda_{\beta_0}(L) \left( M_{\theta_0}((X_{t-i})_{i\geq 1}) \zeta_t + f_{\theta_0}((X_{t-i})_{i\geq 1}) \right)$$

$$\bar{X}_t + \sum_{j=1}^{\infty} \psi_j(\beta_0) \bar{X}_{t-j} = M_{\theta_0}((\Lambda^{-1}_{\beta_0}(L)\bar{X}_{t-i})_{i\geq 1}) \zeta_t + f_{\theta_0}((\Lambda^{-1}_{\beta_0}(L)\bar{X}_{t-i})_{i\geq 1})$$

$$\bar{X}_t = \tilde{M}_{\theta_0}((\bar{X}_{t-i})_{i\geq 1}) \zeta_t + \tilde{f}_{\theta_0}((\bar{X}_{t-i})_{i\geq 1})$$

with

$$\tilde{M}_{\theta_0}((x_{t-i})_{i\geq 1}) = M_{\theta_0}((\Lambda^{-1}_{\beta_0}(L)x_{t-i})_{i\geq 1})$$

$$\tilde{f}_{\theta_0}((x_{t-i})_{i\geq 1}) = f_{\theta_0}((\Lambda^{-1}_{\beta_0}(L)x_{t-i})_{i\geq 1}) - \sum_{j=1}^{\infty} \psi_j(\beta_0)x_{t-j}.$$  

(5.1)
Table 2

Root Mean Square Error of the components of $\hat{\theta}_n$ and $\hat{\theta}_n$ for ARMA(1,1)-GARCH(1,1) and ARMA(1,1)-APARCH(1,1) processes.

<table>
<thead>
<tr>
<th>n</th>
<th>$\theta$</th>
<th>$\phi$</th>
<th>$\alpha_0$</th>
<th>$\alpha_1$</th>
<th>$\beta$</th>
<th>$\delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.120</td>
<td>0.150</td>
<td>0.156</td>
<td>0.169</td>
<td>0.302</td>
<td>0.732</td>
</tr>
<tr>
<td>1000</td>
<td>0.035</td>
<td>0.045</td>
<td>0.040</td>
<td>0.048</td>
<td>0.072</td>
<td>0.352</td>
</tr>
<tr>
<td>5000</td>
<td>0.016</td>
<td>0.022</td>
<td>0.018</td>
<td>0.023</td>
<td>0.017</td>
<td>0.170</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>n</th>
<th>$\gamma$</th>
<th>$\beta$</th>
<th>$\delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.413</td>
<td>0.287</td>
<td>0.732</td>
</tr>
<tr>
<td>1000</td>
<td>0.105</td>
<td>0.074</td>
<td>0.402</td>
</tr>
<tr>
<td>5000</td>
<td>0.042</td>
<td>0.033</td>
<td>0.170</td>
</tr>
</tbody>
</table>
Finally, for $i = 0$,
\[ |\tilde{f}_{\tilde{\theta}_0}(\{x_{t-i}\}_{i \geq 1}) - \tilde{f}_{\tilde{\theta}_0}(\{y_{t-i}\}_{i \geq 1})| \leq \sum_{j=1}^{\infty} \alpha_{j}^{(0)}(f, \{\theta_0\}) \left| (\Lambda_{\beta_0}^{-1}(L)x_{t-j-i})_{i \geq 1} - (\Lambda_{\beta_0}^{-1}(L)y_{t-j-i})_{i \geq 1} \right| + |\psi_j(\beta_0)| |x_{t-j} - y_{t-j}| \]
\[ \leq \sum_{j=1}^{\infty} \alpha_{j}^{(0)}(f, \{\theta_0\}) \sum_{k=0}^{j} |\psi_k(\beta_0)| |x_{t-k-j} - y_{t-k-j}| + |\psi_j(\beta_0)| |x_{t-j} - y_{t-j}| \]
\[ \leq \sum_{j=1}^{\infty} \left( |\psi_j(\beta_0)| + \sum_{k=1}^{j} \alpha_{k}^{(0)}(f, \{\theta_0\}) |\psi_{j-k}(\beta_0)| \right) |x_{t-j} - y_{t-j}| \]
\[ \Rightarrow \alpha_{j}^{(0)}(\tilde{f}, \{\tilde{\theta}_0\}) \leq |\psi_j(\beta_0)| + \sum_{k=1}^{j} \alpha_{k}^{(0)}(f, \{\theta_0\}) |\psi_{j-k}(\beta_0)|. \tag{5.2} \]

Moreover, we also have:
\[ |\tilde{M}_{\tilde{\theta}_0}(\{x_{t-i}\}_{i \geq 1}) - \tilde{M}_{\tilde{\theta}_0}(\{y_{t-i}\}_{i \geq 1})| \leq \sum_{j=1}^{\infty} \alpha_{j}^{(0)}(M, \{\theta_0\}) \left| (\Lambda_{\beta_0}^{-1}(L)x_{t-j-i})_{i \geq 1} - (\Lambda_{\beta_0}^{-1}(L)y_{t-j-i})_{i \geq 1} \right| \]
\[ \Rightarrow \alpha_{j}^{(0)}(\tilde{M}, \{\tilde{\theta}_0\}) \leq \sum_{k=1}^{j} \alpha_{k}^{(0)}(M, \{\theta_0\}) |\psi_{j-k}|. \tag{5.3} \]

The same kinds of computations could also be done by considering the first and second derivatives of $\tilde{f}$ and $\tilde{M}$ with respect to $\tilde{\theta}$. Note, and this is important, that the first and second derivatives of $\Lambda_{\beta_0}^{-1}$ with respect to $\tilde{\theta}$ are also EDS. Finally,

- if when $j \to \infty$, $\alpha_{j}^{(0)}(K, \{\theta_0\}) = O(j^{-\beta})$ with $\beta > 1$ and $\psi_j = O(\rho^j)$ with $0 \leq \rho < 1$, then there exists $C > 0$ such as $\sum_{k=1}^{j} \alpha_{k}^{(0)}(K, \{\theta_0\}) |\psi_{j-k}| \leq C \sum_{k=1}^{j} k^{-\beta} \rho^{j-k} \sim -C(\log \rho)^{-1} j^{-\beta}$ and therefore $\alpha_{j}^{(0)}(\tilde{K}, \{\tilde{\theta}_0\}) = O(j^{-\beta})$.

- if when $j \to \infty$, $\alpha_{j}^{(0)}(K, \{\theta_0\}) = O(r^j)$ with $0 \leq r < 1$ and $\psi_j = O(\rho^j)$ with $0 \leq \rho < 1$, then there exists $C > 0$ such as $\sum_{k=1}^{j} \alpha_{k}^{(0)}(K, \{\theta_0\}) |\psi_{j-k}| \leq C \sum_{k=1}^{j} r^{-k} \rho^{j-k} = O(j \max(r, \rho)^j)$ and therefore $\alpha_{j}^{(0)}(\tilde{K}, \{\tilde{\theta}_0\})$ is EDS.

The same kind of computation can be also done for $(\alpha_{j}^{(0)}(\tilde{K}, \{\tilde{\theta}_0\}))_j$ since the derivatives and second-derivatives of $\Lambda_{\beta_0}^{-1}$ with respect to $\beta$ and therefore to $\tilde{\theta}$ are also EDS. \hfill \square

Now we remind two lemmas already proved in Bardet and Wintenberger (2009):

**Lemma 5.1.** Assume that $\theta_0 \in \Theta(r)$ for $r \geq 1$ and $X$ is the causal stationary solution of the equation (1.1). If $(A_0(K, \Theta))$ holds (with $K = f$ or $K = M$) then $K_{\theta}^t \in L^r(C(\Theta, R^m))$ and there exists $C > 0$ not depending on $t$ such that
\[ E \left[ \|K_{\theta}^t - K_{\theta_0}^t\|_6^2 \right] \leq C E \left[ \|X_0\|^r \right] \left( \sum_{j \geq t} \alpha_j(K, \Theta) \right)^r \tag{5.4} \]

**Lemma 5.2.** Let $D^{(2)}(\Theta)$ denote the Banach space of $2$ times continuously differentiable functions on $\Theta$ equipped with the uniform norm
\[ \|h\|_{D^{(2)}(\Theta)} = \|h\|_\Theta + \left\| \frac{\partial h}{\partial \Theta} \right\|_\Theta + \left\| \frac{\partial^2 h}{\partial \Theta \partial \Theta} \right\|_\Theta. \]
Let $\theta_0 \in \Theta(r)$ ($r \geq 1$) and assume that for $i = 0, 1, 2, (A_i(f, \Theta))$ and $(A_i(M, \Theta))$ hold. Then $f_{\theta_0}^i \in L^r(D^{(2)}(\Theta))$ and $M_{\theta_0}^i \in L^r(D^{(2)}(\Theta))$.

Now, we begin with the proofs of Theorem 3.1, 3.2 and 3.3.

**Proof of Theorem 3.1.** The proof of the theorem is divided into two parts and follows the same kind of procedure than in Jeantheau (1998). In (i), a uniform (on $\Theta$) strong law of large numbers satisfied by $\frac{1}{n} \tilde{L}_n(\theta)$ converging to $L(\theta) := -E[q_0(\theta)]$ is established. In (ii), it is proved that $L(\theta)$ admits a unique maximum in $\theta_0$. Those two conditions lead to the strong consistency of $\hat{\theta}_n$ (from Jeantheau (1998)).
(i) In the same way and for the same reason in the proof of Theorem 1 of Bardet and Wintenberger (2009), the uniform strong law of large numbers satisfied by the sample mean of \( (\tilde{q}_t)_{t\in N^*} \) (defined in (2.3)) is implied by establishing \( \text{E}[(q_t(\theta))_\Theta] < \infty \). But new computations have to be done in case of Laplacian conditional log-density \( q_t(\theta) \). From Lemma 5.1, for all \( t \in \mathbb{Z} \),

\[
|q_t(\theta)| = |(M_0^{-1})X_t - f_0^t| + \log(M_0^t) \leq \frac{|X_t - f^t(\theta)|}{M} + |\log(M)| + M_0^t
\]

\[
\Rightarrow \sup_{\theta \in \Theta} |q_t(\theta)| \leq \frac{1}{M}(|X_t| + \|f^t(\theta)\|_\Theta) + |\log(M)| + M_0^t.
\]

With \( r \geq 1 \), we have \( \forall t \in \mathbb{Z}, \text{E}[|X_t|] < \infty \) from Proposition 2.1 and \( \text{E}[\|f^t_\Theta\|_\Theta + \|M^t_\Theta\|_\Theta] < \infty \) from Lemma 5.1, implying \( \text{E}[\|f^t_\Theta\|_\Theta + \|M^t_\Theta\|_\Theta] < \infty \). As a consequence, for all \( t \in \mathbb{Z} \),

\[
\text{E}[|q_t(\theta)|_\Theta] < \infty.
\]

Hence, the uniform strong law of large numbers for \( (q_t(\theta)) \) follows:

\[
\left\| \frac{L_n(\theta)}{n} - L(\theta) \right\|_{\Theta} \xrightarrow{a.s.} 0. \tag{5.5}
\]

Now, we are going to establish \( \frac{1}{n} \left\| \tilde{L}_n(\theta) - L_n(\theta) \right\|_{\Theta} \xrightarrow{a.s.} 0 \). Indeed, for all \( \theta \in \Theta \) and \( t \in N^* \),

\[
|\tilde{q}_t(\theta) - q_t(\theta)| \leq |\log(M_0^t) - \log(M_0^t + (M_0^t)^{-1}|X_t - f_0^t|) - (M_0^t)^{-1}|X_t - f_0^t||
\]

\[
\leq \frac{1}{M}(M_0^t - M_0^t|X_t - f_0^t| + |\tilde{M}_0^t - M_0^t| + \tilde{M}_0^t - M_0^t|X_t - f_0^t| + M_0^t - f_0^t)
\]

with \( C > 0 \). Hence, we have:

\[
\|\tilde{q}_t(\theta) - q_t(\theta)\|_\Theta \leq C(1 + |X_t| + \|f_0^t\|_\Theta)(|\tilde{M}_0^t - M_0^t| + \|\tilde{f}_0^t - f_0^t\|_\Theta).
\]

By Corollary 1 of Kounias and Weng (1969), the proof is achieved if there exists \( s \in (0, 1) \) such as

\[
\sum_{t \geq 1} \frac{1}{t^s} \text{E}[|q_t(\theta) - \tilde{q}_t(\theta)|^r_\Theta] < \infty. \tag{5.6}
\]

Let us prove (5.6) with \( s = r/2 \) when \( r \in [1, 2] \).

From Cauchy-Schwarz Inequality and assumptions \( A_0(f, \Theta) \) and \( A_0(M, \Theta) \),

\[
\text{E}[|\tilde{q}_t(\theta) - q_t(\theta)|_\Theta^{r/2}] \leq C(1 + |X_t| + \|f_0^t\|_\Theta^{r/2}) \text{E}[|\tilde{M}_0^t - M_0^t| + \|\tilde{f}_0^t - f_0^t\|_\Theta^{r/2}].
\]

Using Lemma 5.1 and previous proved results implying \( \text{E}[|X_t|] < \infty \), \( \text{E}[\|f_0^t\|_\Theta + \|M_0^t\|_\Theta] < \infty \), we obtain

\[
\text{E}[|\tilde{q}_t(\theta) - q_t(\theta)|_\Theta^{r/2}] \leq C \left( \sum_{j > 1} \alpha_j^{(0)}(f, \Theta) + \alpha_j^{(0)}(M, \Theta) \right)^{\frac{r}{2}}
\]

\[
\leq C t^{-\frac{r-1}{2}},
\]

where the last inequality is obtained from the condition (3.1) of Theorem 3.1. Hence, we have

\[
\sum_{t \geq 1} \frac{1}{t^{r/2}} \text{E}[|\tilde{q}_t(\theta) - q_t(\theta)|_\Theta^{r/2}] \leq A \sum_{t \geq 1} t^{-r/2},
\]
which is finite when \( r \ell > 2 \). When \( r \geq 2 \), it is sufficient to consider the case \( r = 2 \). As a consequence, we obtain

\[
\frac{1}{n} \sum_{t=1}^{n} \| \hat{q}_t(\theta) - q_t(\theta) \|_{\Theta} \xrightarrow{a.s.} 0 \quad \text{and} \quad \frac{1}{n} \| \hat{L}_n(\theta) - L_n(\theta) \|_{\Theta} \xrightarrow{a.s.} 0,
\]

and therefore, using (5.5),

\[
\frac{1}{n} \| \hat{L}_n(\theta) - L(\theta) \|_{\Theta} \xrightarrow{a.s.} 0.
\]

(ii) Now for \( \theta \in \Theta \), we study

\[
L(\theta) = -E[q_0(\theta)].
\]

which can also be consider as a Kullback-Leibler discripency. We have

\[
L(\theta) = -E\left[ \log(M_0^\theta) + (M_0^\theta)^{-1} |X_t - f_0^\theta| \right] = -E\left[ \log(M_0^\theta) + \frac{M_{0,0}^\theta}{M_0^\theta} \left( \zeta_t + \frac{f_0^\theta - f_0^\theta}{M_{0,0}^\theta} \right) \right].
\]

Hence, using \( E[|\zeta_t|] = 1 \), we obtain:

\[
L(\theta_0) - L(\theta) = E\left[ \log\left( \frac{M_0^\theta}{M_0^\theta} \right) + \frac{M_{0,0}^\theta}{M_0^\theta} \left( \zeta_t + \frac{f_0^\theta - f_0^\theta}{M_{0,0}^\theta} \right) - 1 \right] = E\left[ \log\left( \frac{M_0^\theta}{M_0^\theta} \right) - 1 + \frac{M_{0,0}^\theta}{M_0^\theta} E\left[ |\zeta_t + \frac{f_0^\theta - f_0^\theta}{M_{0,0}^\theta} | \right] (X_{t-k})_{k \geq 1} \right].
\]

But for \( \zeta_t \) following a symmetric probability distribution, for any \( m \in \mathbb{R}^* \), \( E[|\zeta_t + m|] > E[|\zeta_t|] = 1 \). Therefore, for \( \theta \neq \theta_0 \), if \( f_0 \neq f_0^\theta \) (else \( > \) is replaced by \( \geq \)),

\[
L(\theta_0) - L(\theta) > E\left[ \log\left( \frac{M_0^\theta}{M_0^\theta} \right) - 1 + \frac{M_{0,0}^\theta}{M_0^\theta} \right] > h\left( \frac{M_{0,0}^\theta}{M_0^\theta} \right)
\]

with \( h(x) = -\log(x) - 1 + x \). But for any \( x \in (0, 1) \cup (1, \infty) \), \( h(x) > 0 \) and \( h(1) = 0 \). Therefore if \( M_0 \neq M_{0,0} \), \( h\left( \frac{M_{0,0}^\theta}{M_0^\theta} \right) > 0 \) (> 0 is replaced by = 0 if \( M_0 = M_{0,0} \)). This implies from Condition C3 (Identifiability) that \( L(\theta_0) - L(\theta) > 0 \) almost surely for all \( \theta \in \Theta, \theta \neq \theta_0 \). Hence a supremum of \( L(\theta) \) is only reached for \( \theta = \theta_0 \) which is the unique maximum. \( \square \)

**Proof of Theorem 3.2.** We follow the same scheme of proof than in Davis and Dunsmuir (1997). Hence, denote

\[
S_n = \sum_{i=1}^{n} V_{t-1} (|Z_t - n^{-1/2} Y_{t-1}| - |Z_t|)
\]

\[
= -n^{-1/2} \sum_{i=1}^{n} V_{t-1} Y_{t-1} \text{sgn}(Z_t) + 2 \sum_{i=1}^{n} V_{t-1} (n^{-1/2} Y_{t-1} - Z_t)(1_{0 < Z_t < n^{-1/2} Y_{t-1}} - 1_{n^{-1/2} Y_{t-1} < Z_t < 0})
\]

\[
= A_n + B_n.
\]

Since \( E[V_{t-1} Y_{t-1} \text{sgn}(Z_t) | \mathcal{F}_{t-1}] = E[\text{sgn}(Z_t)] E[V_{t-1} Y_{t-1}] = 0 \) and \( E[V_0^2 Y_0^2] < \infty \), we can apply a central limit theorem for stationary martingale difference sequence (see Billingsley (1968)) and

\[
A_n \xrightarrow{D} \mathcal{N}(0, E[V_0^2 Y_0^2]).
\]

(5.9)
Now, considering $B_n$, define also $W_{nt} = V_{t-1}(n^{-1/2}Y_{t-1} - Z_t)\mathbf{1}_{0 < Z_t < n^{-1/2}Y_{t-1}}$. Using the same arguments as in Davis and Dunsmuir (1997), we also obtain

- $\limsup_{n \to \infty} n E[W_{nt}^2] = 0$;
- $E[W_{nt} | \mathcal{F}_{t-1}] \simeq \frac{1}{2n} f(0) V_{t-1} Y_{t-1}^2$ for $|n^{-1/2}Y_{t-1}| < \varepsilon$;
- $\sum_{t=1}^{n} W_{nt} \overset{P}{\to}_{n \to \infty} \frac{1}{2} f(0) E[V_0 Y_0^2 \mathbf{1}_{Y_0 > 0}]$.

Then we deduce

$$B_n \overset{P}{\to}_{n \to \infty} f(0) E[V_0 Y_0^2].$$

(5.10)

The proof is achieved from (5.9) and (5.10).

**Proof of Theorem 3.3.** We follow a proof which is similar to the one of Theorem 2 in Davis and Dunsmuir (1997) or Li and Li (2008).

Let $v = \sqrt{n}(\theta - \theta_0) \in \mathbb{R}^d$. Then we are going to prove in 2/ that maximizing $\hat{L}_n(\theta)$ is equivalent to maximizing $L_n(\theta)$ which is equivalent to maximizing

$$W_n(v) = -\sum_{t=1}^{n} (q_t(\theta_0 + n^{-1/2}v) - q_t(\theta_0)) \tag{5.11}$$

$$= \sum_{t=1}^{n} \log \left( \frac{(M_{\theta_0}^t)^{-1}}{(M_{\theta_0}^t)^{-1}} \right) + (M_{\theta_0}^t)^{-1} |X_t - f_{\theta_0}^t| - (M_{\theta_0}^t)^{-1} |X_t - f_{\theta_0 + n^{-1/2}v}^t|$$

with respect to $v$. As a consequence, there exists a sequence $(\hat{v}_n)_n$ where $\hat{v}_n$ is a maximizer of $W_n(v)$ such as $\hat{v}_n = \sqrt{n}(\hat{\theta}_n - \theta_0)$. In 1/ we will provide a limit theorem satisfied by $W_n(v)$. Then we are going to prove in 3/ that $(W_n(\cdot))_n$ converges as a process of $C(\mathbb{R}^d)$ (space of continuous functions on $\mathbb{R}^d$) to a limit process $W$. Hence $(\hat{v}_n)_n$ converges to the maximizer of $W$.

1/ First, we are going to study the asymptotic behavior of $W_n(v)$. We have

$$W_n(v) = \sum_{t=1}^{n} \log \left( \frac{(M_{\theta_0}^t)^{-1}}{(M_{\theta_0}^t)^{-1}} \right) + |X_t - f_{\theta_0}^t| - (M_{\theta_0}^t)^{-1} |X_t - f_{\theta_0 + n^{-1/2}v}^t|$$

$$= \mathcal{I}_1(v) + \mathcal{I}_2(v).$$

We have:

$$\mathcal{I}_1(v) = -\sum_{t=1}^{n} \log \left( \frac{(M_{\theta_0}^t)^{-1}}{(M_{\theta_0}^t)^{-1}} \right) + |\xi_t| \left( 1 - \frac{M_{\theta_0}^t}{M_{\theta_0 + n^{-1/2}v}^t} \right)$$

Using Taylor expansions, we deduce that for each $t \in \{1, \cdots, n\}$, there exists $\overline{\theta}_1^t$ and $\overline{\theta}_2^t$ in the segment $[\theta_0, \theta_0 + n^{-1/2}v]$ such as:

$$\log \left( \frac{(M_{\theta_0}^t)^{-1}}{(M_{\theta_0}^t)^{-1}} \right) = n^{-1/2}(M_{\theta_0}^t)^{-1} v' \left( \frac{\partial M_{\theta_0}^t}{\partial \theta} \right)_{\theta_0} + \frac{1}{2} n^{-1} \left\{ v' \left( \frac{\partial^2 M_{\theta_0}^t}{\partial \theta^2} \right)_{\theta_0} M_{\theta_0}^t - (M_{\theta_0}^t)^{-2} \left( v' \left( \frac{\partial M_{\theta_0}^t}{\partial \theta} \right)_{\theta_0} M_{\theta_0}^t - v' \left( \frac{\partial^2 M_{\theta_0}^t}{\partial \theta^2} \right)_{\theta_0} M_{\theta_0}^t \right) \right\}$$

$$\frac{M_{\theta_0}^t}{(M_{\theta_0 + n^{-1/2}v}^t)^{-1}} = 1 - n^{-1/2}(M_{\theta_0}^t)^{-1} v' \left( \frac{\partial M_{\theta_0}^t}{\partial \theta} \right)_{\theta_0} + \frac{1}{2} n^{-1} \left\{ 2(M_{\theta_0}^t)^{-2} \left( v' \left( \frac{\partial M_{\theta_0}^t}{\partial \theta} \right)_{\theta_0} M_{\theta_0}^t - v' \left( \frac{\partial^2 M_{\theta_0}^t}{\partial \theta^2} \right)_{\theta_0} M_{\theta_0}^t \right) \right\}$$

$$+ \frac{1}{2} n^{-1} \left\{ 2(M_{\theta_0}^t)^{-2} \left( v' \left( \frac{\partial M_{\theta_0}^t}{\partial \theta} \right)_{\theta_0} M_{\theta_0}^t - v' \left( \frac{\partial^2 M_{\theta_0}^t}{\partial \theta^2} \right)_{\theta_0} M_{\theta_0}^t \right) \right\}$$
Using a Central Limit Theorem for martingale-differences (see for instance Billingsley (1968)), and since from Lemma 5.2, 
\[ \text{E} \left[ \left( M_{0}^{t} \right)^{-1} v' \left( \frac{\partial M_{0}^{t}}{\partial \theta} \right)_{\theta_{0}} \right] \leq \infty \text{ and } \text{E} \left[ \left( M_{0}^{t} \right)^{-1} v' \left( \frac{\partial M_{0}^{t}}{\partial \theta} \right)_{\theta_{0}} \right] \left( |\zeta_{t}| - 1 \right) = 0, \] 
we have:

\[ I_{1}^{(1)}(v) \xrightarrow{D} N(0, \text{E} \left[ \left( M_{0}^{t} \right)^{-2} v' \left( \frac{\partial M_{0}^{t}}{\partial \theta} \right)_{\theta_{0}} \right] \left( \sigma_{\zeta}^{2} - 1 \right)). \] (5.12)

Now, using that \( \theta \in \Theta \mapsto \frac{\partial M_{0}^{t}}{\partial \theta} \) and \( \theta \mapsto \frac{\partial^{2} M_{0}^{t}}{\partial \theta^{2}} \) are continuous functions, \( \theta_{1} \xrightarrow{D} \theta_{0} \) and \( \theta_{2} \xrightarrow{D} \theta_{0} \), we claim that

\[ I_{1}^{(2)}(v) \xrightarrow{D} 0. \] (5.13)

Previous arguments induce that \( I_{1}^{(3)}(v) \) has the same limit distribution that

\[ \frac{1}{2n} \sum_{t=1}^{n} \left( M_{0}^{t} \right)^{-1} v' \left( \frac{\partial^{2} M_{0}^{t}}{\partial \theta^{2}} \right)_{\theta_{0}} \left( |\zeta_{t}| - 1 \right) \xrightarrow{a.s.} 0, \] and this implies:

\[ I_{1}^{(3)}(v) \xrightarrow{D} - \frac{1}{2} \text{E} \left[ \left( M_{0}^{t} \right)^{-2} v' \left( \frac{\partial M_{0}^{t}}{\partial \theta} \right)_{\theta_{0}} \right] \left( \sigma_{\zeta}^{2} - 1 \right). \] (5.14)

Finally, from (5.12), (5.13) and (5.14), we obtain:

\[ I_{1}(v) \xrightarrow{D} N \left( - \frac{1}{2} \text{E} \left[ \left( M_{0}^{t} \right)^{-2} v' \left( \frac{\partial M_{0}^{t}}{\partial \theta} \right)_{\theta_{0}} \right] , \text{E} \left[ \left( M_{0}^{t} \right)^{-2} v' \left( \frac{\partial M_{0}^{t}}{\partial \theta} \right)_{\theta_{0}} \right] \right) \left( \sigma_{\zeta}^{2} - 1 \right). \] (5.15)

Now we consider \( I_{2}(v) \) using again Taylor expansion, we can write:

\[ I_{2}(v) = \sum_{t=1}^{n} (M_{0}^{t})^{-1} \left( |X_{t} - f_{0}^{t}| - |X_{t} - f_{0, n-1/2}^{t}| \right) \]

\[ -n^{-1/2} \sum_{t=1}^{n} (M_{0}^{t})^{-1} \left( \frac{\partial M_{0}^{t}}{\partial \theta} \right)_{\theta_{0}} \left( \left| X_{t} - f_{0}^{t} \right| - \left| X_{t} - f_{0, n-1/2}^{t} \right| \right) \]

\[ = I_{2}^{(1)}(v) + I_{2}^{(2)}(v), \]
First we have:

\[ I_2^{(1)}(v) = \sum_{i=1}^{n} \left( |\xi_i| - |\xi_i - n^{-1/2}(M_{0_0}^t)^{-1} v' \left( \frac{\partial f_0}{\partial \theta} \right)_{\theta_0} | \right) \]

with \( \overline{\theta}_M \) in the segment \([\theta_0, \theta_0 + n^{-1/2}v]\].

Then, we have:

\[ I_2^{(2)}(v) \sim n^{-1/2} \sum_{i=1}^{n} (M_{0_0}^t)^{-2} v' \left( \frac{\partial M_0^t}{\partial \theta} \right)_{\theta_0} \left( |\xi_i - f_0^t| - |\xi_i - f_0^t + n^{-1/2} v| \right) \]

using the proof of Theorem 3.2 and denoting \( Z_t = \xi_t, Y_t = (M_{0_0}^t)^{-1} v' \left( \frac{\partial M_0^t}{\partial \theta} \right)_{\theta_0} \) and \( V_t = (M_{0_0}^t)^{-1} v' \left( \frac{\partial M_0^t}{\partial \theta} \right)_{\theta_0} \) for \( t \in \mathbb{Z} \) and condition \( \mathbb{E}[|V_t Y_t|] < \infty \) insuring a strong law of large number instead of central limit theorem for a martingale difference process. Therefore, from (5.16) and (5.17), we deduce

\[ I_2(v) \xrightarrow{D} \mathcal{N}\left( -g(0) \mathbb{E}\left[(M_{0_0}^0)^{-2} \left( v' \frac{\partial f_0}{\partial \theta} \right)_{\theta_0} \right], \mathbb{E}\left[(M_{0_0}^0)^{-2} \left( v' \frac{\partial f_0}{\partial \theta} \right)_{\theta_0} \right) \right). \]  

Finally, we obtain the behavior of \( W_n(v) \) defined in (5.11) from (5.15) and (5.18). However, we have to specify the asymptotic dependency relation between \( I_2^{(1)} \) and \( I_2^{(2)} \). Indeed these two terms converge to a Gaussian law. This implies to consider the asymptotic behavior of the sum of these two terms which could be reduced to the asymptotic behavior of:

\[ n^{-1/2} \sum_{i=1}^{n} (M_{0_0}^t)^{-1} v' \left( \frac{\partial M_0^t}{\partial \theta} \right)_{\theta_0} (|\xi_i| - 1) \]

from the proof of Theorem 3.2. Using again a central limit theorem for martingale differences, we obtain as asymptotic variance:

\[ \mathbb{E}\left[(M_{0_0}^0)^{-2} \left( v' \frac{\partial f_0}{\partial \theta} \right)_{\theta_0} \right] \left( \mathbb{E}\left[(M_{0_0}^0)^{-2} \left( v' \frac{\partial f_0}{\partial \theta} \right)_{\theta_0} \right] \right) = \mathbb{E}\left[(M_{0_0}^0)^{-2} \left( v' \frac{\partial f_0}{\partial \theta} \right)_{\theta_0} \right] (\sigma_\xi^2 - 1) \]

since \((\xi_t)_t\) admits a symmetric probability distribution with a null median and expectation. Therefore, there is no covariance term and finally we obtain:

\[ W_n(v) \xrightarrow{D} W(v) = v' \left( -\frac{1}{2} \Gamma_M - g(0) \Gamma_F \right) v + v' N \]  

with

\[ N = \xi \mathcal{N}(0, ((\sigma_\xi^2 - 1) \Gamma_M + \Gamma_F)) \]

\[ \Gamma_F = \mathbb{E}\left[(M_{0_0}^0)^{-2} \left( \frac{\partial f_0}{\partial \theta} \right)_{\theta_0} \right] \]

\[ \Gamma_M = \mathbb{E}\left[\frac{\partial \log(M_{0_0}^0)}{\partial \theta_{ij}} \right] \]  

\( 1 \leq i, j \leq d \)
2/ Now, we consider the approximation $\hat{W}_n(v)$ of $W_n(v)$ defined by:

$$\hat{W}_n(v) = -\sum_{t=1}^{n} \left( \hat{q}_t(\theta_0 + n^{-1/2}v) - \hat{q}_t(\theta_0) \right)$$

for any $v \in \mathbb{R}^d$.

From the assumptions of Theorem 3.1 and (5.7) we have $\frac{1}{n} \sum_{t=1}^{n} \| \hat{q}_t(\theta) - q_t(\theta) \|_\Theta \xrightarrow{a.s.} 0$. Then we have $\hat{W}_n(v) = W_n(v) + R_n(v)$ with $\left[ \sup_{v \in \mathbb{R}^d} |R_n(v)| \right] \leq 2 \sum_{t=1}^{n} \left[ \| \hat{q}_t(\theta) - q_t(\theta) \|_\Theta \right] \xrightarrow{a.s.} 0$ and then:

$$\hat{W}_n(v) \xrightarrow{D} W(v)$$

(5.22)

with $W$ defined in (5.20).

3/ Now, from (5.22), the proof of Theorem 3.2 and the same arguments than in the proof of Theorem 2 of Davis and Dunsmuir (1997), we deduce that finite distributions $(\hat{W}_n(v_1), \cdots, \hat{W}_n(v_k))$ converge to $(W(v_1), \cdots, W(v_k))$ for any $(v_1, \cdots, v_k) \in (\mathbb{R}^d)^k$. Moreover, always following the proof of Theorem 2, $(\hat{W}_n(v))_v$ converges to $(W(v))_v$ as a process on the continuous function space $C_0$.

As a consequence, a maximum $\hat{v}$ of $\hat{W}_n(v)$ satisfies:

$$\hat{v} = (\Gamma_M + 2g(0)\Gamma_F)^{-1}N,$$

with $N$ defined in (5.21) and this implies (3.3).

\[\square\]

**Proof of Proposition 3.1.** First, Condition C2 is satisfied since $b_0 > 0$. Other conditions on Lipschitz coefficients are also satisfied from Lemma 2.1 (see the arguments above). The identifiability condition C3 is also satisfied from the following which are divided into two parts. In (i) we prove that $(\delta, b_0, (b_i^+)(\delta, b_i^-)(\delta, b_i^-)_{i \geq 1})$ (defined in (3.5)) are unique, thereafter in (ii) we proof that $\theta = (\omega, (\alpha_i)_{1 \leq i \leq p}, (\gamma_i)_{1 \leq i \leq p}, (\beta_i)_{1 \leq i \leq q})$ is also unique.

(i) The proof of this result follow the same reasoning in Berkes et al. (2003). First we have

$$\sigma^\delta = b_0(\theta) + \sum_{i \geq 1} b_i^+(\theta)(\max(X_{t-i}, 0))^\delta + \sum_{i \geq 1} b_i^-(\theta)(\max(-X_{t-i}, 0))^\delta.$$  (5.23)

We prove the result by contradiction. Suppose that there exist two vectors $\beta = (\delta, b_0, (b_i^+), (b_i^-)_{i \geq 1})$ and $\beta' = (\delta', b_0', (b_i'^+)_{i \geq 1}, (b_i'^-)_{i \geq 1})$ verifying (5.23). Let $m > 0$ be the smallest integer satisfying $b_i^m \neq b_i'^m$ or $b_i^m \neq b_i'^m$ (if $b_i^m = b_i'^m$ and $b_i^m = b_i'^m$ for $i \geq 1$ then $b_0 = b_0'$). In one hand, since $x \in (0, \infty) \mapsto x^\delta$ is a one-to-one map and since $P(X_t = \pm 1, \forall t \in \mathbb{Z}) = 0$, we have $\delta = \delta'$. In the other hand, by definition of $m$, we have

$$\sum_{i \geq m+1} (b_i^m - b_i'^m)(\max(X_{t-i}, 0))^\delta + (b_i^m - b_i'^m)(\max(-X_{t-i}, 0))^\delta = b_0 - b_0' + \sum_{i \geq m+1} (b_i^m - b_i'^m)(\max(X_{t-i}, 0))^\delta + \sum_{i \geq m+1} (b_i^m - b_i'^m)(\max(-X_{t-i}, 0))^\delta.$$  (5.24)

From (3.4), we have $X_{t-m} = \sigma_{t-m} \zeta_{t-m}$, therefore

$$(b_i^m - b_i'^m)(\max(X_{t-m}, 0))^\delta + (b_i^m - b_i'^m)(\max(-X_{t-m}, 0))^\delta = \begin{cases} (b_i^m - b_i'^m)\sigma_{t-m}^\delta \zeta_{t-m}^\delta & \text{when } \zeta_{t-m} \geq 0 \\ (b_i^m - b_i'^m)\sigma_{t-m}^\delta (-\zeta_{t-m})^\delta & \text{when } \zeta_{t-m} < 0 \end{cases}$$

Moreover (5.24) and the fact that $b_i^m \neq b_i'^m$ or $b_i^m \neq b_i'^m$ implies that at least one of the following equalities hold

$$\sum_{i \geq m+1} (b_i^m - b_i'^m)(\max(X_{t-i}, 0))^\delta = (b_i^m - b_i'^m)\sigma_{t-m}^\delta \zeta_{t-m}^\delta$$

or

$$\sum_{i \geq m+1} (b_i^m - b_i'^m)(\max(X_{t-i}, 0))^\delta = (b_i^m - b_i'^m)\sigma_{t-m}^\delta (-\zeta_{t-m})^\delta$$

When $\zeta_{t-m} \geq 0$.

Since $\sigma_{t-m}^\delta > b_0 > 0$, $\zeta_{t-m}^\delta$ is well defined. Let $F_k$ be the $F$-algebra generated by $(\zeta_i, i < k)$. The causal representation of the APARCH($\delta, p, q$) shows that $X_j$ is $F_j$-measurable and thus the right-hand side of the above equations (and
consequently also $\zeta_{t-m}^\delta$ in the case $\zeta_{t-m} \geq 0$ or the case $\zeta_{t-m} < 0$ is a real-valued random variable, measurable with respect to $\mathcal{F}_{t-m-1}$. Since $(\zeta_t)$ is a sequence of independent random variables, this implies that $\zeta_{t-m}$ is a.s. constant when $\zeta_{t-m} \geq 0$ or when $\zeta_{t-m} < 0$, contradicting the hypothesis saying $\zeta_{t-m}^\delta$ has a non-degenerate distribution. This achieves (i).

(ii) The representation (5.23) is the same as

$$\sigma_t^{\delta} = b_0 + \Psi^+(L)(\max(X_t, 0))^\delta + \Psi^-(L)(\max(-X_t, 0))^\delta.$$

with $\Psi^+ = \Psi_{\theta_1}^{-1}\Delta_{\theta_2}^+,$ $\Psi^- = \Psi_{\theta_1}^{-1}\Delta_{\theta_2}^{-}$ and $\Delta_{\theta_2}^+(L) = \sum_{i=1}^p \alpha_i(1 - \gamma_i)L_i$, $\Delta_{\theta_2}^-(L) = \sum_{i=1}^q \alpha_i(1 + \gamma_i)L_i$ and $\Psi_{\theta_1}(L) = \sum_{i=1}^q \beta_i L_i$, where $(\Delta_{\theta_2}^+, \Psi_{\theta_1})$ and $(\Delta_{\theta_2}^-, \Psi_{\theta_1})$ respectively coprime and $\theta_1 = (\beta_i)_{1 \leq i \leq q}$, $\theta_2 = (\alpha_i)_{1 \leq i \leq p} (\gamma_i)_{1 \leq i \leq p}$, then $\theta = (\omega, \theta_1, \theta_2)$.

Suppose that there exist others polynomials $\Delta_{\theta_2}^+ = \sum_{i=1}^q \alpha_i' (1 - \gamma_i') L_i$, $\Delta_{\theta_2}^- = \sum_{i=1}^q \alpha_i' (1 + \gamma_i') L_i$, $\Psi_{\theta_1}^2 = \sum_{i=1}^q \beta_i' L_i$ satisfying $\Psi^+ = \Psi_{\theta_1}^{-1}\Delta_{\theta_2}^+$, $\Psi^- = \Psi_{\theta_1}^{-1}\Delta_{\theta_2}^-$ with $(\Delta_{\theta_2}^+, \Psi_{\theta_1})$, $(\Delta_{\theta_2}^-, \Psi_{\theta_1})$ respectively coprime. Then

$$\begin{cases}
\Psi_{\theta_1}^{-1}\Delta_{\theta_2}^+ = \Psi_{\theta_1}^{-1}\Delta_{\theta_2}^+
\Psi_{\theta_1}^{-1}\Delta_{\theta_2}^- = \Psi_{\theta_1}^{-1}\Delta_{\theta_2}^-
\end{cases}$$

from the first equality, since $\deg(\Delta_{\theta_2}^+) = \deg(\Delta_{\theta_2}^-) = q$, we conclude that $\Psi_{\theta_1}^{-1}\Delta_{\theta_2}^+ = 1$, therefore $\Psi_{\theta_1} = \Psi_{\theta_1}'$ and so $\Delta_{\theta_2}^+ = \Delta_{\theta_2}^+$. Likewise from the second equality we conclude that $\Delta_{\theta_2}^- = \Delta_{\theta_2}^-$.  

- The equalities $\Delta_{\theta_2}^+ = \Delta_{\theta_2}^+$, $\Delta_{\theta_2}^- = \Delta_{\theta_2}^-$ implies that $\alpha_i(1 - \gamma_i) = \alpha_i' (1 - \gamma_i')$ and $\alpha_i(1 + \gamma_i) = \alpha_i' (1 + \gamma_i')$ which give $\alpha_i = \alpha_i'$ and $\gamma_i = \gamma_i'$.
- The equality $\Psi_{\theta_1} = \Psi_{\theta_1}'$ implies that $\beta_i = \beta_i'$.
- Since $(\beta_i)_{i=1,p}, b_0 = w(1 - \sum_{j=1}^q \beta_j)^{-1}$ are unique then $\omega$ is unique.

Thus, Condition C3 is established and the proof of proposition is achieved. \(\square\)

**Proof of Proposition 3.2.** Since we prove that Lemma 2.1 implies that conditions on Lipschitzian coefficients $(\alpha_j^i(f, \Theta))$, $(\alpha_j^i(M, \Theta))$, it remains to prove conditions C2 and C3. Condition C2 holds since $c_0$ is supposed to be a positive number. Finally, condition C3 also holds since $f_0 = f_{\theta'}$ implies $\psi_j(\theta) = \psi_j(\theta')$ for all $j \in \mathbb{Z}$. Therefore the parameters of the ARMA part of the process are identified and then the identification of the parameters GARCH can be deduced from the proof of Proposition 3.1. \(\square\)

**Proof of Proposition 3.4.** This proofs mimics exactly the proof of Proposition 3.2.

Acknowledgements. The authors are grateful to the referees and associated editor for their comments which helped to improve the contents of this paper.

References


