Electronic Journal of Statistics
Vol. 11 (2017) 452–479
ISSN: 1935-7524
DOI: 10.1214/17-EJS1241

Asymptotic behavior of the Laplacian quasi-maximum likelihood estimator of affine causal processes

Jean-Marc Bardet

S.A.M.M., Université Panthéon-Sorbonne, 90, rue de Tolbiac, 75634, Paris, France
e-mail: bardet@univ-paris1.fr

Yakoub Boularouk and Khedidja Djaballah

Université des Sciences et de la Technologie Houari Boumédiene, 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.

MSC 2010 subject classifications: Primary 62M10, 62M10; secondary 60G10.
Keywords and phrases: Laplacian quasi-maximum likelihood estimator, strong consistency, asymptotic normality, ARMA-ARCH processes.

Received April 2016.

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 [8], [1] and [2]. Hence, we will consider an observed sample $(X_1, \cdots, X_n)$ where $(X_t)_{t\in\mathbb{Z}}$ is a solution of the following equation:

$$X_t = M_{\theta_0}(X_{t-1}, X_{t-2}, \cdots)\,\zeta_t + f_{\theta_0}(X_{t-1}, X_{t-2}, \cdots), \qquad t \in Z, \tag{1.1}$$

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.* $\zeta_0 \stackrel{\mathcal{L}}{=} -\zeta_0$, satisfying $\mathrm{E}[|\zeta_0|^r] < \infty$ with $r \geq 1$ and $\mathrm{E}[|\zeta_0|] = 1$. If $r \geq 2$, denote $\sigma_\zeta^2 = \mathrm{Var}(\zeta_0)$;

- $(\theta,(x_n)_{n\in\mathbb{N}}) \to M_\theta((x_n)_{n\in\mathbb{N}}) \in (0,\infty)$ and $(\theta,(x_n)_{n\in\mathbb{N}}) \to f_\theta((x_n)_{n\in\mathbb{N}}) \in \mathbb{R}$ are two known applications.

For instance, if $M_{\theta_0}(X_{t-1},X_{t-2},\cdots)=1$ and $f_{\theta_0}(X_{t-1},X_{t-2},\cdots)=\alpha_0\,X_{t-1}$ with $|\alpha_0|<1$ then (X_t) is a causal AR(1) process. In [8] and [1], 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 [1], it was also established that under several conditions on M_θ, f_θ and if $\mathrm{E}[|\zeta_0|^r]$ with $r\geq 2$, the usual Gaussian Quasi-Maximum Likelihood Estimator (QMLE) of θ is strongly consistent and when $r\geq 4$ it is asymptotically normal. This estimator was first defined by [24] for ARCH processes, and the asymptotic study of this estimator was first obtained by [17] for GARCH$(1,1)$ processes, [3] for GARCH(p,q) processes, [9] for ARMA-GARCH processes, [22] for general heteroskedastic models, and [21] for ARCH(∞) processes. The results of [1] 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 (ζ_t) is a Gaussian sequence and even if it could be applied when the probability distribution of (ζ_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 $\sum_{t=1}^n M_\theta^{-2}(X_t-f_\theta)^2$) and this induces that $r=2$ is required for the consistency and $r=4$ for the asymptotic normality (and therefore confidence intervals or tests). For numerous real data such requirement is sometimes too strong (for instance the kurtosis of economic data is frequently considered as infinite). Moreover, such estimator is not robust to potential outliers. Hence, the reference probability distribution of (ζ_t) could be a Laplace one and this allows to avoid both these drawbacks. Roughly speaking this choice implies to minimize a least absolute deviations contrast (typically $\sum_{t=1}^n M_\theta^{-1}|X_t-f_\theta|$) instead of the previous least squares contrast. And therefore, $r=1$ will be sufficient for insuring the strong consistency of this Laplacian-QMLE, while only $r=2$ is required for the asymptotic normality (see below).

Such probability distribution choice is not new since this leads to a Least Absolute Deviations (LAD) estimation. Hence, for ARMA processes, [6] proved the consistency and asymptotic normality of the LAD estimator. For ARCH or GARCH processes, the same results concerning the LAD estimator were already established by [19], while [4] proved the consistency and asymptotic normality of the Laplacian-QMLE. [18] considered also the estimator for other conditional heteroskedasticity models. Recently, [10] proposed a two-stage non-Gaussian-QML estimation for GARCH models and [12] proposed an alternative one-step procedure, based on an appropriate non-Gaussian-QML estimator, the asymptotic properties of both these approaches were studied.

In this paper we unify all these studies of the Laplacian-QMLE in a simple framework, *i.e.* causal stationary solutions of (1.1). This notably allows to obtain known results on ARMA or GARCH but also to establish for the first time the consistency and the asymptotic normality of the Laplacian-QMLE for APARCH, ARMA-GARCH, ARMA-ARCH(∞) 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 (ζ_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 θ 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 ζ_0, then, from the affine causal definition of X, this log-likelihood can be written:

$$\log\big(L_\theta(X_1, \cdots, X_n)\big) = \sum_{t=1}^n \log\Big(\frac{1}{M_\theta^t}\, h\Big(\frac{X_t - f_\theta^t}{M_\theta^t}\Big)\Big)$$

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

$$\log\big(QL_\theta(X_1, \cdots, X_n)\big) = \sum_{t=1}^n \log\Big(\frac{1}{M_\theta^t}\, h\Big(\frac{X_t - \widehat{f}_\theta^t}{\widehat{M}_\theta^t}\Big)\Big),$$

with $\widehat{f}_\theta^t := f_\theta(X_{t-1}, \ldots, X_1, u)$ and $\widehat{M}_\theta^t := M_\theta(X_{t-1}, \ldots, X_1, u)$, where $u = (u_n)_{n\in N}$ is a finitely non-zero sequence $(u_n)_{n\in N}$. The choice of $(u_n)_{n\in N}$ 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:

$$\widetilde{\theta}_n := \mathrm{Argmax}_{\theta\in\Theta} \log\big(QL_\theta(X_1, \cdots, X_n)\big).$$

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

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

and this provides the Gaussian-QMLE of θ.

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

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

and this implies $\mathrm{E}\big[|\zeta_0|\big] = 1$.

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

$$L_n(\Theta) = -\sum_{t=1}^{n} q_t(\Theta) \quad \text{with} \quad q_t(\Theta) = \log|M_\theta^t| + |M_\theta^t|^{-1}|X_t - f_\theta^t| \tag{2.2}$$

$$\widehat{L}_n(\theta) = -\sum_{t=1}^{n} \widehat{q}_t(\theta) \quad \text{with} \quad \widehat{q}_t(\theta) := \log|\widehat{M}_\theta^t| + |\widehat{M}_\theta^t|^{-1}|X_t - \widehat{f}_\theta^t|. \tag{2.3}$$

Hence, if it exists, a Laplacian-QMLE $\widehat{\theta}_n$ is a maximizer of $\widehat{L}_n$:

$$\widehat{\theta}_n := \arg\max_{\theta \in \Theta} \widehat{L}_n(\theta).$$

We restrict the set Θ 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 $\widehat{\theta}_n$. More details are given now.

2.2. Existence and stationarity

As it was already done in [8] and [1], several Lipschitz-type inequalities on f_θ and M_θ 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 Θ of $\mathbb{R}^d$, define a Lipschitz assumption on function K_θ:

Assumption $(\mathbf{A}_k(K, \Theta))$ $\forall x \in \mathbb{R}^\infty$, $\theta \in \Theta \mapsto K_\theta(x) \in \mathcal{C}^k(\Theta)$ and $\partial_\theta^k K_\theta$ satisfies $\big\|\partial_\theta^k K_\theta(0)\big\|_\Theta < \infty$ and there exists a sequence $\big(\alpha_j^{(k)}(K,\Theta)\big)_j$ of nonnegative numbers such that $\forall x,\, y \in \mathbb{R}^{\mathbb{N}}$

$$\big\|\partial_\theta^k K_\theta(x) - \partial_\theta^k K_\theta(y)\big\|_\Theta \le \sum_{j=1}^{\infty} \alpha_j^{(k)}(K,\Theta)|x_j - y_j|, \text{ with } \sum_{j=1}^{\infty} \alpha_j^{(k)}(K,\Theta) < \infty$$

For ensuring a stationary r-order solution of (1.1), for $r \ge 1$, define the set

$$\Theta(r) := \Big\{\theta \in \mathbb{R}^d,\ (\mathbf{A}_0(f,\{\theta\})) \text{ and } (\mathbf{A}_0(M,\{\theta\})) \text{ hold},\\ \sum_{j=1}^{\infty} \alpha_j^{(0)}(f,\{\theta\}) + \big(\mathrm{E}[|\zeta_0|^r]\big)^{1/r} \sum_{j=1}^{\infty} \alpha_j^{(0)}(M,\{\theta\}) < 1\Big\}.$$

Then, from [8], 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 $(\zeta_i)_{i>t}$ for $t \in \mathbb{Z}$) solution X of (1.1), which is stationary, ergodic and satisfies $\mathrm{E}\big[|X_0|^r\big] < \infty$.

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:

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

Lemma 2.1. Let X be a.s. a causal stationary solution of (1.1) for $\theta_0 \in \mathbb{R}^d$. Let $\widetilde{X}$ be such as $\widetilde{X}_t = \Lambda_\beta(L)\, X_t$ for $t \in \mathbb{Z}$ with $\Lambda_{\beta_0}(L) = P_{\beta_0}^{-1}(L)\, Q_{\beta_0}(L)$ 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 $\widetilde{X}$ is a.s. a causal stationary solution of the equation

$$\widetilde{X}_t = \widetilde{M}_{\widetilde{\theta}_0}\big((\widetilde{X}_{t-i})_{i\geq 1}\big)\, \zeta_t + \widetilde{f}_{\widetilde{\theta}_0}\big((\widetilde{X}_{t-i})_{i\geq 1}\big) \quad \text{for } t \in \mathbb{Z},$$

where $\widetilde{f}_{\theta_0}$ and $\widetilde{M}_{\theta_0}$ are given in (5.1) and $\widetilde{\theta}_0 = (\theta_0, \beta_0)$. Moreover, for $i = 0,\, 1,\, 2$ and with $K = f$ or M and $\widetilde{K} = \widetilde{f}$ or $\widetilde{M}$,

- if $\alpha_j^{(i)}(K,\{\theta_0\}) = O(j^{-\beta})$ and $\beta > 1$, then $\alpha_j^{(i)}(\widetilde{K},\{\widetilde{\theta}_0\}) = O(j^{-\beta})$;
- if $\alpha_j^{(i)}(K,\{\theta_0\})$ is EDS, then $\alpha_j^{(i)}(\widetilde{K},\{\widetilde{\theta}_0\})$ 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_θ and M_θ:

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

$$\alpha_j^{(0)}(f,\Theta) + \alpha_j^{(0)}(M,\Theta) = \mathcal{O}(j^{-\ell}) \ \text{ for some } \ell > \frac{2}{\min(r\,,\,2)} \tag{3.1}$$

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

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 [1]). Moreover, if $r = 2$, the condition (3.1) on Lipshitzian 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(∞) process can be defined for satisfying the strong consistency of Laplacian-QMLE while the conditions given in [1] do not ensure the strong consistency of Gaussian-QMLE.

Now we state an extension of Theorem 1 established in [6] 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 $\mathrm{Var}(Z_0) = \sigma^2 < \infty$, 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 $\mathrm{E}\big[Y_0^2 V_0^2\big] < \infty$. Then

$$\sum_{t=1}^{n} V_{t-1}\big(|Z_t - n^{-1/2} Y_{t-1}| - |Z_t|\big) \ \overset{\mathcal{D}}{\underset{n\to\infty}{\longrightarrow}} \ \mathcal{N}\Big(f(0)\,\mathrm{E}\big[V_0 Y_0^2\big]\,,\,\mathrm{E}\big[V_0^2 Y_0^2\big]\Big) \tag{3.2}$$

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

Theorem 3.3. Assume that $\theta_0 \in \overset{\circ}{\Theta} \cap \Theta(r)$ where $r \geq 2$ and $\overset{\circ}{\Theta}$ denotes the interior of Θ. 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 ζ_0 is continuously differentiable in a neighborhood of 0 with derivative $g(0)$ in 0 and if matrix Γ_F or Γ_M, defined in (5.22), are definite positive symmetric matrix, then

$$\sqrt{n}\big(\widehat{\theta}_n-\theta_0\big) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}_d\Big(0,\big(\Gamma_M+2g(0)\Gamma_F\big)^{-1}\big((\sigma_\zeta^2-1)\Gamma_M+\Gamma_F\big)\big(\Gamma_M+2g(0)\Gamma_F\big)^{-1}\Big). \tag{3.3}$$

As it was already proved for the median estimator (see [23]) or for least absolute deviations estimator of ARMA process (see [6]), it is not surprising that the probability density function g of the white noise $(\zeta_i)_i$ 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 [10] 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 [6] 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 [6]. Thus our results could appear as extensions of this paper.
- The second and certainly closest paper, [1]. 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 [1] 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 $q_t(\theta)$ 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 [1]. Indeed, in our new proof, we use a condition

of almost sure uniform consistency based on a general and powerful result established in [14] while a Feller-type condition was "only" used in [1]. This difference leads to a very sharp condition on the decreasing rate of the Lipshitzian coefficients $(\alpha_k^{(0)})$ 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 [6], *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_θ and M_θ are clearly weaker with Laplacian than with Gaussian QMLE.

- The third related paper is [10]. 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 ($X_t = M_\theta^t\,\xi_t + f_\theta^t$) 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.* $\sup_\theta |\sigma_t(\theta) - \hat\sigma_t(\theta)| \le C_1\,\rho^t$ is clearly weaker than our Lipshitzian condition (for instance $ARCH(\infty)$ processes with Riemanian decay of the coefficients could satisfy our conditions but not their conditions). In [10], 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_θ^t typically present for ARMA processes, is not considered in their models. Note also that [12] 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 [1] and Bardet *et al.* (2012) for ARCH(∞) and TARCH(∞).

1/ APARCH processes. APARCH(δ, p, q) model has been introduced (see [7]) as the solution of equations

$$\begin{cases} X_t = \sigma_t \zeta_t, \\ \sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i (|X_{t-i}| - \gamma_i X_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta, \end{cases} \tag{3.4}$$

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 α_p, β_q strictly positive and $\sum_{j=1}^q \beta_j < 1$. Hence, we denote here $\theta = \big(\delta, \omega, \alpha_1, \ldots, \alpha_p, \gamma_1, \ldots, \gamma_p, \beta_1, \ldots, \beta_q\big)$.

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

$$\begin{aligned} \sigma_t^\delta &= \big(1 - \sum_{j=1}^q \beta_j L^j\big)^{-1} \Big[\omega + \sum_{i=1}^p \alpha_i (1-\gamma_i)^\delta (\max(X_{t-i}, 0))^\delta \\ &\qquad + \alpha_i (1+\gamma_i)^\delta (-\min(X_{t-i}, 0))^\delta \Big] \\ &= b_0 + \sum_{i \geq 1} b_i^+ (\max(X_{t-i}, 0))^\delta + \sum_{i \geq 1} b_i^- (\max(-X_{t-i}, 0))^\delta. \end{aligned}$$

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

$$\begin{cases} b_i^+ = \sum_{k=1}^q \beta_k b_{i-k}^+ + \alpha_i (1-\gamma_i)^\delta & \text{with } \alpha_i(1-\gamma_i) = 0 \text{ for } i > p \\ b_i^- = \sum_{k=1}^q \beta_k b_{i-k}^- + \alpha_i (1+\gamma_i)^\delta & \text{with } \alpha_i(1+\gamma_i) = 0 \text{ for } i > p \end{cases} \tag{3.5}$$

with $b_i^+ = b_i^- = 0$ for $i \leq 0$. As a consequence, for APARCH model, $f_\theta^t \equiv 0$ and $M_\theta^t = \sigma_t$. It is clear that $\alpha_j^{(0)}(f, \Theta) = 0$ and simple computations imply $\alpha_j^{(0)}(M, \Theta) = \sup_{\theta \in \Theta} \max\big(|b_j^+(\theta)|^{1/\delta}, |b_j^-(\theta)|^{1/\delta}\big)$. 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

$$\Theta(r) = \Big\{\theta \in \mathbb{R}^{2p+q+2} \Big/ \big(\mathrm{E}\big[|\zeta_0|^r\big]\big)^{1/r} \sum_{j=1}^\infty \max\big(|b_j^+|^{1/\delta}, |b_j^-|^{1/\delta}\big) < 1\Big\}. \tag{3.6}$$

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 Θ is a compact subset of $\Theta(r)$ defined in (3.6). Then,

1. If $r = 1$, then $\widehat{\theta}_n \xrightarrow[n\to\infty]{a.s.} \theta_0$.
2. If $r = 2$, and if Γ_M defined in (5.22) is a definite positive symmetric matrix, then

$$\sqrt{n}\big(\widehat{\theta}_n - \theta_0\big) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}_{2p+q+2}\big(0\ ,\ (\sigma_\zeta^2 - 1)\ \Gamma_M^{-1}\big).$$

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 [7] and [16] as the solution of the system of equations

$$\begin{cases} P_\theta(L)\ X_t = Q_\theta(L)\ \varepsilon_t, \\ \varepsilon_t = \sigma_t \zeta_t, \quad \text{with} \quad \sigma_t^2 = c_0 + \sum_{i=1}^{p'} c_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q'} d_i \sigma_{t-i}^2 \end{cases} \tag{3.7}$$

where

- $c_0 > 0$, $c_i \geq 0$ for $i = 1, \dots, p'$, $d_i \geq 0$ for $i = 1, \dots, q'$, $\sum_{i=1}^{q'} d_i < 1$ and $c_{p'}, d_{q'}$ positive;
- $P_\theta(x) = 1 - a_1 x - \cdots - a_p x^p$ and $Q_\theta(x) = 1 - b_1 x - \cdots - b_q x^q$ are coprime polynomials with $\sum_{i=1}^{p} |a_i| < 1$ and $\sum_{i=1}^{p} |b_i| < 1$.

Let $\theta = (c_0, c_1, \dots, c_{p'}, d_1, \dots, d_{q'}, a_1, \dots, a_p, b_1, \dots, b_q)$. We are going to use Lemma 2.1. Since (ε_t) is supposed to be a GARCH(p',q'), then $f_\theta^\varepsilon = 0$ and $M_\theta^\varepsilon = \big(\big(1 - \sum_{j=1}^{q'} d_j L^j\big)^{-1}\big(c_0 + c_1 \varepsilon_{t-1}^2 + \cdots + c_{p'} \varepsilon_{t-p'}^2\big)^{1/2}$ and direct computations imply that the Lipshitz coefficients of (ε_t) are such as $\alpha_j^{(0)}(f^\varepsilon, \{\theta_0\}) = 0$ and $\alpha_j^{(0)}(M^\varepsilon, \{\theta_0\}) = |\beta_j|$ with $\big(1 + \sum_{j=1}^\infty \beta_j x^j\big)\big(1 - \sum_{j=1}^{q'} d_j x^j\big) = \sum_{j=0}^{p'} c_j x^j$. Therefore $(\alpha_j^{(0)}(f^\varepsilon, \{\theta_0\}))_j$ and $(\alpha_j^{(0)}(M^\varepsilon, \{\theta_0\}))_j$ are EDS (see for instance [4]). Thus $(A_0(f^\varepsilon, \{\theta_0\}))$ and $(A_0(M^\varepsilon, \{\theta_0\}))$ hold.

Considering the ARMA part and denoting (ψ_j) such as $\big(1 + \sum_{j=1}^\infty \psi_j x^j\big)\big(1 - \sum_{j=1}^\infty a_j x^j\big) = \big(1 - \sum_{j=1}^\infty b_j x^j\big)$, then from Lemma 2.1 we deduce that:

$$\begin{cases} \alpha_j^{(0)}(f, \{\theta_0\}) & = & |\psi_j| \\ \alpha_j^{(0)}(M, \{\theta_0\}) & \leq & \sum_{k=1}^{j} |\psi_k| \times |\beta_{j-k}| \end{cases}.$$

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

$$\Theta(r) = \Big\{\theta \in \mathbb{R}^{p+q+p'+q'+1} \Big/ \sum_{i=1}^\infty |\psi_i(\theta)| + \big(\mathrm{E}\big[|\zeta_0|^r\big]\big)^{1/r} \sum_{j=1}^\infty \sum_{k=1}^{j} |\psi_k| \times |\beta_{j-k}| < 1\Big\}. \tag{3.8}$$

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 Θ is a compact subset of $\Theta(r)$ defined in (3.8). Then,

1. If $r = 1$, then $\widehat{\theta}_n \xrightarrow[n\to\infty]{a.s.} \theta_0$.
2. If $r = 2$, and if Γ_f and Γ_M defined in (5.22) are definite positive symmetric matrix, then the asymptotic normality (3.3) of $\widehat{\theta}_n$ holds.

This result is a new one and extends the previous results already obtained with Gaussian-QMLE for such processes (see for instance, [16] and [1]).

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

$$\begin{cases} P_\theta(L)\ X_t = Q_\theta(L)\ \varepsilon_t, \\ \varepsilon_t = \sigma_t \zeta_t, \quad \text{with} \quad \sigma_t^2 = c_0 + \sum_{i=1}^\infty c_i \varepsilon_{t-i}^2 \end{cases} \tag{3.9}$$

where

- $c_0 > 0$, $c_i \geq 0$ for $i \geq 1$;
- $P_\theta(x) = 1 - a_1 x - \cdots - a_p x^p$ and $Q_\theta(x) = 1 - b_1 x - \cdots - b_q x^q$ are coprime polynomials with $\sum_{i=1}^p |a_i| < 1$ and $\sum_{i=1}^p |b_i| < 1$.

ARCH(∞) processes were introduced by [20] and the asymptotic properties of Gaussian-QMLE were studied in [21], [22] or [1]. Hence, we assume that there exists $\beta = (\beta_1, \cdots, \beta_m)$ such as for all $i \in \mathbb{N}$, $c_i = c(i, \beta)$, with $c(\cdot)$ a known function. Let $\theta = (\beta, a_1, \ldots, a_p, b_1, \ldots, b_q)$. We are going to use Lemma 2.1. Since (ε_t) is supposed to be an ARCH(∞), then $f_\theta^\varepsilon = 0$ and $M_\theta^\varepsilon = \big(c(0, \beta) + \sum_{i=1}^\infty c(i, \beta) \varepsilon_{t-i}^2\big)^{1/2}$ and direct computations imply that the Lipshitz coefficients of (ε_t) are such as $\alpha_j^{(0)}(f^\varepsilon, \{\theta_0\}) = 0$ and $\alpha_j^{(0)}(M^\varepsilon, \{\theta_0\}) = c(j, \beta_0)$. Therefore we assume that there exists $\ell > 1$ such as

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

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

Considering the ARMA part and denoting (ψ_j) such as $\big(1 + \sum_{j=1}^\infty \psi_j x^j\big)\big(1 - \sum_{j=1}^\infty a_j x^j\big) = \big(1 - \sum_{j=1}^\infty b_j x^j\big)$, then from Lemma 2.1 we deduce that:

$$\begin{cases} \alpha_j^{(0)}(f, \{\theta_0\}) & = & |\psi_j| \\ \alpha_j^{(0)}(M, \{\theta_0\}) & \leq & \sum_{k=1}^j |\psi_k| \times c(j, \beta_0) \end{cases}.$$

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

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

Now the strong consistency and asymptotic normality of the Laplacian-QMLE for ARMA-ARCH(∞) 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 Θ is a compact subset of $\Theta(r)$ defined in (3.11). Then,

1. If $r \geq 1$ and $\ell \geq 2/\min(r,2)$, then $\widehat{\theta}_n \xrightarrow[n\to\infty]{a.s.} \theta_0$.
2. If $r = 2$, $\ell > 1$ and if $\partial^i_\beta c(j,\beta) = \mathcal{O}\big(j^{-\ell}\big)$ for $i = 1,\ 2$, and if Γ_f and Γ_M defined in (5.22) are definite positive symmetric matrix, then the asymptotic normality (3.3) of $\widehat{\theta}_n$ 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 [1]). 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 [7] as the solutions of the equations

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

where:

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

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

Then, we have

$$\begin{cases} \alpha_j^{(0)}(f, \{\theta_0\}) & \leq \quad |\psi_j| \\ \alpha_j^{(0)}(M, \{\theta_0\}) & \leq \quad \sum_{k=1}^{j} |\psi_k| \times \max\big(|b_{j-k}^+|^{1/\delta}, |b_{j-k}^-|^{1/\delta}\big) \end{cases}.$$

(ψ_j) such as $\big(1 + \sum_{j=1}^\infty \psi_j x^j\big)\big(1 - \sum_{j=1}^\infty a_j x^j\big) = \big(1 - \sum_{j=1}^\infty b_j x^j\big)$. 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

$$\Theta(r) = \Big\{\theta \in \mathbb{R}^s \ \Big/ \ \sum_{j=1}^\infty |\psi_j| + \big(\mathrm{E}\big|\zeta_0|^r\big]\big)^{1/r} \sum_{j=1}^\infty \sum_{k=1}^j |\psi_k| \\ \times \max\big(|b_{j-k}^+|^{1/\delta}, |b_{j-k}^-|^{1/\delta}\big) < 1\Big\}.$$

with $s = p+q+p'+q'+2$. 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 Θ is a compact subset of $\Theta(r)$ defined in (3.8). Then,

1. If $r = 1$, then $\widehat{\theta}_n \xrightarrow[n\to\infty]{a.s.} \theta_0$.
2. If $r = 2$, and if Γ_f and Γ_M defined in (5.22) are definite positive symmetric matrix, then the asymptotic normality (3.3) of $\widehat{\theta}_n$ 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 [1]. 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 bevarior of Laplacian-QMLE (denoted $\widehat{\theta}_n^{LQL}$) 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 (ζ_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 $0.05 * \mathcal{N}(-2, 0.16) + \mathcal{N}(0, 1) + 0.05 * \mathcal{N}(2, 0.16)$ and denoted $\mathcal{M}$.

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

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 $\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta\sigma_{t-1}^2$ with $\alpha_0 = 0.2$, $\alpha_1 = 0.4$ and $\beta = 0.2$;
- a ARMA$(1,1)$-GARCH$(1,1)$ process defined by $X_t = \phi X_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$ where $\varepsilon_t = \zeta_t\, \sigma_t$ and $\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$ 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 $X_t = \phi X_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$ where $\varepsilon_t = \zeta_t\, \sigma_t$ and $\sigma_t^\delta = \alpha_0 + \alpha_1\big(|\varepsilon_{t-1}| - \gamma\varepsilon_{t-1}\big)^\delta + \beta\sigma_{t-1}^\delta$ and $\phi = 0.4$, $\theta = 0.6$, $\alpha_0 = 0.2$, $\alpha_1 = 0.4$, $\gamma = 0.5$, $\beta = 0.1$ and $\delta = 1.2$.

Hence we computed the root-mean-square error (RMSE) from 1000 independent replications of $\widehat{\theta}_n^{LQL}$ and $\widehat{\theta}_n^{LQL}$ for those processes and the results are presented in Table 1 on page 465 and 2 on page 466.

TABLE 1
Root Mean Square Error of the components of $\widehat{\theta}_n^{LQL}$ and $\widehat{\theta}_n^{GQL}$ for ARMA$(1,1)$, ARCH(1) and GARCH$(1,1)$ processes.

			$\mathcal{L}$		$\mathcal{N}$		t_3		$\mathcal{U}$		$\mathcal{M}$	
		n	$\widehat{\theta}_n^{GQL}$	$\widehat{\theta}_n^{LQL}$	$\widehat{\theta}_n^{GQL}$	$\widehat{\theta}_n^{LQL}$	$\widehat{\theta}_n^{GQL}$	$\widehat{\theta}_n^{LQL}$	$\widehat{\theta}_n^{GQL}$	$\widehat{\theta}_n^{LQL}$	$\widehat{\theta}_n^{GQL}$	$\widehat{\theta}_n^{LQL}$
ARMA(1,1)	θ	100	0.106	0.091	0.114	0.117	0.113	0.090	0.112	0.059	0.110	0.078
		1000	0.031	0.024	0.032	0.032	0.036	0.027	0.031	0.014	0.031	0.023
		5000	0.014	0.010	0.014	0.015	0.016	0.011	0.016	0.012	0.013	0.010
	ϕ	100	0.119	0.102	0.121	0.128	0.123	0.102	0.120	0.067	0.121	0.090
		1000	0.037	0.028	0.036	0.036	0.040	0.030	0.036	0.017	0.036	0.027
		5000	0.016	0.012	0.014	0.016	0.017	0.013	0.016	0.007	0.014	0.006
ARCH(1)	ω	100	0.068	0.061	0.048	0.049	0.254	0.085	0.035	0.025	0.062	0.052
		1000	0.020	0.018	0.015	0.015	0.134	0.049	0.011	0.016	0.036	0.018
		5000	0.010	0.009	0.006	0.006	0.115	0.044	0.005	0.015	0.031	0.008
	α	100	0.161	0.155	0.141	0.142	0.979	0.418	0.102	0.064	0.484	0.423
		1000	0.063	0.058	0.043	0.043	0.852	0.169	0.029	0.033	0.157	0.133
		5000	0.016	0.014	0.012	0.012	0.378	0.109	0.013	0.031	0.087	0.062
GARCH(1,1)	α_0	100	0.112	0.105	0.095	0.100	0.211	0.126	0.081	0.047	0.134	0.114
		1000	0.036	0.032	0.028	0.028	0.098	0.058	0.023	0.018	0.066	0.051
		5000	0.016	0.014	0.012	0.012	0.055	0.043	0.010	0.015	0.040	0.023
	α_1	100	0.162	0.157	0.149	0.150	0.453	0.364	0.115	0.070	0.507	0.429
		1000	0.061	0.056	0.449	0.449	0.333	0.150	0.030	0.033	0.160	0.136
		5000	0.029	0.026	0.020	0.020	0.193	0.095	0.013	0.030	0.086	0.058
	β	100	0.225	0.209	0.188	0.190	0.499	0.429	0.163	0.105	0.483	0.390
		1000	0.060	0.055	0.051	0.051	0.285	0.174	0.044	0.022	0.170	0.169
		5000	0.027	0.024	0.022	0.022	0.180	0.075	0.019	0.009	0.072	0.075

TABLE 2

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

			$\mathcal{L}$		$\mathcal{N}$		t_3		$\mathcal{U}$		$\mathcal{M}$	
		n	$\widehat{\theta}_n^{GQL}$	$\widehat{\theta}_n^{LQL}$	$\widehat{\theta}_n^{GQL}$	$\widehat{\theta}_n^{LQL}$	$\widehat{\theta}_n^{GQL}$	$\widehat{\theta}_n^{LQL}$	$\widehat{\theta}_n^{GQL}$	$\widehat{\theta}_n^{LQL}$	$\widehat{\theta}_n^{GQL}$	$\widehat{\theta}_n^{LQL}$
ARMA$(1,1)$	θ	100	0.120	0.097	0.107	0.107	0.121	0.098	0.097	0.067	0.123	0.087
-GARCH$(1,1)$		1000	0.035	0.024	0.028	0.028	0.048	0.029	0.024	0.015	0.035	0.026
		5000	0.016	0.010	0.012	0.012	0.023	0.012	0.015	0.011	0.011	0.007
	ϕ	100	0.135	0.109	0.117	0.119	0.141	0.116	0.110	0.077	0.132	0.102
		1000	0.044	0.030	0.033	0.033	0.063	0.035	0.029	0.023	0.053	0.046
		5000	0.020	0.014	0.015	0.015	0.029	0.015	0.013	0.012	0.019	0.014
	α_0	100	0.104	0.096	0.085	0.084	0.158	0.129	0.073	0.055	0.131	0.116
		1000	0.031	0.028	0.025	0.025	0.241	0.060	0.021	0.019	0.053	0.046
		5000	0.014	0.012	0.010	0.010	0.052	0.042	0.009	0.016	0.036	0.019
	α_1	100	0.179	0.177	0.166	0.167	0.469	0.385	0.134	0.107	0.494	0.405
		1000	0.064	0.060	0.045	0.045	0.328	0.161	0.031	0.046	0.160	0.137
		5000	0.031	0.027	0.020	0.020	0.182	0.096	0.013	0.038	0.090	0.062
	β	100	0.302	0.269	0.252	0.233	0.604	0.497	0.217	0.164	0.553	0.472
		1000	0.057	0.051	0.051	0.051	0.312	0.187	0.045	0.049	0.165	0.170
		5000	0.025	0.022	0.020	0.020	0.199	0.073	0.062	0.066	0.019	0.025
ARMA$(1,1)$	θ	100	0.110	0.086	0.096	0.101	0.112	0.090	0.097	0.068	0.125	0.091
-APARCH$(1,1)$		1000	0.029	0.021	0.023	0.024	0.031	0.021	0.022	0.014	0.033	0.024
		5000	0.013	0.008	0.010	0.010	0.014	0.009	0.010	0.006	0.015	0.011
	ϕ	100	0.138	0.114	0.121	0.126	0.128	0.107	0.111	0.086	0.146	0.107
		1000	0.040	0.027	0.032	0.032	0.041	0.028	0.029	0.026	0.043	0.030
		5000	0.018	0.012	0.012	0.012	0.020	0.012	0.013	0.014	0.019	0.013
	ω	100	0.198	0.192	0.199	0.210	0.254	0.262	0.221	0.170	0.290	0.272
		1000	0.079	0.067	0.056	0.056	0.226	0.218	0.044	0.045	0.142	0.129
		5000	0.033	0.028	0.025	0.025	0.209	0.207	0.017	0.029	0.061	0.056
	α	100	0.206	0.201	0.183	0.184	0.464	0.449	0.167	0.131	0.352	0.327
		1000	0.060	0.053	0.041	0.041	0.447	0.432	0.029	0.043	0.143	0.134
		5000	0.025	0.023	0.018	0.018	0.421	0.414	0.012	0.027	0.071	0.059
	γ	100	0.413	0.386	0.346	0.356	0.439	0.426	0.310	0.233	0.613	0.601
		1000	0.105	0.094	0.071	0.070	0.101	0.092	0.057	0.041	0.217	0.220
		5000	0.042	0.039	0.029	0.029	0.045	0.038	0.024	0.018	0.086	0.089
	β	100	0.297	0.282	0.255	0.238	0.312	0.288	0.186	0.145	0.476	0.468
		1000	0.074	0.067	0.058	0.058	0.074	0.066	0.043	0.033	0.151	0.150
		5000	0.033	0.031	0.024	0.024	0.034	0.029	0.018	0.012	0.061	0.063
	δ	100	0.732	0.732	0.712	0.717	0.740	0.739	0.703	0.613	0.830	0.815
		1000	0.402	0.352	0.296	0.296	0.394	0.338	0.235	0.290	0.542	0.534
		5000	0.170	0.147	0.132	0.131	0.169	0.145	0.092	0.168	0.251	0.262

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 $\widetilde{X}$ is also a stationary process (the coefficients of Λ_{β_0} are EDS). Moreover, it is well known that $(\psi_j(\beta_0))_{j\in\mathbb{N}}$ is EDS. Then we have:

$$\begin{aligned} \widetilde{X}_t &= \Lambda_{\beta_0}(L)\,\Big(M_{\theta_0}\big((X_{t-i})_{i\geq 1}\big)\,\zeta_t + f_{\theta_0}\big((X_{t-i})_{i\geq 1}\big)\Big) \\ \widetilde{X}_t + \sum_{j=1}^{\infty}\psi_j(\beta_0)\widetilde{X}_{t-j} &= M_{\theta_0}\big((\Lambda_{\beta_0}^{-1}(L)\widetilde{X}_{t-i})_{i\geq 1}\big)\,\zeta_t + f_{\theta_0}\big((\Lambda_{\beta_0}^{-1}(L)\widetilde{X}_{t-i})_{i\geq 1}\big) \\ \widetilde{X}_t &= \widetilde{M}_{\widetilde{\theta}_0}\big((\widetilde{X}_{t-i})_{i\geq 1}\big)\,\zeta_t + \widetilde{f}_{\widetilde{\theta}_0}\big((\widetilde{X}_{t-i})_{i\geq 1}\big) \\ &\text{with}\quad \begin{cases} \widetilde{M}_{\widetilde{\theta}_0}\big((x_{t-i})_{i\geq 1}\big) = M_{\widetilde{\theta}_0}\big((\Lambda_{\beta_0}^{-1}(L)x_{t-i})_{i\geq 1}\big) \\ \widetilde{f}_{\widetilde{\theta}_0}\big((x_{t-i})_{i\geq 1}\big) = f_{\theta_0}\big((\Lambda_{\beta_0}^{-1}(L)x_{t-i})_{i\geq 1}\big) \\ \qquad\qquad - \sum_{j=1}^{\infty}\psi_j(\beta_0)x_{t-j}. \end{cases} \end{aligned} \tag{5.1}$$

Finally, for $i=0$,

$$\begin{aligned} &\big|\widetilde{f}_{\widetilde{\theta}_0}\big((x_{t-i})_{i\geq 1}\big) - \widetilde{f}_{\widetilde{\theta}_0}\big((y_{t-i})_{i\geq 1}\big)\big| \leq \\ &\quad\leq \sum_{j=1}^{\infty}\alpha_j^{(0)}(f,\{\theta_0\})\,\big|(\Lambda_{\beta_0}^{-1}(L)x_{t-j-i})_{i\geq 1} - (\Lambda_{\beta_0}^{-1}(L)y_{t-j-i})_{i\geq 1}\big| \\ &\qquad + |\psi_j(\beta_0)|\,|x_{t-j}-y_{t-j}| \\ &\quad\leq \sum_{j=1}^{\infty}\alpha_j^{(0)}(f,\{\theta_0\})\,\Big|\sum_{k=0}^{\infty}|\psi_k(\beta_0)|\,|x_{t-k-j}-y_{t-k-j}\Big| + |\psi_j(\beta_0)|\,|x_{t-j}-y_{t-j}| \\ &\quad\leq \sum_{j=1}^{\infty}\Big(|\psi_j(\beta_0)| + \sum_{k=1}^{j}\alpha_k^{(0)}(f,\{\theta_0\})\psi_{j-k}(\beta_0)\Big)\,|x_{t-j}-y_{t-j}| \\ &\quad\Longrightarrow\quad \alpha_j^{(0)}(\widetilde{f},\{\widetilde{\theta}_0\}) \leq |\psi_j(\beta_0)| + \sum_{k=1}^{j}\alpha_k^{(0)}(f,\{\theta_0\})\big|\psi_{j-k}(\beta_0)\big|. \end{aligned} \tag{5.2}$$

Moreover, we also have:

$$\big|\widetilde{M}_{\widetilde{\theta}_0}\big((x_{t-i})_{i\geq 1}\big) - \widetilde{M}_{\widetilde{\theta}_0}\big((y_{t-i})_{i\geq 1}\big)\big|$$

$$\leq \sum_{j=1}^{\infty} \alpha_j^{(0)}(M,\{\theta_0\}) \big|(\Lambda_{\beta_0}^{-1}(L)x_{t-j-i})_{i\geq 1} - (\Lambda_{\beta_0}^{-1}(L)y_{t-j-i})_{i\geq 1}\big|$$

$$\Longrightarrow \quad \alpha_j^{(0)}(\widetilde{M},\{\widetilde{\theta}_0\}) \leq \sum_{k=1}^{j} \alpha_k^{(0)}(M,\{\theta_0\})\big|\psi_{j-k}\big|. \qquad (5.3)$$

The same kinds of computations could also be done by considering the first and second derivatives of $\widetilde{f}$ and $\widetilde{M}$ with respect to $\widetilde{\theta}$. Note, and this is important, that the first and second derivatives of Λ_β^{-1} with respect to $\widetilde{\theta}$ 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 $\sum_{k=1}^{j} \alpha_k^{(0)}(K,\{\theta_0\})\big|\psi_{j-k}\big| \leq C\sum_{k=1}^{j} k^{-\beta}\rho^{j-k} \sim -C(\log\rho)^{-1}j^{-\beta}$ and therefore $\alpha_j^{(0)}(\widetilde{K},\{\widetilde{\theta}_0\}) = O(j^{-\beta})$.
- 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 $\sum_{k=1}^{j} \alpha_k^{(0)}(K,\{\theta_0\})\big|\psi_{j-k}\big| \leq C\sum_{k=1}^{j} r^{-k}\rho^{j-k} = O(j\,\max(r,\rho)^j)$ and therefore $\alpha_j^{(O)}(\widetilde{K},\{\widetilde{\theta}_0\})$ is EDS.

The same kind of computation can be also done for $(\alpha_j^{(i)}(\widetilde{K},\{\widetilde{\theta}_0\}))_j$ since the derivatives and second-derivatives of $\Lambda_{\beta_0}^{-1}$ with respect to β and therefore to $\widetilde{\theta}$ are also EDS. □

Now we remind two lemmas already proved in [1]:

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 $K_\theta^t \in L^r(\mathcal{C}(\Theta, R^m))$ and there exists $C > 0$ not depending on t such that

$$\mathrm{E}\big[\|\widehat{K}_\theta^t - K_\theta^t\|_\Theta^r\big] \leq C\,\mathrm{E}\big[|X_0|^r\big]\Big(\sum_{j\geq t}\alpha_j(K,\Theta)\Big)^r \qquad \text{for all } t \in N^*. \qquad (5.4)$$

Lemma 5.2. Let $\mathcal{D}^{(2)}(\Theta)$ denote the Banach space of 2 times continuously differentiable functions on Θ equipped with the uniform norm

$$\|h\|_{2,\Theta} = \|h\|_\Theta + \Big\|\frac{\partial h}{\partial\theta}\Big\|_\Theta + \Big\|\frac{\partial^2 h}{\partial\theta\partial\theta'}\Big\|_\Theta.$$

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 $f_\theta^t \in \mathrm{L}^r\big(\mathcal{D}^{(2)}(\Theta)\big)$ and $M_\theta^t \in \mathrm{L}^r\big(\mathcal{D}^{(2)}(\Theta)\big)$.

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 [13]. In (i), a uniform (on Θ) strong law of large numbers satisfied by $\frac{1}{n}\widehat{L}_n(\theta)$ converging to $L(\theta) := -\mathrm{E}[q_0(\theta)]$ is established. In (ii), it is proved that $L(\theta)$ admits a unique maximum in θ_0. Those two conditions lead to the strong consistency of $\widehat{\theta}_n$ (from [13]).

(i) In the same way and for the same reason in the proof of Theorem 1 of [1], the uniform strong law of large numbers satisfied by the sample mean of $(\widehat{q}_t)_{t\in N^*}$ (defined in (2.3)]) is implied by establishing $\mathrm{E}[\|q_t(\theta)\|_\Theta] < \infty$. But new computations have to be done in case of Laplacian conditional log-density $q_t(\theta)$. From Lemma 5.1, for all $t \in Z$,

$$\begin{aligned}|q_t(\theta)| &= \big|(M_\theta^t)^{-1}|X_t - f_\theta^t| + \log(M_\theta^t)\big| \\ &\leq \frac{|X_t - f^t(\theta)|}{\underline{M}} + \big|\log(\underline{M})\big| + M_\theta^t \\ \Longrightarrow \qquad \sup_{\theta\in\Theta}|q_t(\theta)| &\leq \frac{1}{\underline{M}}\big(|X_t| + \|f^t(\theta)\|_\Theta\big) + \big|\log(\underline{M})\big| + \|M_\theta^t\|_\Theta.\end{aligned}$$

With $r \geq 1$, we have $\forall t \in \mathbb{Z},\ \mathrm{E}[|X_t|] < \infty$ from Proposition 2.1 and $\mathrm{E}\big[\|f_\theta^t\|_\Theta^r + \|M_\theta^t\|_\Theta^r\big] < \infty$ from Lemma 5.1, implying $\mathrm{E}\big[\|f_\theta^t\|_\Theta + \|M_\theta^t\|_\Theta\big] < \infty$. As a consequence, for all $t \in \mathbb{Z}$,

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

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

$$\Big\|\frac{L_n(\theta)}{n} - L(\theta)\Big\|_\Theta \xrightarrow[n\to\infty]{a.s.} 0. \tag{5.5}$$

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

$$\begin{aligned}|\widehat{q}_t(\theta) - q_t(\theta)| &\leq \big|\log(\widehat{M}_\theta^t) - \log(M_\theta^t) + (\widehat{M}_\theta^t)^{-1}|X_t - \widehat{f}_\theta^t| - (M_\theta^t)^{-1}|X_t - f_\theta^t|\big| \\ &\leq \big|\widehat{M}_\theta^t - {M_\theta}^t\big|\underline{M}^{-1} + |\widehat{M}_\theta^t - M_\theta^t|\underline{M}^{-2}|X_t - f_\theta^t| + \underline{M}^{-1}|\widehat{f}_\theta^t - f_\theta^t|\end{aligned}$$

with $C > 0$. Hence, we have:

$$\|\widehat{q}_t(\theta) - q_t(\theta)\|_\Theta \leq C\big(1 + |X_t| + \|f_\theta^t\|_\Theta\big)\big(\|\widehat{M}_\theta^t - M_\theta^t\|_\Theta + \|\widehat{f}_\theta^t - f_\theta^t\|_\Theta\big).$$

By Corollary 1 of [14], the proof is achieved if there exists $s \in (0,1]$ such as

$$\sum_{t\geq 1}\frac{1}{t^s}\,\mathrm{E}\big[\|q_t(\theta) - \widehat{q}_t(\theta)\|_\Theta^s\big] < \infty. \tag{5.6}$$

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 $A_0(M,\Theta)$,

$$\begin{aligned}\mathrm{E}\big[\|\widehat{q}_t(\theta) - q_t(\theta)\|_\Theta^{r/2}\big] \leq C\left(\mathrm{E}\big[(1 + |X_t| + \|f_\theta^t\|_\Theta)^r\big]\right)^{\frac{1}{2}}\Big(\mathrm{E}\big[(\|\widehat{M}_\theta^t - M_\theta^t\|_\Theta \\ + \|\widehat{f}_\theta^t - f_\theta^t\|_\Theta)^r\big]\Big)^{\frac{1}{2}}.\end{aligned}$$

Using Lemma 5.1 and previous proved results implying $\mathrm{E}[|X_t|^r] < \infty$, $\mathrm{E}[\|f_\theta^t\|_\Theta^r + \|M_\theta^t\|_\Theta^r] < \infty$, we obtain

$$\mathrm{E}\big[\|\widehat{q}_t(\theta) - q_t(\theta)\|_\Theta^{r/2}\big] \quad \leq \quad C\Big(\sum_{j>t}\alpha_j^{(0)}(f,\Theta) + \alpha_j^{(0)}(M,\Theta)\Big)^{\frac{r}{2}}$$

$$\leq \quad C\, t^{-\frac{(\ell-1)r}{2}},$$

where the last inequality is obtained from the condition (3.1) of Theorem 3.1.

Hence, we have

$$\sum_{t\geq 1} \frac{1}{t^{r/2}}\, \mathrm{E}\big[|\widehat{q}_t(\theta)-q_t(\theta)|_\Theta^{r/2}\big] \leq A \sum_{t\geq 1} t^{-r\,\ell/2},$$

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

$$\frac{1}{n}\sum_{t=1}^{n} \big\|\widehat{q}_t(\theta)-q_t(\theta)\big\|_\Theta \xrightarrow[n\to\infty]{a.s.} 0 \quad \text{and} \quad \frac{1}{n}\big\|\widehat{L}_n(\theta)-L_n(\theta)\big\|_\Theta \xrightarrow[n\to\infty]{a.s.} 0, \qquad (5.7)$$

and therefore, using (5.5),

$$\frac{1}{n}\big\|\widehat{L}_n(\theta)-L(\theta)\big\|_\Theta \xrightarrow[n\to\infty]{a.s.} 0. \qquad (5.8)$$

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

$$L(\theta) = -\mathrm{E}[q_0(\theta)].$$

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

$$\begin{aligned} L(\theta) &= -\mathrm{E}\big[\log(M_\theta^t) + (M_\theta^t)^{-1}\big|X_t - f_\theta^t\big|\big] \\ &= -\mathrm{E}\Big[\log(M_\theta^t) + \frac{M_{\theta_0}^t}{M_\theta^t}\Big|\zeta_t + \frac{f_{\theta_0}^t - f_\theta^t}{M_{\theta_0}^t}\Big|\Big]. \end{aligned}$$

Hence, using $\mathrm{E}[|\zeta_t|] = 1$, we obtain:

$$\begin{aligned} L(\theta_0) - L(\theta) &= \mathrm{E}\Big[\log\big(\frac{M_\theta^t}{M_{\theta_0}^t}\big) + \frac{M_{\theta_0}^t}{M_\theta^t}\Big|\zeta_t + \frac{f_{\theta_0}^t - f_\theta^t}{M_{\theta_0}^t}\Big| - 1\Big] \\ &= \mathrm{E}\Big[\log\big(\frac{M_\theta^t}{M_{\theta_0}^t}\big) - 1 + \frac{M_{\theta_0}^t}{M_\theta^t}\mathrm{E}\Big[\Big|\zeta_t + \frac{f_{\theta_0}^t - f_\theta^t}{M_{\theta_0}^t}\Big| \mid (X_{t-k})_{k\geq 1}\Big]\Big]. \end{aligned}$$

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

$$\begin{aligned} L(\theta_0) - L(\theta) &> \mathrm{E}\Big[\log\big(\frac{M_\theta^t}{M_{\theta_0}^t}\big) - 1 + \frac{M_{\theta_0}^t}{M_\theta^t}\Big] \\ &> h\Big(\frac{M_{\theta_0}^t}{M_\theta^t}\Big), \end{aligned}$$

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 $M_\theta \neq M_{\theta_0}$, $h\Big(\frac{M_{\theta_0}^t}{M_\theta^t}\Big) > 0$ (> 0 is replaced by $= 0$ if $M_\theta = M_{\theta_0}$). 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. □

Proof of Theorem 3.2. We follow the same scheme of proof than in [6]. Hence, denote

$$\begin{aligned} S_n &= \sum_{t=1}^n V_{t-1}\big(|Z_t - n^{-1/2}Y_{t-1}| - |Z_t|\big) \\ &= -n^{-1/2}\sum_{t=1}^n V_{t-1}Y_{t-1}\mathrm{sgn}(Z_t) \\ &\quad +2\sum_{t=1}^n V_{t-1}(n^{-1/2}Y_{t-1} - Z_t)\big(\mathbf{1}_{0<Z_t<n^{-1/2}Y_{t-1}} - \mathbf{1}_{n^{-1/2}Y_{t-1}<Z_t<0}\big) \\ &= A_n + B_n. \end{aligned}$$

Since $\mathrm{E}\big[V_{t-1}Y_{t-1}\mathrm{sgn}(Z_t) \mid \mathcal{F}_{t-1}\big] = \mathrm{E}[\mathrm{sgn}(Z_t)]\,\mathrm{E}\big[V_{t-1}Y_{t-1}\big] = 0$ and $\mathrm{E}\big[V_0^2Y_0^2\big] < \infty$, we can apply a central limit theorem for stationary martingale difference sequence (see [5]) and

$$A_n \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\big(0\,,\,\mathrm{E}\big[V_0^2Y_0^2\big]\big). \tag{5.9}$$

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 [6], we also obtain

- $\limsup_{n\to\infty} n\,\mathrm{E}\big[W_{nt}^2\big] = 0$;
- $\mathrm{E}\big[W_{nt} \mid \mathcal{F}_{t-1}\big] \simeq \dfrac{1}{2n}\,f(0)\,V_{t-1}Y_{t-1}^2 \quad \text{for } |n^{-1/2}Y_{t-1}| < \varepsilon$;
- $\sum_{t=1}^n W_{nt} \xrightarrow[n\to\infty]{\mathcal{P}} \dfrac{1}{2}\,f(0)\,\mathrm{E}\big[V_0Y_0^2\,\mathbf{1}_{Y_0>0}\big]$.

Then we deduce

$$B_n \xrightarrow[n\to\infty]{\mathcal{P}} f(0)\,\mathrm{E}\big[V_0Y_0^2\big]. \tag{5.10}$$

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 [6] or [15].

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

$$\begin{aligned} W_n(v) &= -\sum_{t=1}^n \big(q_t(\theta_0 + n^{-1/2}v) - q_t(\theta_0)\big) &(5.11) \\ &= \sum_{t=1}^n \log\Big(\frac{(M^t_{\theta_0+n^{-1/2}v})^{-1}}{(M^t_{\theta_0})^{-1}}\Big) + (M^t_{\theta_0})^{-1}|X_t - f^t_{\theta_0}| \\ &\qquad -(M^t_{\theta_0+n^{-1/2}v})^{-1}\big|X_t - f^t_{\theta_0+n^{-1/2}v}\big| &(5.12) \end{aligned}$$

with respect to v. As a consequence, there exists a sequence $(\widehat{v}_n)_n$ where $\widehat{v}_n$ is a maximizer of $W_n(v)$ such as $\widehat{v}_n = \sqrt{n}(\widehat{\theta}_n - \theta_0)$. 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 $\mathcal{C}(\mathbb{R}^d)$ (space of continuous functions on $\mathbb{R}^d$) to a limit process W. Hence $(\widehat{v}_n)_n$ converges to the maximizer of W.

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

$$\begin{aligned} W_n(v) &= \sum_{t=1}^n \log\Big(\frac{(M^t_{\theta_0+n^{-1/2}v})^{-1}}{(M^t_{\theta_0})^{-1}}\Big) + |X_t - f^t_{\theta_0}|\big((M^t_{\theta_0})^{-1} - (M^t_{\theta_0+n^{-1/2}v})^{-1}\big) \\ &\quad + \sum_{t=1}^n (M^t_{\theta_0+n^{-1/2}v})^{-1}\big(|X_t - f^t_{\theta_0}| - \big|X_t - f^t_{\theta_0+n^{-1/2}v}\big|\big) \\ &= I_1(v) + I_2(v). \end{aligned}$$

We have:

$$I_1(v) \quad = \quad -\sum_{t=1}^n \log\Big(\frac{M^t_{\theta_0+n^{-1/2}v}}{M^t_{\theta_0}}\Big) + |\zeta_t|\Big(1 - \frac{M^t_{\theta_0}}{M^t_{\theta_0+n^{-1/2}v}}\Big)$$

Using Taylor expansions, we deduce that for each $t \in \{1, \cdots, n\}$, there exists $\overline{\theta}^t_1$ and $\overline{\theta}^t_2$ in the segment $[\theta_0, \theta_0 + n^{-1/2}v]$ such as:

$$\begin{aligned} \log\Big(\frac{M^t_{\theta_0+n^{-1/2}v}}{M^t_{\theta_0}}\Big) &= n^{-1/2}(M^t_{\theta_0})^{-1} v' \Big(\frac{\partial M^t_\theta}{\partial\theta}\Big)_{\theta_0} \\ &\quad + \frac{1}{2}\, n^{-1}\Big\{v'\Big(\frac{\partial^2 M^t_\theta}{\partial\theta^2}\Big)_{\overline{\theta}^t_1} - (M^t_{\theta_0})^{-2}\Big(v'\Big(\frac{\partial M^t_\theta}{\partial\theta}\Big)_{\overline{\theta}^t_1}\Big)^2\Big\} \\ \frac{M^t_{\theta_0}}{M^t_{\theta_0+n^{-1/2}v}} &= 1 - n^{-1/2}(M^t_{\theta_0})^{-1} v'\Big(\frac{\partial M^t_\theta}{\partial\theta}\Big)_{\theta_0} \\ &\quad + \frac{1}{2}\, n^{-1}\Big\{2(M^t_{\theta_0})^{-2}\Big(v'\Big(\frac{\partial M^t_\theta}{\partial\theta}\Big)_{\overline{\theta}^t_2}\Big)^2 - v'\Big(\frac{\partial^2 M^t_\theta}{\partial\theta^2}\Big)_{\overline{\theta}^t_2}\Big\} \end{aligned}$$

Then,

$$\begin{aligned} I_1(v) &= n^{-1/2}\sum_{t=1}^n (M^t_{\theta_0})^{-1} v'\Big(\frac{\partial M^t_\theta}{\partial\theta}\Big)_{\theta_0}\big(|\zeta_t| - 1\big) \\ &\quad + \frac{1}{2n}\sum_{t=1}^n (M^t_{\theta_0})^{-1}\Big\{v'\Big(\frac{\partial^2 M^t_\theta}{\partial\theta^2}\Big)_{\overline{\theta}^t_2} v\,|\zeta_t| - v'\Big(\frac{\partial^2 M^t_\theta}{\partial\theta^2}\Big)_{\overline{\theta}^t_1} v\Big\} \\ &\quad + \frac{1}{2n}\sum_{t=1}^n (M^t_{\theta_0})^{-2}\Big\{\Big(v'\Big(\frac{\partial M^t_\theta}{\partial\theta}\Big)_{\overline{\theta}^t_1}\Big)^2 - 2\Big(v'\Big(\frac{\partial M^t_\theta}{\partial\theta}\Big)_{\overline{\theta}^t_2}\Big)^2|\zeta_t|\Big\} \\ &= I^{(1)}_1(v) + I^{(2)}_1(v) + I^{(3)}_1(v). \end{aligned}$$

Using a Central Limit Theorem for martingale-differences (see for instance

[5]), and since from Lemma 5.2 we have $\mathrm{E}\big[\big\|(M^t_{\theta_0})^{-1}v'\left(\frac{\partial M^t_\theta}{\partial\theta}\right)_{\theta_0}\big\|^2_\Theta\big]<\infty$ and $\mathrm{E}\big[(M^t_{\theta_0})^{-1}v'\left(\frac{\partial M^t_\theta}{\partial\theta}\right)_{\theta_0}\big(|\zeta_t|-1\big)\mid\mathcal{F}_{t-1}\big]=0$, then

$$I_1^{(1)}(v)\ \xrightarrow[n\to\infty]{\mathcal{D}}\ \mathcal{N}\Big(0\,,\,\mathrm{E}\Big[(M^0_{\theta_0})^{-2}\Big(v'\left(\frac{\partial M^0_\theta}{\partial\theta}\right)_{\theta_0}\Big)^2\Big]\big(\sigma^2_\zeta-1\big)\Big). \tag{5.13}$$

Now, using that $\theta\in\Theta\mapsto\frac{\partial M^t_\theta}{\partial\theta}$ and $\theta\mapsto\frac{\partial^2 M^t_\theta}{\partial\theta^2}$ are continuous functions, $\overline{\theta}^t_1\ \xrightarrow[n\to\infty]{\mathcal{D}}\ \theta_0$ and $\overline{\theta}^t_2\ \xrightarrow[n\to\infty]{\mathcal{D}}\ \theta_0$, we claim that $I_1^{(2)}(v)$ have the same limit distribution that $\frac{1}{2n}\sum_{t=1}^n(M^t_{\theta_0})^{-1}v'\left(\frac{\partial^2 M^t_\theta}{\partial\theta^2}\right)_{\theta_0}v\left(|\zeta_t|-1\right)$. From Lemma 5.2, note that $\mathrm{E}\Big[(M^t_{\theta_0})^{-1}v'\left(\frac{\partial^2 M^t_\theta}{\partial\theta^2}\right)_{\theta_0}v\left(|\zeta_t|-1\right)\mid\mathcal{F}_{t-1}\Big]=0$ and $\mathrm{E}\big[\big\|(M^t_{\theta_0})^{-1}v'\left(\frac{\partial^2 M^t_\theta}{\partial\theta^2}\right)_{\theta_0}v\big\|\big]<\infty$.

Thus, from the strong large number law for martingale-differences (see again [5]), we obtain:

$$\frac{1}{2n}\sum_{t=1}^n(M^t_{\theta_0})^{-1}v'\left(\frac{\partial^2 M^t_\theta}{\partial\theta^2}\right)_{\theta_0}v\left(|\zeta_t|-1\right)\ \xrightarrow[n\to\infty]{a.s.}\ 0,$$

and this implies:

$$I_1^{(2)}(v)\ \xrightarrow[n\to\infty]{\mathcal{D}}\ 0. \tag{5.14}$$

Previous arguments induce that $I_1^{(3)}(v)$ has the same limit distribution that $\displaystyle\frac{1}{2n}\sum_{t=1}^n(M^t_{\theta_0})^{-2}\Big(v'\left(\frac{\partial M^t_\theta}{\partial\theta}\right)_{\theta_0}\Big)^2\big(1-2|\zeta_t|\big)$. From the strong large number law for martingale-differences (see [5]), we obtain:

$$\begin{aligned}\frac{1}{2n}\sum_{t=1}^n(M^t_{\theta_0})^{-2}&\Big(v'\left(\frac{\partial M^t_\theta}{\partial\theta}\right)_{\theta_0}\Big)^2\left(1-2|\zeta_t|\right)\\ &\xrightarrow[n\to\infty]{a.s.}\ \frac{1}{2}\,\mathrm{E}\Big[(M^0_{\theta_0})^{-2}\Big(v'\left(\frac{\partial M^0_\theta}{\partial\theta}\right)_{\theta_0}\Big)^2\left(1-2|\zeta_0|\right)\Big]\\ &\xrightarrow[n\to\infty]{a.s.}\ -\frac{1}{2}\,\mathrm{E}\Big[(M^0_{\theta_0})^{-2}\Big(v'\left(\frac{\partial M^0_\theta}{\partial\theta}\right)_{\theta_0}\Big)^2\Big],\end{aligned}$$

and this implies:

$$I_1^{(3)}(v)\ \xrightarrow[n\to\infty]{\mathcal{D}}\ -\frac{1}{2}\,\mathrm{E}\Big[(M^0_{\theta_0})^{-2}\Big(v'\left(\frac{\partial M^0_\theta}{\partial\theta}\right)_{\theta_0}\Big)^2\Big]. \tag{5.15}$$

Finally, from (5.13), (5.14) and (5.15), we obtain:

$$\begin{aligned}I_1(v)\ \xrightarrow[n\to\infty]{\mathcal{D}}\ &\mathcal{N}\Big(-\frac{1}{2}\,\mathrm{E}\Big[(M^0_{\theta_0})^{-2}\Big(v'\left(\frac{\partial M^0_\theta}{\partial\theta}\right)_{\theta_0}\Big)^2\Big],\\ &\mathrm{E}\Big[(M^0_{\theta_0})^{-2}\Big(v'\left(\frac{\partial M^0_\theta}{\partial\theta}\right)_{\theta_0}\Big)^2\Big]\big(\sigma^2_\zeta-1\big)\Big).\end{aligned} \tag{5.16}$$

Now we consider $I_2(v) = \sum_{t=1}^{n} (M^t_{\theta_0+n^{-1/2}v})^{-1}\big(|X_t - f^t_{\theta_0}| - \big|X_t - f^t_{\theta_0+n^{-1/2}v}\big|\big)$. Using again Taylor expansion, we can write:

$$\begin{aligned}
I_2(v) &= \sum_{t=1}^{n} (M^t_{\theta_0})^{-1}\big(|X_t - f^t_{\theta_0}| - \big|X_t - f^t_{\theta_0+n^{-1/2}v}\big|\big) \\
&\quad - n^{-1/2}\sum_{t=1}^{n} (M^t_{\overline{\theta}^t_M})^{-2}\, v' \Big(\frac{\partial M^t_\theta}{\partial \theta}\Big)_{\overline{\theta}^t_M} \big(|X_t - f^t_{\theta_0}| - \big|X_t - f^t_{\theta_0+n^{-1/2}v}\big|\big) \\
&= I_2^{(1)}(v) + I_2^{(2)}(v),
\end{aligned}$$

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

First we have:

$$I_2^{(1)}(v) = \sum_{t=1}^{n} \Big(|\zeta_t| - \Big|\zeta_t - n^{-1/2}(M^t_{\theta_0})^{-1} v' \big(\frac{\partial f^t_\theta}{\partial \theta}\big)_{\overline{\theta}^t_f}\Big|\Big)$$

with $\overline{\theta}^t_f$ in the segment $[\theta_0, \theta_0+n^{-1/2}v]$. Using Theorem 3.2, which an extension of Theorem 1 established in [6], denoting $Z_t = \zeta_t$, $Y_t = (M^t_{\theta_0})^{-1} v' \big(\frac{\partial f^t_\theta}{\partial \theta}\big)_{\overline{\theta}^t_f}$ and $V_t = 1$ for $t \in \mathbb{Z}$,

$$I_2^{(1)} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\Big(- g(0)\, \mathrm{E}\big[(M^0_{\theta_0})^{-2}\big(v' \big(\frac{\partial f^0_\theta}{\partial \theta}\big)_{\theta_0}\big)^2\big] \, , \, \mathrm{E}\big[(M^0_{\theta_0})^{-2}\big(v' \big(\frac{\partial f^0_\theta}{\partial \theta}\big)_{\theta_0}\big)^2\big] \Big). \tag{5.17}$$

since $\mathrm{E}\big[Y_t^2 V_t^2\big] \leq \underline{M}^{-2}\mathrm{E}\big[\big\|v' \frac{\partial f^t_\theta}{\partial \theta}\big\|^2_\Theta\big] < \infty$ from Lemma 5.2.

Then, we have:

$$\begin{aligned}
I_2^{(2)}(v) &\sim n^{-1/2}\sum_{t=1}^{n} (M^t_{\theta_0})^{-2}\, v' \Big(\frac{\partial M^t_\theta}{\partial \theta}\Big)_{\theta_0} \big(|X_t - f^t_{\theta_0}| - \big|X_t - f^t_{\theta_0+n^{-1/2}v}\big|\big) \\
&\sim n^{-1/2}\sum_{t=1}^{n} (M^t_{\theta_0})^{-1}\, v' \Big(\frac{\partial M^t_\theta}{\partial \theta}\Big)_{\theta_0} \Big(|\zeta_t| - \Big|\zeta_t - n^{-1/2}(M^t_{\theta_0})^{-1} v' \big(\frac{\partial f^t_\theta}{\partial \theta}\big)_{\overline{\theta}^t_f}\Big|\Big) \\
&\xrightarrow[n\to\infty]{\mathcal{P}} 0,
\end{aligned} \tag{5.18}$$

using the proof of Theorem 3.2 and denoting $Z_t = \zeta_t$, $Y_t = (M^t_{\theta_0})^{-1} v' \big(\frac{\partial f^t_\theta}{\partial \theta}\big)_{\overline{\theta}^t_f}$ and $V_t = (M^t_{\theta_0})^{-1}\, v' \Big(\frac{\partial M^t_\theta}{\partial \theta}\Big)_{\theta_0}$ for $t \in \mathbb{Z}$ and condition $\mathrm{E}\big[|V_t Y_t|\big] < \infty$ insuring a strong law of large number instead of central limit theorem for a martingale difference process. Therefore, from (5.17) and (5.18), we deduce

$$I_2(v) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\Big(- g(0)\, \mathrm{E}\big[(M^0_{\theta_0})^{-2}\big(v' \big(\frac{\partial f^0_\theta}{\partial \theta}\big)_{\theta_0}\big)^2\big] \, , \, \mathrm{E}\big[(M^0_{\theta_0})^{-2}\big(v' \big(\frac{\partial f^0_\theta}{\partial \theta}\big)_{\theta_0}\big)^2\big] \Big). \tag{5.19}$$

Finally, we obtain the behavior of $W_n(v)$ defined in (5.11) from (5.16) and (5.19). However, we have to specify the asymptotic dependency relation between $I_1^{(1)}$

and $I_2^{(1)}$. 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:

$$n^{-1/2}\sum_{t=1}^{n}(M_{\theta_0}^t)^{-1}v'\Big(\frac{\partial M_\theta^t}{\partial\theta}\Big)_{\theta_0}\big(|\zeta_t|-1\big)+n^{-1/2}\sum_{t=1}^{n}(M_{\theta_0}^t)^{-1}v'\,\big(\frac{\partial f_\theta^t}{\partial\theta}\big)_{\theta_0}\operatorname{sgn}(\zeta_t),$$

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

$$\begin{aligned}
&\mathrm{E}\Big[(M_{\theta_0}^t)^{-2}\Big(v'\Big(\frac{\partial M_\theta^t}{\partial\theta}\Big)_{\theta_0}\big(|\zeta_t|-1\big)+v'\Big(\frac{\partial f_\theta^t}{\partial\theta}\Big)_{\theta_0}\operatorname{sgn}(\zeta_t)\Big)^2\Big]\\
&\quad=\mathrm{E}\Big[(M_{\theta_0}^0)^{-2}\Big(v'\Big(\frac{\partial M_\theta^0}{\partial\theta}\Big)_{\theta_0}\Big)^2\Big](\sigma_\zeta^2-1)\\
&\qquad+2\,\mathrm{E}\Big[(M_{\theta_0}^0)^{-2}v'\Big(\frac{\partial M_\theta^t}{\partial\theta}\Big)_{\theta_0}v'\Big(\frac{\partial f_\theta^0}{\partial\theta}\Big)_{\theta_0}\Big]\,\mathrm{E}\big[\big(|\zeta_t|-1\big)\operatorname{sgn}(\zeta_t)\big]\\
&\qquad+\mathrm{E}\Big[(M_{\theta_0}^0)^{-2}\Big(v'\Big(\frac{\partial f_\theta^0}{\partial\theta}\Big)_{\theta_0}\Big)^2\Big]\\
&\quad=\mathrm{E}\Big[(M_{\theta_0}^0)^{-2}\Big\{\big(\sigma_\zeta^2-1\big)\Big(v'\Big(\frac{\partial M_\theta^0}{\partial\theta}\Big)_{\theta_0}\Big)^2+\Big(v'\Big(\frac{\partial f_\theta^0}{\partial\theta}\Big)_{\theta_0}\Big)^2\Big\}\Big] \qquad (5.20)
\end{aligned}$$

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

$$W_n(v)\ \xrightarrow[n\to\infty]{\mathcal{D}}\ W(v)=v'\big(-\frac{1}{2}\,\Gamma_M-g(0)\,\Gamma_F\big)\,v+v'\,N \qquad (5.21)$$

$$\text{with}\quad\begin{cases}N\ \stackrel{\mathcal{L}}{=}\ \mathcal{N}\big(0,\ \big(\big(\sigma_\zeta^2-1\big)\,\Gamma_M+\Gamma_F\big)\big)\\ \Gamma_F=\Big(\mathrm{E}\Big[(M_{\theta_0}^0)^{-2}\Big(\frac{\partial f_\theta^0}{\partial\theta_i}\Big)_{\theta_0}\Big(\frac{\partial f_\theta^0}{\partial\theta_j}\Big)_{\theta_0}\Big]\Big)_{1\le i,j\le d}\\ \Gamma_M=\Big(\mathrm{E}\Big[\Big(\frac{\partial\log(M_\theta^0)}{\partial\theta_i}\Big)_{\theta_0}\Big(\frac{\partial\log(M_\theta^0)}{\partial\theta_j}\Big)_{\theta_0}\Big]\Big)_{1\le i,j\le d}\end{cases} \qquad (5.22)$$

2/ Now, we consider the approximation $\widehat{W}_n(v)$ of $W_n(v)$ defined by:

$$\widehat{W}_n(v)=-\sum_{t=1}^{n}\big(\widehat{q}_t(\theta_0+n^{-1/2}v)-\widehat{q}_t(\theta_0)\big)\qquad\text{for any } v\in\mathbb{R}^d.$$

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

$$\widehat{W}_n(v)\ \xrightarrow[n\to\infty]{\mathcal{D}}\ W(v) \qquad (5.23)$$

with W defined in (5.21).

3/ Now, from (5.23), the proof of Theorem 3.2 and the same arguments than in the proof of Theorem 2 of [6], we deduce that finite distributions $(\widehat{W}_n(v_1),\cdots,\widehat{W}_n(v_k))$ 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, $(W_n(v))_v$ converges to $(W(v))_v$ as a process on the continuous function space $\mathcal{C}^0$.

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

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

with N defined in (5.22) and this implies (3.3). □

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 proof that $\big(\delta,b_0,(b_i^+(\theta),b_i^-(\theta))_{i\geq1}\big)$ (defined in (3.5)) are unique, thereafter in (ii) we are going to prove that $\theta=\big(\omega,(\alpha_i)_{1\leq i\leq p},(\gamma_i)_{1\leq i\leq p},(\beta_i)_{1\leq i\leq q}\big)$ is also unique.

(i) The proof of this result follow the same reasoning in [3]. First we have

$$\sigma_t^\delta=b_0(\theta)+\sum_{i\geq1}b_i^+(\theta)(\max(X_{t-i},0))^\delta+\sum_{i\geq1}b_i^-(\theta)(\max(-X_{t-i},0))^\delta. \quad (5.24)$$

We prove the result by contradiction. Suppose that there exist two vectors $\beta=\big(\delta,b_0,(b_i^+)_{i\geq1},(b_i^-)_{i\geq1}\big)$ and $\beta'=\big(\delta',b_0',(b_i^{+'})_{i\geq1},(b_i^{-'})_{i\geq1}\big)$ verifying (5.24). Let $m>0$ be the smallest integer satisfying $b_m^+\neq b_m^{+'}$ or $b_m^-\neq b_m^{-'}$ (if $b_i^+=b_i^{+'}$ and $b_i^-=b_i^{-'}$ $\forall i\geq1$ then $b_0=b_0'$). In one hand, since $x\in(0,\infty)\mapsto x^\delta$ is a one-to-one map and since $\mathrm{P}(X_t=\pm1,\ \forall t\in\mathbb{Z})=0$, we have $\delta=\delta'$. In the other hand, by definition of m, we have

$$\begin{aligned}(b_m^{+'}-b_m^+)(\max(X_{t-i},0))^\delta+(b_m^{-'}-b_m^-)(\max(-X_{t-m},0))^\delta&\\ =b_0-b_0'+\sum_{i\geq m+1}(b_i^+-b_i^{+'})(\max(X_{t-i},0))^\delta&\\ +\sum_{i\geq m+1}(b_i^--b_i^{-'})(\max(-X_{t-i},0))^\delta.&\quad (5.25)\end{aligned}$$

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

$$\begin{aligned}(b_m^{+'}-b_m^+)(\max(X_{t-m},0))^\delta+(b_m^{-'}-b_m^-)(\max(-X_{t-m},0))^\delta&\\ =\begin{cases}(b_m^{+'}-b_m^+)\sigma_{t-m}^\delta\zeta_{t-m}^\delta & \text{when } \zeta_{t-m}\geq0\\ (b_m^{-'}-b_m^-)\sigma_{t-m}^\delta(-\zeta_{t-m})^\delta & \text{when } \zeta_{t-m}<0\end{cases}&\end{aligned}$$

Moreover (5.25) and the fact that $b_m^+\neq b_m^{+'}$ or $b_m^-\neq b_m^{-'}$ implies that at least one of the following equalities hold

$$\begin{cases} \zeta_{t-m}^{\delta} = ((b_m^{+'} - b_m^{+})\sigma_{t-m}^{\delta})^{-1}\Big(\sum_{i\geq m+1}(b_i^{+} - b_i^{+'})(\max(X_{t-i},0))^{\delta}\Big) \\ \hfill \text{when } \zeta_{t-m} \geq 0 \\ \text{or} \\ (-\zeta_{t-m})^{\delta} = ((b_m^{-'} - b_m^{-})\sigma_{t-m}^{\delta})^{-1}\Big(\sum_{i\geq m+1}(b_i^{-} - b_i^{-'})(\max(X_{t-i},0))^{\delta}\Big) \\ \hfill \text{when } \zeta_{t-m} < 0 \end{cases}$$

Since $\sigma_{t-m}^{\delta} > b_0 > 0$, ζ_{t-m}^{δ} is well defined. Let F_k be the F-algebra generated by $(\zeta_i, i < k)$. The causal representation of tha APARCH(δ, p, q) shows that X_j is F_j-measurable and thus the right-hand side of the above equations (and consequently also ζ_{t-m}^{δ} 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 F_{t-m-1}. Since (ζ_j) is a sequence of independent random variables, this implies that ζ_{t-m} is a.s. constant when $\zeta_{t-m} \geq 0$ or when $\zeta_{t-m} < 0$, contradicting the hypothesis saying ζ_0^{δ} has a non-degenerate distribution. This achieves (i).

(ii) The representation (5.24) is the same as

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

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

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

$$\begin{cases} \Upsilon_{\theta_1}^{-1}\Delta_{\theta_2}^{+} = \Upsilon_{\theta_1'}^{-1}\Delta_{\theta_2'}^{+} \\ \\ \Upsilon_{\theta_1}^{-1}\Delta_{\theta_2}^{-} = \Upsilon_{\theta_1'}^{-1}\Delta_{\theta_2'}^{-} \end{cases} \Rightarrow \begin{cases} \Delta_{\theta_2}^{+} = \big(\Upsilon_{\theta_1}\Upsilon_{\theta_1'}^{-1}\big)\Delta_{\theta_2'}^{+} \\ \\ \Delta_{\theta_2}^{-} = \big(\Upsilon_{\theta_1}\Upsilon_{\theta_1'}^{-1}\big)\Delta_{\theta_2'}^{-} \end{cases}$$

from the first equality, since $\deg(\Delta_{\theta_2}^{+}) = \deg(\Delta_{\theta_2'}^{+}) = q$, we conclude that $\Upsilon_{\theta_1}\Upsilon_{\theta_1'}^{-1} = 1$, therefore $\Upsilon_{\theta_1} = \Upsilon_{\theta_1'}$ 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 $\Upsilon_{\theta_1} = \Upsilon_{\theta_1'}$ implies that $\beta_i = \beta_i'$.
- Since $(\beta_i)_{i=1,p}, b_0 = w(1 - \sum_{j=1}^{p}\beta_j)^{-1}$ are unique then ω is unique.

Thus, Condition **C3** is established and the proof of proposition is achieved. □

Proof of Proposition 3.2. Since we prove that Lemma 2.1 implies that conditions on Liptshitzian coefficients $(\alpha_j^{(i)}(f,\Theta))_j$ and $(\alpha_j^{(i)}(M,\Theta))_j$, 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 $f_\theta = f_{\theta'}$ 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. □

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

[1] Bardet, J.-M. and Wintenberger, O. (2009) Asymptotic normality of the Quasi-Maximum likelihood estimator for multidimensional causal process, *Ann. Statist.*, **37**, 2730–2759. MR2541445

[2] Bardet, J.-M., Kengne, W. and Wintenberger, O. (2012) Detecting multiple change-points in general causal time series using penalized quasi-likelihood. *Electronic Journal of Statistics 6*, **2012**, 435–477. MR2988415

[3] Berkes, I., Horváth, L. and Kokoszka, P. (2003), GARCH processes: structure and estimation, *Bernoulli*, **9**, 201–227. MR1997027

[4] Berkes, I. and Horváth, L. (2004) The efficiency of the estimators of the parameters in GARCH processes, *Ann. Statist.*, **32**, 633–655. MR2060172

[5] Billingsley, P. (1968) *Convergence of Probability Measures.* Wiley, New York.

[6] Davis, R. and Dunsmuir, W. (1997) Least absolute deviation estimation for regression with ARMA errors. *Journal of Theoretical Probability*, **10**, 481–497. MR1455154

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

[8] Doukhan, P. and Wintenberger, O. (2007) Weakly dependent chains with infinite memory. *Stochastic Process. Appl.*, **118**, 1997–2013. MR2462284

[9] Francq, C. and Zakoian, J.-M. (2004) Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes, *Bernoulli*, **10**, 605–637. MR2076065

[10] Francq, C., Lepage, G. and Zakoian, J.-M. (2011) Two-stage non Gaussian QML estimation of GARCH models and testing the efficiency of the Gaussian QMLE. *Journal of Econometrics*, **165**, 246–257.

[11] Francq, C. and Zakoian, J.-M. (2013) Optimal predictions of powers of conditionally heteroskedastic processes. *Journal of the Royal Statistical Society B*, **75**, 345–367.

[12] Francq, C. and Zakoian, J.-M. (2015) Risk-parameter estimation in volatility models. *Journal of Econometrics*, **184**, 158–173.

[13] Jeantheau, T. (1998). Strong consistency of estimators for multivariate arch models. *Econometric Theory*, **14**, 70–86. MR1613694

[14] Kounias, E.G. and Weng, T.-S. (1969) An inequality and almost sure convergence. *Annals of Mathematical Statistics*, **40**, 1091–1093.

[15] Li, G. and Li, W.K. (2008) Least absolute deviation estimation for fractionally integrated autoregressive moving average time series models with conditional heteroscedasticity. *Biometrika*, **95**, 399–414. MR2521590

[16] Ling, S. and McAleer, M. (2003). Asymptotic theory for a vector ARMA-GARCH model. *Econometric Theory*, **19**, 280–310. MR1966031

[17] Lumsdaine, R.L. (1996) Consistency and asymptotic normality of the quasi-maximum likelihood in IGARCH$(1,1)$ and covariance stationary GARCH$(1,1)$ models. *Econometrica*, **64**, 575–596.

[18] Newey, W.K. and Steigerwald, D.G. (1997). Asymptotic bias for quasi-maximum-likelihood estimators in conditional heteroskedasticity models. *Econometrica*, **65**, 587–599.

[19] Peng, L. and Yao, Q. (2003). Least absolute deviations estimation for ARCH and GARCH models. *Biometrika*, **90**, 967–975.

[20] Robinson, P.M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression, *Journal of Econometrics*, **47**, 67–84.

[21] Robinson, P.M. and Zaffaroni, P. (2006) Pseudo-maximum likelihood estimation of ARCH(∞) models, *Ann. Statist.*, **34**, 1049–1074.

[22] Straumann, D. and Mikosch, T. (2006). Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. *Ann. Statist.*, **34**, 2449–2495.

[23] van der Vaart, A.W. (2000) *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.

[24] Weiss, A.A. (1986) Asymptotic theory for ARCH models estimation and testing. *Econometric Theory*, **2**, 107–131.