QML estimation of the continuously invertible EGARCH(1,1) model

Olivier Wintenberger

QML ESTIMATION OF THE CONTINUOUSLY INVERTIBLE EGARCH(1,1) MODEL

OLIVIER WINTENBERGER

Abstract. We introduce the notion of continuous invertibility on a compact set for volatility models driven by a Stochastic Recurrence Equation (SRE) introduced by Straumann in [30]. We prove the strong consistency of the Quasi Maximum Likelihood Estimator (QMLE) for such continuously invertible models. This approach gives for the first time the strong consistency of the QMLE used by Nelson in [27] for the EGARCH(1,1) model. We also give sufficient conditions for the asymptotic normality of the QMLE for this model. We propose a new estimator IQMLE with the same asymptotic properties than QMLE but for which the volatility forecasting is stable.

AMS 2000 subject classifications: Primary 62F12; Secondary 60H25, 62F10, 62M20, 62M10, 91B84.

Keywords and phrases: Invertibility, volatility models, parametric estimation, strong consistency, asymptotic normality, asymmetric GARCH, exponential GARCH, stochastic recurrence equation, stationarity

1. Introduction

Since the seminal papers [14, 7], the General Autoregressive Conditional Heteroskedasticity (GARCH) type models have been successfully applied to financial time series modeling. One of the stylized facts observed on the data is the asymmetry with respect to (wrt) shocks [11]: a negative past observation impacts the present volatility more importantly than a positive one. Nelson is the first to introduce in [27] a model that reproduce this asymmetry with the Exponential-GARCH (EGARCH) model. Not surprisingly, theoretical investigations of EGARCH has attracted lot of attention since then, see for example [20, 25]. However, the properties of the Quasi Maximum Likelihood Estimator (QMLE) used empirically in [27] was not proved except in some degenerate case, see [31]. We give in this paper some sufficient conditions for the strong consistency and the asymptotic normality of the QMLE in the EGARCH(1,1) model. We give sufficient conditions for the estimator to be strongly consistent and asymptotically normal. Our approach is based on the natural notion of continuous invertibility that we introduce in the general setting of volatility model solutions of a Stochastic Recurrent Equation (SRE) as studied in [30, 31].

Consider a real valued volatility model of the form $X_t = \sigma_t Z_t$ where σ_t is the volatility and where the innovations Z_t are normalized, centered independent identical distributed (iid) random vectors. It is assumed that a transformation of the volatility satisfies some parametric SRE (also called Iterated Random Function), i.e. there exist a function h and some ψ_t measurable wrt Z_t such that the following relations

$$h(\sigma_{t+1}^2) = \psi_t(h(\sigma_t^2), \theta_0), \qquad \forall t \in \mathbb{Z} \tag{1}$$

hold. Classical examples are the GARCH(1,1) and EGARCH(1,1) models:

$$\text{GARCH(1,1):} \qquad \sigma_{t+1}^2 = \alpha_0 + \beta_0 \sigma_t^2 + \gamma_0 X_t^2, \tag{2}$$

$$\text{EGARCH(1,1):} \qquad \log(\sigma_{t+1}^2) = \alpha_0 + \beta_0 \log(\sigma_t^2) + (\gamma_0 Z_t + \delta_0 |Z_t|). \tag{3}$$

To assert the existence of a stationary solution of both models, we rewrite (2) as an SRE in term of the innovations: $\sigma_{t+1}^2 = \alpha_0 + (\beta_0 + \gamma_0 Z_t^2)\sigma_t^2$, i.e. $\psi_t(x,\theta) = \alpha + (\beta + \gamma Z_t^2)x$. This SRE yields the Lyapunov condition $\mathbb{E}[\log(\beta_0 + \gamma_0 Z_0^2)] < 0$, necessary and sufficient for the stationarity, see [26]. Such a SRE is said to be convergent when its solution is unique, non anticipative (i.e. the function of the innovations $h(\sigma_t^2)$ are function of the past innovations $Z_{t-1}, Z_{t-2}, Z_{t-3}, \dots$ at any time t) and its law is the asymptotic law of the solution of the SRE

$$h_{t+1} = \psi_t(h_t, \theta_0) \qquad \forall t \geq 0 \tag{4}$$

for any initial value h_0. This last property is also called the stability of the SRE and is related with the ergodicity for Markov chains. Sufficient conditions for the existence of a stationary solution is the negativity of the top Lyapunov coefficient [13]. If so, the SRE is stable [8]. If the innovations were observed, the stability of the SRE ensures that the volatility can be efficiently forecast by using recursively the relation (4) from an arbitrary initial value h_0.

In practice, the innovations Z_t are not observed. Writing $Z_t = X_t/\sigma_t$ in the expression of ψ_t wrt Z_t, we invert the model, i.e we consider a new SRE driven by a function ϕ_t of the observation X_t,

$$h(\sigma_{t+1}^2) = \phi_t(h(\sigma_t^2), \theta_0). \tag{5}$$

For instance, we obtain from (2) that $\phi_t(x,\theta) = \alpha + \beta x + \gamma X_t^2$ for the GARCH(1,1) model. For the EGARCH(1,1) model, we obtain from (3) that

$$\phi_t(x,\theta) = \alpha + \beta x + (\gamma X_t + \delta |X_t|)\exp(-x/2). \tag{6}$$

In accordance with the notions of invertibility given previously in [18, 33, 30, 31], we will say that the model is invertible if the SRE (5) is stable. Then, as the functions ϕ_t are observed, the volatility is efficiently forecast by using recursively the relation

$$h_{t+1} = \phi_t(h_t, \theta_0) \tag{7}$$

from an arbitrary initial value h_0. As previously, sufficient conditions for the convergence of this SRE are the negativity of a Lyapunov coefficient and the existence of logarithmic moments, see [8]. So the GARCH(1,1) model is invertible as soon as $0 \leq \beta_0 < 1$. The invertibility of the EGARCH(1,1) model is more complicated to assert due to the exponential function in (6). The recursive relation on h_{t+1} can explode to $-\infty$ for small negative values of h_t and negative values of $\gamma_0 X_t + \delta_0 |X_t|$. However, assuming that $\delta_0 \geq |\gamma_0|$, the relation $\gamma_0 X_t + \delta_0 |X_t| > 0$ holds and conditions for invertibility of the EGARCH(1,1), denoted hereafter INV(θ_0), are obtained in [30]. On the contrary, note that Sorokin introduces sufficient conditions of non-invertibility for the EGARCH(1,1) model in [29]. Then the SRE (6) is completely chaotic for any possible choice of the initial value h_0 and the volatility forecasting procedure based on the model (5) is not reliable.

If financial datas are observed, the parameter $\theta_0 = (\alpha_0, \beta_0, \gamma_0, \delta_0)$ of the EGARCH(1,1) model is unknown. Nelson proposed in [27] to estimate θ_0 with the Quasi Maximum Likelihood Estimator (QMLE). Let us recall the definition of this classical estimator that estimates efficiently many GARCH models, see [5] for the GARCH case and [17] for the ARMA-GARCH case. To construct the QMLE one approximates the volatility using the observed SRE (7) at any θ. Assume that the

SREs driven by $\phi_t(\cdot,\theta)$ are stable for any θ. Let us consider the functions $\hat{g}_t(\theta)$ defined for any θ as the recursive solutions of the SRE

$$\hat{g}_{t+1}(\theta) = \phi_t(\hat{g}_t(\theta),\theta) \tag{8}$$

for some arbitrary initial value $\hat{g}_0(\theta)$. Assume that h is a bijective function of inverse $\ell > 0$. Then $\ell(\hat{g}_t(\theta_0))$ has a limiting law that coincides with the one of σ_t^2. The Quasi Likelihood (QL) criteria is defined as

$$2n\hat{L}_n(\theta) = \sum_{t=1}^{n} \hat{l}_t(\theta) = \sum_{t=1}^{n} X_t^2/\ell(\hat{g}_t(\theta))^2 + \log(\ell(\hat{g}_t(\theta)).$$

The correspondinf M-estimator is the QMLE $\hat{\theta}_n$ defined as

$$\hat{\theta}_n = \text{argmin}_{\theta\in\Theta}\hat{L}_n(\theta)$$

for some compact set Θ. This estimator has been used since the seminal paper of Nelson [27] for estimating the EGARCH model, see [10]. However, the consistency and the asymptotic normality are not proved except in the degenerate case $\beta_0 = 0$ in [30]. On the contrary with other GARCH models, the QML estimation procedure is not always reliable for the EGARCH model, see the discussion in [19]. Thus, other estimation procedure has been investigated such as the bayesian, bias correction and the Whittle procedure in [35, 12, 36] respectively. Another approach is to introduce models that behave as the EGARCH model but in which the QMLE could be more reliable, see [19, 32].

We prove the strong consistency of the QMLE for the general model (1). We give sufficient conditions called the continuous invertibility of the model such that the QMLE is strongly consistent. More precisely we assume that the SRE (8) produces continuous functions $\hat{g}_t$ of θ on Θ. The continuous invertibility holds when the limiting law of $\hat{g}_t$ corresponds to the law of some continuous function g_t on Θ that does not depend on the initial function $\hat{g}_0$. The continuous invertibility ensures the stability of the estimation procedure regardless the initial function $\hat{g}_0$ chosen arbitrarily in practice. Under few other assumptions, we prove that the QMLE is strongly consistent for continuously invertible models on the compact set Θ. The continuous invertiblity should be checked systematically on models before using QMLE. One example of such continuously invertible models with properties similar than the EGARCH model is the Log-GARCH model studied in [15].

Our approach differs from the one of [30] by new sufficient conditions for continuous invertibility. These conditions are collected in assumption **(CI)** defined below. These conditions ensure the invertibility of the model at any point θ of the compact set Θ and some regularity of the model with respect to the parameter θ. As the EGARCH(1,1) model is regular with respect to θ, it satisfies **(CI)** on any Θ such that the invertibility condition INV(θ) is satisfied for any $\theta \in \Theta$. Thus we prove the strong consistency of the QMLE for the invertible EGARCH(1,1) model when $\theta_0 \in \Theta$ in cases where $\beta_0 \neq 0$. It is a serious advantage of the sufficient condition **(CI)** compared with the one of [30] that is more restrictive when applied to the EGARCH(1,1) model. Moreover, we prove the strong consistency of the volatility forecasting $\hat{\sigma}_n^2 = \ell(\hat{g}_n(\hat{\theta}_n))$ of σ_{n+1}^2 under **(CI)**. Continuous invertibility seems to be well suited to assert volatility forecasting because $\hat{\sigma}_n^2$ expresses as functions $\hat{g}_n$ evaluated at points $\hat{\theta}_n \neq \theta_0$. To infer the EGARCH(1,1) model, we propose to constrain the QMLE on some compact set satisfying the empirical counterpart of the condition INV(θ_0). This new estimator $\hat{\theta}_n^c$ called Invertible QMLE (IQMLE) produces only reliable volatility forecasting, i.e. such as $\hat{\sigma}_n^2$ does not depend asymptotically on the arbitrary choice of the initial value $\hat{\sigma}_0^2 = \ell(\hat{g}_0)$.

The asymptotic normality of the QMLE in the EGARCH(1,1) model is proved under the additional assumption **(MM)**. The moment conditions in **(MM)** are sufficient for the existence of the asymptotic covariance matrix. No uniform moment condition on the score vector is assumed. The proof is based on functional SREs as (12) and their perturbations. As for the strong consistency, we need to refine arguments from [31]. As we do not apply any uniform Strong Law of Large Number, we develop in the proof new arguments based on SREs satisfied by differences such as $|\hat{L}_n(\hat{\theta}_n)-\hat{L}_n(\theta_0)|$. We consider conditions of moments **(MM)** only at the point θ_0 where the expression of the score vector simplifies. In the EGARCH(1,1) model, the conditions **(MM)** take the simple form $\mathbb{E}[Z_0^4]<\infty$ and $\mathbb{E}[(\beta_0-2^{-1}(\gamma_0 Z_0+\delta_0|Z_0|)^2]<1$ and can be checked in practice by estimating the innovations. We believe that this new approach gives sharp conditions for asymptotic normality for other models.

The paper is organized as follows. In Section 2, we discuss the standard notions of invertibility and introduce the continuous invertibility and its sufficient condition **(CI)**. We prove the strong consistency of the QMLE for general continuously invertible models in Section 3. The consistency of the volatility forecasting is also proved under the sufficient condition **(CI)**. We apply this results in the EGARCH(1,1) model in Section 4. For this model, we propose a new method called IQMLE that produces only reliable volatility forecasting. The asymptotic normality of $\hat{\theta}_n$ in the EGARCH(1,1) model is given in Section 5. The proofs of the Technical Lemmas are collected in Section 6.

2. Preliminaries

2.1. **The general volatility model.** In this paper, the innovations $Z_t\in\mathbb{R}$ are iid random variables (r.v.) such that Z_t is centered and normalized such as $\mathbb{E}[Z_0^2]=1$. Consider the general volatility model $X_t=\sigma_t Z_t$ satisfying $h(\sigma_{t+1}^2)=\psi_t(h(\sigma_t^2),\theta_0)$ for all $t\in\mathbb{Z}$. The function h is a bijection from some subset $\mathbb{R}^+$ to some subset of $\mathbb{R}$ of inverse ℓ called the link function.

A first question regarding such general SRE is the existence of the model, i.e. wether or not a stationary solution exists. The classical stability assumption is mainly the negativity of the top Lyapunov exponent, see [9, 8]. Hereafter, we work under the general assumption

(ST): The SRE (1) admits a unique stationary solution denoted (σ_t^2) that is non anticipative, i.e. σ_t^2 is independent of $(Z_t,Z_{t+1},Z_{t+2},\ldots)$ for all $t\in\mathbb{Z}$, and has finite log-moments: $\mathbb{E}\log^+\sigma_0^2<\infty$.

The GARCH(1,1) model (2) satisfies the condition **(ST)** if and only if (iff) $\mathbb{E}[\log(\beta_0+\gamma_0 Z_0^2)]<0$, see [26] for the existence of the stationary solution and [5] for the existence of log moments. The EGARCH(1,1) model (3) satisfies the condition **(ST)** iff $|\beta_0|<1$, see [27]. In this case, the model has nice ergodic properties: any process recursively defined by the SRE from an arbitrary initial value approximates exponentially fast a.s. the original process (σ_t^2). In the sequel, we say that the sequence of non negative r.v. (W_t) converges exponentially almost surely to 0, $W_t\xrightarrow{e.a.s.}0$ as $t\to\infty$, if $W_t=o(e^{-Ct})$ a.s. for some r.v. $C>0$. We will also use the notation x^+ for the positive part of x, i.e. $x^+=x\vee 0$ for any $x\in\mathbb{R}$.

2.2. **Invertible models.** Under **(ST)** the process (X_t) is stationary, non anticipative and thus ergodic as a Bernoulli shift of an ergodic sequence (Z_t), see [23]. Let us now investigate the question of invertibility of the general model (1). The classical notions of invertibility are related with convergences of SRE and thus are implied by Lyapunov conditions of Theorem 3.1 in [8]. Following [33], we say that a volatility model is invertible if the volatility can be expressed as a function of the past observed values:

Definition 1. *Under* **(ST)**, *the model is invertible if the sequence of the volatilities* (σ_t^2) *is adapted to the filtration generated by* $(X_{t-1},X_{t-2},\cdots)$.

Using the relation $Z_t = X_t/\ell(h(\sigma_t))$ in the expression of ψ_t yields the new SRE (5): $h(\sigma_{t+1}^2) = \phi_t(h(\sigma_t^2), \theta_0)$. Now the random functions $\phi_t(\cdot, \theta_0)$ depends only on X_t. As (X_t) is an ergodic and stationary process, it is also the case of the sequence of parametrized maps $(\phi_t(\cdot, \theta_0))$. Using Theorem 3.1 in [8], the invertibility of the model follows if the $\phi_t(\cdot, \theta_0)$ are Lipschitz maps such that it exists $r > 0$ satisfying

(9)

$$\mathbb{E}[\log^+ |\phi_0(x, \theta_0)|] < \infty \text{ for some } x \in E,\ \mathbb{E}[\log^+ \Lambda(\phi_0(\cdot, \theta_0))] < \infty \text{ and } \mathbb{E}[\log \Lambda(\phi_0(\cdot, \theta_0)^{(r)})] < 0.$$

Here $\Lambda(f)$ denotes the Lipschitz coefficient of any function f defined by the relation (in the case where f is real valued)

$$\Lambda(f) = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}.$$

The conditions (9) are called the conditions of invertibility in [31] and is proved there that

Proposition 1. *Under* **(ST)** *and* (9)*, the general model* (5) *is invertible.*

The GARCH(1,1) model (2) is invertible as soon as $0 \leq \beta_0 < 1$.

The invertibility of the EGARCH(1,1) model is more difficult to assert due to the exponantial function in the SRE (6). Let us describe the sufficient condition of invertibility of the EGARCH(1,1) given in [30, 31]. It expresses as a constraint on the coefficients $\theta_0 = (\alpha_0, \beta_0, \gamma_0, \delta_0)$:

INV(θ_0): $\delta_0 \geq |\gamma_0|$ and

$$\mathbb{E}[\log(\max\{\beta_0, 2^{-1}(\gamma_0 X_0 + \delta_0 |X_0|) \exp(-2^{-1}\alpha_0/(1 - \beta_0)) - \beta_0\})] < 0. \tag{10}$$

This invertible region does not depend on α_0 when the process (X_t) is itself the stationary solution of the EGARCH(1,1) model for θ_0. Indeed, $(\log \sigma_t^2)$ admits a MA(∞) representation

$$\log \sigma_t^2 = \alpha_0(1 - \beta_0)^{-1} + \sum_{k=1}^{\infty} \beta_0^{k-1}(\gamma_0 Z_{t-k} + \delta_0 |Z_{t-k}|).$$

Plugging in this MA(∞) representation into (10), we obtain the equivalent sufficient condition

$$\mathbb{E}\Big[\log\Big(\max\Big\{\beta_0, 2^{-1}\exp\Big(2^{-1}\sum_{k=0}^{\infty}\beta_0^k(\gamma_0 Z_{-k-1} + \delta_0\,|Z_{-k-1}|)\Big)(\gamma_0 Z_0 + \delta_0\,|Z_0|) - \beta_0\Big\}\Big)\Big] < 0. \tag{11}$$

Using the Monte Carlo algorithm and assuming that Z_0 is $\mathcal{N}(0, 1)$-distributed, we report in Figure 1 the largest values of β_0 that satisfies the condition (10) on a grid of values of γ_0 and δ_0.

3. Strong consistency of the QMLE for continuously invertible models

3.1. **Continuously invertible models.** Assume that the observations follow the model (5) for an unknown value θ_0. Consider the inference of the QMLE $\hat{\theta}_n$ defined as $\hat{\theta}_n = \text{argmin}_{\theta \in \Theta} \hat{L}_n(\theta)$ for some compact set Θ and where $\hat{L}_n$ is the QL criteria

$$2n\hat{L}_n(\theta) = \sum_{t=1}^{n} \hat{l}_t(\theta) = \sum_{t=1}^{n} \log(\ell(\hat{g}_t(\theta)) + X_t^2/\ell(\hat{g}_t(\theta)).$$

is based on the functional (observable) SRE of the form

$$\hat{g}_{t+1}(\theta) = \phi_t(\hat{g}_t(\theta), \theta), \qquad \forall t \geq 1, \forall \theta \in \Theta$$

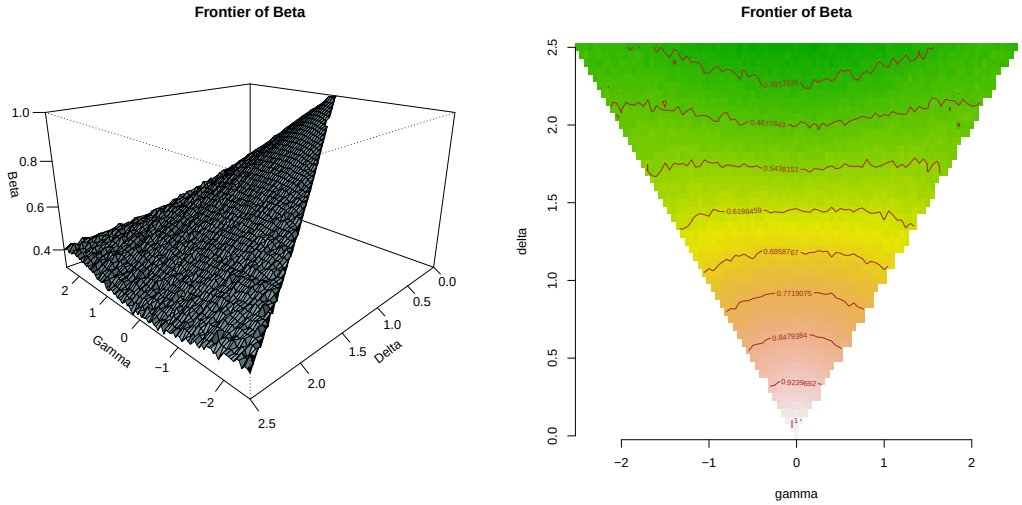

FIGURE 1. Perspective and contour plots of the domain for invertibility.

starting at an arbitrary initial value $\hat{g}_0(\theta)$, for any $\theta \in \Theta$. If the conditions of invertibility (9) are satisfied for $(\phi_t(\cdot, \theta))$ for any $\theta \in \Theta$ then the unique stationary solution (g_t) exists and is defined by the relation

$$g_{t+1}(\theta) = \phi_t(g_t(\theta), \theta), \qquad \forall t \in \mathbb{Z}, \forall \theta \in \Theta. \tag{12}$$

Assume from now that the random functions $(x, \theta) \to \phi_t(x, \theta)$ restricted on $\mathcal{K} \times \Theta$ take values in $\mathcal{K}$ where $\mathcal{K}$ is some subset of $\mathbb{R}$. Consider that the initial values $\hat{g}_0(\theta)$ is a continuous function on Θ that takes its values in $\mathcal{K}$. Denote $\|\cdot\|_\Theta$ the uniform norm and $\mathcal{C}_\Theta$ the space of continuous functions from Θ to $\mathcal{K}$. By a recursive argument, it is obvious that the random functions $\hat{g}_t$ belong to $\mathcal{C}_\Theta$ for all $t \geq 0$.

We are now ready to define the notion of continuous invertibility:

Definition 2. *The model is continuously invertible on* Θ *iff* $\|\hat{g}_t(\theta) - g_t(\theta)\|_\Theta \to 0$ *a.s. when* $t \to \infty$.

This continuous invertibility notion is crucial for the strong consistency of the QMLE.

3.2. **Strong consistency of the QMLE.** Let us do another assumption of regularity and boundedness from below on the link function ℓ:

(LB): The maps $x \to 1/\ell(x)$ and $\log(\ell(x))$ are Lipschitz functions on $\mathcal{K}$ and there exists $m > 0$ such that $\ell(x) \geq m$ for all $x \in \mathcal{K}$.

Classically, we also assume the identifiability of the model

(ID): $\ell(g_0(\theta)) = \sigma_0^2$ a.s. for some $\theta \in \Theta$ iff $\theta = \theta_0$.

The consistency of the QMLE follows from the continuous invertibility notion:

Theorem 1. *Assume that (***ST***),* **(LB)** *and* **(ID)** *for a volatility model that is continuously invertible on the compact set* Θ. *Then the QMLE is strongly consistent,* $\hat{\theta}_n \to \theta_0$ *a.s., when* $\theta_0 \in \Theta$.

Proof. First, note that from the continuous invertibility of the model the SRE (12) admits a stationary solution in $E = \mathcal{C}_\Theta$ that is a separable complete metric space equipped with the metric

$d(x,y) = \|x-y\|_\Theta$. Let us denote Φ_t the mapping acting on $\mathcal{C}_\Theta$ and satisfying

$$g_{t+1} = \Phi_t(g_t) \quad \text{iff} \quad g_{t+1}(\theta) = \phi_t(g_t(\theta),\theta), \quad \forall \theta \in \Theta.$$

We will apply the principle of Letac [24] extended to the stationary sequences as in [31]: the existence of a unique stationary solution which is non anticipative is ensured by the convergence a.s. of the backward equation for any initial value $g \in \mathcal{C}_\Theta$:

$$Z_t(g) := \Phi_0 \, o \, \Phi_{-1} o \cdots o \, \Phi_{-t}(g) \to_{t\to\infty} Z$$

where Z does not depend on g. By stationarity, $Z_t(g)$ is distributed as $\hat{g}_t$ when $\hat{g}_0 = g$. Thus, we have

$$\mathbb{P}(\lim_{t\to\infty} \|Z_t(g) - g_0\|_\Theta = 0) = \mathbb{P}(\lim_{t\to\infty} \|\hat{g}_t - g_t\|_\Theta = 0)$$

where $g_0(\theta)$ is well defined for each $\theta \in \Theta$ as the solution of the SRE (12). Thus under continuous invertibility, an application of Letac's principle leads to the existence of a unique stationary solution distributed denoted also (g_t) that coincides with $g_t(\theta)$ at any point $\theta \in \Theta$. In particular, g_0 belongs to $\mathcal{C}_\Theta$ thus is a continuous function on Θ.

Let us turn to the proof of the strong consistency based on standard arguments, see for example the book of Francq and Zakoïan [16]. As the model satisfies the identifiability condition **(ID)**, the strong consistency follows the intermediate results

(a) $\lim_{n\to\infty} \sup_{\theta\in\Theta} |L_n(\theta) - \hat{L}_n(\theta)| = 0$ a.s.
(b) $\mathbb{E}|\ell_0(\theta_0)| < \infty$ and if $\theta \neq \theta_0$ then $\mathbb{E}[\ell_0(\theta)] > \mathbb{E}[\ell_0(\theta_0)]$.
(c) Any $\theta \neq \theta_0$ has a neighborhood $V(\theta)$ such that

$$\lim \inf_{n\to\infty} \inf_{\theta^* \in V(\theta)} \hat{L}_n(\theta^*) > E\ell_t(\theta_0) \quad \text{a.s.}$$

Let us prove that (a) is satisfied when the model is continuously invertible and **(LB)** condition holds. As $1/\ell$ and $\log(\ell)$ are Lipschitz continuous functions there exists some constant $C > 0$ such that

$$|L_n(\theta) - \hat{L}_n(\theta)| \leq C \frac{1}{n} \sum_{t=1}^{n} |\hat{g}_t(\theta) - g_t(\theta)|. \tag{13}$$

The desired convergence to 0 of the upper bound that is a Cesaro mean follows from the definition of the continuous invertibility.

The first assertion of (b) follows from the identity

$$\ell_0(\theta_0) = \log(\sigma_0^2) + Z_0^2.$$

Thus, as $\sigma_0 = \ell(g_0(\theta_0)) \geq m$ from **(LB)** and $\mathbb{E}\log^+ \sigma_0^2 < \infty$ under **(ST)**, we have that $\log(\sigma_0^2)$ is integrable. Moreover, by assumption $\mathbb{E}Z_0^2 = 1$ and the assertion $\mathbb{E}|\ell_0(\theta_0)| < \infty$ is proved. To prove the second assertion, note that $\theta \to \mathbb{E}[\ell_0(\theta)]$ has a unique minimum iff $\mathbb{E}[\sigma_0^2/\ell(g_0(\theta)) - \log(\sigma_0^2/\ell(g_0(\theta))]$ has a unique minimum. Under the identifiability condition, as $x - \log(x) \geq 1$ for all $x > 0$ with equality iff $x = 1$, we deduce that for any $\theta \neq \theta_0$ we have $\mathbb{E}[\ell_0(\theta)] > \mathbb{E}[\ell_0(\theta_0)]$.

Finally, let us prove (c) under continuous invertibility, **(LB)** and **(ID)**. First, under continuous invertibility, we have

$$\begin{aligned} \lim \inf_{\theta^*\to\theta} \hat{L}_n(\theta^*) &= \lim \inf_{n\to\infty} \inf_{\theta^*\in V(\theta)} L_n(\theta^*) + \lim \inf_{n\to\infty} \inf_{\theta^*\in V(\theta)} \hat{L}_n(\theta^*) - L_n(\theta^*) \\ &= \lim \inf_{n\to\infty} \inf_{\theta^*\in V(\theta)} L_n(\theta^*) \end{aligned}$$

by using the convergence to 0 of the Cesaro mean (13). Second, by ergodicity of the stationary solution (g_t) we have the ergodicity of the sequence $(\inf_{\theta^* \in V(\theta)} \ell_t(\theta^*))$. Moreover, under condition **(LB)**, we have that $\inf_{\theta^* \in V(\theta)} \ell_t(\theta^*) \geq \log(m) > -\infty$ a.s. Then, for any $K < 0$, the sequence $\inf_{\theta^* \in V(\theta)} \ell_t(\theta^*) \vee K$ is integrable. We use the classical SLLN and obtain

$$\lim_{n\to\infty} \inf_{\theta^* \in V(\theta)} L_n(\theta^*) \vee K = \mathbb{E}\Big[\inf_{\theta^* \in V(\theta)} \ell_0(\theta^*) \vee K \Big] \qquad a.s.$$

Letting $K \to -\infty$ we obtain that

$$\lim_{n\to\infty} \inf_{\theta^* \in V(\theta)} L_n(\theta^*) = \mathbb{E}\Big[\inf_{\theta^* \in V(\theta)} \ell_0(\theta^*) \Big] \in \mathbb{R} \cup \{+\infty\}.$$

Finally, remark that by continuity of $\theta \to g_t(\theta)$ the functions L_n are continuous. Thus, for any $\varepsilon > 0$ we can find a neighborhood $V(\theta)$ such that

$$\mathbb{E}\Big[\inf_{\theta^* \in V(\theta)} \ell_0(\theta^*) \Big] \geq \mathbb{E}[\ell_0(\theta)] + \varepsilon.$$

From the second assertion of (b) $\varepsilon > 0$ can be chosen such that $\mathbb{E}[\ell_0(\theta)] + \varepsilon > \mathbb{E}[\ell_0(\theta_0)]$ and (c) is proved. The end of the proof of the strong consistency is based on a classical compact argument and thus is omitted. □

3.3. **Sufficient conditions for continuous invertibility.** The definition of continuous invertibility does not give any explicit condition. In order to apply Theorem 3.1 in [8], one classically assume some uniform Lipschitz conditions on the SRE (12).

Consider some generic function $f : \mathcal{K} \times \Theta \mapsto \mathcal{K}$. Assume that there exists a continuous function Λ_f on Θ such that for each x, $y \in E$ we have

$$|f(x,\theta) - f(y,\theta)| \leq \Lambda_f(\theta)|x - y|.$$

The approach followed by Straumann and Mikosch in [31] is to consider the SRE (8) in the complete metric space of continuous functions $\mathcal{C}_\Theta$ on Θ with values in $\mathcal{K}$ equipped with $d(x,y) = \|x - y\|_\Theta$. Sufficient conditions are the following ones

$$\mathbb{E}\log^+(\|\phi_0(y,\cdot)\|_\Theta) < \infty \text{ for some } y \in \mathcal{K},\ \mathbb{E}\log^+(\|\Lambda_{\phi_0}\|_\Theta) < \infty \text{ and } \mathbb{E}\log\left(\left\|\Lambda_{\phi_0^{(r)}}\right\|_\Theta\right) < 0. \tag{14}$$

The EGARCH(1,1) model satisfy this condition under restrictive assumptions on Θ, for example when $\beta_0 = 0$ see [30].

We introduce sufficient conditions for continuous invertibility collected in the assumption **(CI)**

(CI): $\mathbb{E}\log^+(\|\phi_0(y,\cdot)\|_\Theta) < \infty$ for some $y \in \mathcal{K}$, $\mathbb{E}\log^+(\|\Lambda_{\phi_0}\|_\Theta) < \infty$ and $\mathbb{E}\log\left(\Lambda_{\phi_0^{(r)}}(\theta)\right) < 0$ for any $\theta \in \Theta$.

The difference with the conditions (14) is that the condition on the top Lyapunov exponent holds now for any $\theta \in \Theta$. Of course it is implied by the uniform Lyapunov condition (14). When applied to models such as EGARCH(1,1), this new sufficient condition of continuous invertibility **(CI)** is satisfied as soon as the model is invertible on Θ.

Theorem 2. *If* **(ST)** *and* **(CI)** *hold then the model is continuously invertible.*

Proof. For any $\rho > 0$, let us write $\Lambda_*^{(r)}(\theta,\rho) = \sup\{\Lambda_{\phi_0}^{(r)}(\theta^*), \theta^* \in \overline{B}(\theta,\rho) \cap \Theta\}$, where $\overline{B}(\theta,\rho)$ stands for the closed ball centered at θ with radius ρ. Note that for any $K > 0$ we have $\mathbb{E}[\sup_\Theta |\log \Lambda_{\phi_0}^{(r)}(\theta) \vee K|] < \infty$ because $\mathbb{E}[\sup_\Theta \log^+ \Lambda_{\phi_0}^{(r)}(\theta)] < \infty$ under **(CI)**. Applying the dominated convergence

theorem we obtain $\lim_{\rho\to 0}\mathbb{E}[\log\Lambda_*^{(r)}(\theta,\rho)\vee K]=\mathbb{E}[\lim_{\rho\to 0}\log\Lambda_*^{(r)}(\theta,\rho)\vee K]$. By continuity we have $\lim_{\rho\to 0}\log\Lambda_*^{(r)}(\theta,\rho)=\log\Lambda_{\phi_0}^{(r)}(\theta)$ and for sufficiently small $K<0$

$$\lim_{\rho\to 0}\mathbb{E}[\log\Lambda_*^{(r)}(\theta,\rho)\vee K]=\mathbb{E}[\log\Lambda_{\phi_0}^{(r)}(\theta)\vee K]<0.$$

Thus, there exists an $\epsilon>0$ such that

$$\mathbb{E}[\log\Lambda_*^{(r)}(\theta,\epsilon)]\leq\mathbb{E}[\log\Lambda_*^{(r)}(\theta,\epsilon)\vee K]<0.$$

Denote $\mathcal{V}(\theta)$ the compact neighborhood $\overline{B}(\theta,\epsilon)\cap\Theta$ of θ. Let us now work on $\mathcal{C}_{\mathcal{V}(\theta)}$, the complete metric space of continuous functions from $\mathcal{V}(\theta)$ to $\mathcal{K}$ equipped with the supremum norm $\|\cdot\|_{\mathcal{V}(\theta)}$. In this setting $(\hat{g}_t)$ satisfy a functional SRE $\hat{g}_{t+1}=\Phi_(\hat{g}_t)$ with Lipschitz coefficients satisfying

$$\begin{aligned}
\Lambda(\Phi_t^{(r)})&\leq\sup_{s_1,s_2\in\mathcal{C}(\mathcal{V}(\theta))}\frac{\|\Phi_t^{(r)}(s_1)-\Phi_t^{(r)}(s_2)\|_{\mathcal{V}(\theta)}}{\|s_1-s_2\|_{\mathcal{V}(\theta)}}\\
&\leq\sup_{s_1,s_2\in\mathcal{C}(\mathcal{V}(\theta))}\frac{\sup_{\theta^*\in\mathcal{V}(\theta)}\|\phi_t^{(r)}(s_1(\theta^*),\theta^*)-\phi_t^{(r)}(s_2(\theta^*),\theta^*)\|}{\|s_1-s_2\|_{\mathcal{V}(\theta)}}\\
&\leq\sup_{s_1,s_2\in\mathcal{C}(\mathcal{V}(\theta))}\frac{\sup_{\theta^*\in\mathcal{V}(\theta)}\Lambda(\phi_t^{(r)}(\cdot,\theta^*))\|s_1(\theta^*)-s_2(\theta^*)\|}{\|s_1-s_2\|_{\mathcal{V}(\theta)}}\\
&\leq\sup_{s_1,s_2\in\mathcal{C}(\mathcal{V}(\theta))}\frac{\sup_{\theta^*\in\mathcal{V}(\theta)}\Lambda(\phi_t^{(r)}(\cdot,\theta^*))\|s_1-s_2\|_{\mathcal{V}(\theta)}}{\|s_1-s_2\|_{\mathcal{V}(\theta)}}\leq\Lambda_*^{(r)}(\theta,\epsilon).
\end{aligned}$$

As $\mathbb{E}[\log^+(\|\Phi_0(y)\|_{\mathcal{V}(\theta)})]\leq\mathbb{E}[\log^+(\|\phi_0(y,\theta)\|_\Theta)]<\infty$, $\mathbb{E}[\log^+(\Lambda(\Phi_0))]\leq\mathbb{E}[\log^+(\|\Lambda_{\phi_0}(\theta)\|_\Theta)]<\infty$ under **(CI)** and $\mathbb{E}(\log\Lambda(\Phi_0^{(r)}))<0$ we can apply Theorem 3.1 of [8]. The unique stationary solution (g_t) exists and satisfies

$$\|\hat{g}_t-g_t\|_{\mathcal{V}(\theta)}\xrightarrow{e.a.s.}0. \tag{15}$$

Now, let us remark that $\Theta=\cup_{\theta\in\Theta}\mathcal{V}(\theta)$. As Θ is a compact set, there exists a finite number N such that $\Theta=\cup_{k=1}^N\mathcal{V}(\theta_k)$ and we obtain

$$\|\hat{g}_t-g_t\|_{\mathcal{V}(\theta_k)}\xrightarrow{e.a.s.}0\quad\forall 1\leq k\leq N.$$

The desired result follows as $\|\cdot\|_\Theta=\vee_{k=1}^N\|\cdot\|_{\mathcal{V}(\theta_k)}$. □

3.4. **Volatility forecasting.** Based on the QMLE $\hat{\theta}_n$ we deduce a natural forecasting $\hat{\sigma}_n^2=\ell(\hat{g}_t(\hat{\theta}_n))$ of the volatility $\hat{\theta}_n$. It is strongly consistent:

Theorem 3. *Assume that* **(ST)**, **(LB)**, **(ID)** *and* **(CI)** *hold for* $\theta_0\in\Theta$. *Then* $|\hat{\sigma}_n^2-\sigma_n^2|\xrightarrow{a.s.}0$ *as* $n\to\infty$.

Proof. Theorems 1 and 2 assert the strong consistency of $\hat{\theta}_n$ under the assumption of Theorem 3. By continuity of ℓ and by uniform integrability, $|\ell(\hat{g}_t(\hat{\theta}_n))-\ell(g_t(\hat{\theta}_n))|\xrightarrow{a.s.}0$. Thus, by a similar continuity argument on ℓ, the result is proved if $|g_t(\hat{\theta}_n)-g_t(\theta_0)|\xrightarrow{a.s.}0$. Keeping the notation used in the proofs of Theorems 1 and 2, we have $\hat{\theta}_n\in\mathcal{V}(\theta_0)$ for n sufficiently large. The Lyapunov condition $\mathbb{E}[\log\Lambda(\Phi_0^{(r)})]<0$ is satisfied for the functional SRE (Φ_t) restricted on $\mathcal{V}(\theta_0)$. For any $\theta\in\mathcal{V}(\theta_0)$, $t\in\mathbb{Z}$ we have

$$|g_{(t+1)r}(\theta)-g_{(t+1)r}(\theta_0)|\leq\Lambda(\Phi_t^{(r)})|g_t(\theta)-g_t(\theta_0)|+w_t(\theta)$$

where $w_t(\theta) = |\phi_t(g_t(\theta_0), \theta) - \phi_t(g_t(\theta_0), \theta_0)|$. Thus $|g_t(\theta) - g_t(\theta_0)|$ is bounded by a linear SRE. An application of the Borel-Cantelli Lemma as in [5, 31], the convergence of the series

$$\|g_{tr} - g_{tr}(\theta_0)\|_{\mathcal{V}(\theta_0)} \leq \sum_{i=0}^{\infty} \Lambda(\Phi_{tr}^{(r)}) \cdots \Lambda(\Phi_{(t-i+1)r}^{(r)}) \|w_{(t-i)r}\|_{\mathcal{V}(\theta_0)} < \infty$$

follows from the fact that $\Lambda(\Phi_{tr}^{(r)}) \cdots \Lambda(\Phi_{(t-i+1)r}^{(r)}) \xrightarrow{e.a.s.} 0$ when $i \to \infty$ and $\mathbb{E} \log^+ \|w_{t-i}\|_{\mathcal{V}(\theta_0)} < \infty$ Thus the difference $|g_{tr}(\theta) - g_{tr}(\theta_0)|$ is bounding by an a.s. normally convergent function on $\mathcal{C}_{\mathcal{V}(\theta_0)}$ that we denote $a(\theta)$. Moreover, as $\lim_{\theta \to \theta_0} w_t(\theta) = 0$ for any $t \in \mathbb{Z}$ a.s., we also have $\lim_{\theta \to \theta_0} a_t(\theta) = 0$ a.s. from the normal convergence. Finally, using the strong consistency of $\hat{\theta}_n$, we obtain that

$$|g_t(\hat{\theta}_n) - g_t(\theta_0)| \leq a(\hat{\theta}_n) \xrightarrow{a.s.} 0 \quad \forall n \geq t \quad \text{when} \quad t \to \infty. \tag{16}$$

□

4. Strong consistency and IQMLE for the EGARCH(1, 1) model

4.1. **Strong consistency of the QMLE.** Recall that (Z_t) be an iid sequence of r.v. such that $\mathbb{E}(Z_0) = 0$ and $\mathbb{E}(Z_0^2) = 1$. The EGARCH(1, 1) model introduced by [27] is an AR(1) model for $\log \sigma_t^2$,

$$X_t = \sigma_t Z_t \quad \text{with} \quad \log \sigma_t^2 = \alpha_0 + \beta_0 \log \sigma_{t-1}^2 + \gamma_0 Z_t + \delta_0 |Z_t|.$$

The volatilities process (σ_t^2) exists, is stationary and ergodic as soon as $|\beta_0| < 1$ with no other constraint on the coefficients. and **(ST)** holds. However, it does not necessarily have finite moment of any order. The model is identifiable as soon as the distribution of Z_0 is not concentrated on two points, see [31]. Let us assume in the rest of the paper these classical conditions satisfied such that **(ST)**, **(LB)** and **(ID)** automatically hold.

The invertibility of the stationary solution of the EGARCH(1, 1) model does not hold in general. Let us recall how the sufficient condition for invertibility is found in [30], see also [31]. Invertign the model, we obtain the SRE

$$\phi_t(x, \theta) = \alpha + \beta x + (\gamma X_t + \delta |X_t|) \exp(-x/2).$$

Under the constraint $\delta \geq |\gamma|$ it is clear that $\gamma X_t + \delta |X_t| \geq 0$. By a straightforward monotonicity argument, we can consider the restriction of the SRE on the set $\mathcal{K} \times \Theta$ where $\mathcal{K} = [\alpha/(1 - \beta), \infty)$ and $\delta \geq |\gamma|$ for any $\theta \in \Theta$. By regularity of ϕ_t, the Lipschitz coefficients are computed using the first derivatives:

$$\Lambda(\phi_t(\cdot, \theta)) \leq \max\{\beta, 2^{-1}(\gamma X_t + \delta |X_t|) \exp(-2^{-1}\alpha/(1 - \beta)) - \beta\}.$$

This Lipschitz coefficient is continuous in θ and thus coincides with $\Lambda_{\phi_t}(\theta)$. The assumption **(CI)** is satisfied if INV(θ) is satisfied for any $\theta \in \Theta$ as the condition of log moments are automatically satisfied by continuity and because $\mathbb{E}[\log^+ X_0] < \infty$.

Remark that the domain of invertibility represented in Figure 1 does not coincide with the domain of continuous invertibility, i.e. the set of $\theta \in \mathbb{R}^4$ satisfying INV(θ) when (X_t) is the stationary solution of the EGARCH(1,1) model for a fixed $\theta_0 \in \mathbb{R}^4$. This situation is more complicated to represent as the domain INV(θ) depends on the fixed value θ_0 and on the parameter α when $\theta \neq \theta_0$. However, for any θ_0 in the invertibility domain represented in Figure 1 there exists some compact neighborhood Θ of continuous invertibility. For the QMLE computed on such compact set Θ, an application of Theorems 1, 2 and 3 gives directly

Theorem 4. *If INV(θ) is satisfied for any $\theta \in \Theta$ and $\theta_0 \in \Theta$ then $\hat{\theta}_n \to \theta_0$ and $\hat{\sigma}_n^2 - \sigma_n^2 \to 0$ a.s. as $n \to \infty$ with $\hat{\sigma}_n^2 = \exp(\hat{g}_n(\hat{\theta}_n))$ for any initial value $\hat{\sigma}_0^2$.*

Remark we extend the result of [30, 31] under the sufficient condition (14) that expresses in the EGARCH(1,1) model as

$$\mathbb{E}[\sup_{\Theta} \log(\max\{\beta, 2^{-1}(\gamma X_{t-1} + \delta|X_{t-1}|)\exp(-2^{-1}\alpha/(1-\beta)) - \beta\})] < 0.$$

This more restrictive condition is less explicit because of the supremum and thus difficult to check when $\beta \neq 0$.

4.2. **The Invertible QMLE (IQMLE).** Notice that the procedure is valid only if the invertibility condition INV(θ) is satisfies for any $\theta \in \Theta$. We propose to optimize the QL criteria under the empirical constraint

$\widehat{\text{INV}(\theta)}$: $\delta \geq |\gamma|$ and $\sum_{t=1}^{n} \log(\beta, 2^{-1}(\gamma X_t + \delta|X_t|)\exp(-2^{-1}\alpha/(1-\beta)) - \beta) < 0$.

Definition 3. *The IQMLE is the M-estimator*

$$\hat{\theta}_n^c = argmin_{\theta\in\Theta^c} \sum_{t=1}^{n} 2^{-1}\left(X_t^2 \exp(-\hat{g}_t(\theta)/2) + \hat{g}_t(\theta)\right)$$

where $\Theta^c = \{\theta \in \Theta$ *satisfying* $\widehat{INV(\theta)}\}$ *for any compact set* Θ.

The following theorem gives the strong consistency of the IQMLE for the EGARCH(1,1) model under continuous invertibility on Θ and $\theta_0 \in \Theta$. It also shows that the volatility forecasting using the IQMLE does not depend on the arbitrary choice of the initial value even when INV(θ_0) is not satisfied. Thus the IQMLE is more reliable than the QMLE for which chaotic behavior of the volatility forecasting described in [29] can occur if INV(θ_0) is not satisfied.

Theorem 5. *If INV(θ) is satisfied for any $\theta \in \Theta$ then $\hat{\theta}_n^c \xrightarrow{a.s.} \theta_0$. If INV($\theta_0$) is not satisfied, the asymptotic law of $\hat{\sigma}_t^2 = \exp(\hat{g}_t(\hat{\theta}_n^c))$ still does not depend on the initial value $\hat{\sigma}_0^2$.*

Proof. Denote $\Lambda_n(\theta) = \max\{\beta, 2^{-1}(\gamma X_n + \delta|X_n|)\exp(-2^{-1}\alpha/(1-\beta)) - \beta\}$. We will prove that $\|n^{-1}\log\Lambda_n(\theta) - \mathbb{E}\log\Lambda_0(\theta)\|_\Theta \xrightarrow{a.s.} 0$ as $n \to \infty$. Then, if INV(θ) is satisfied for any $\theta \in \Theta$, Θ^c coincides a.s. with Θ asymptotically. An application of Theorem 4 yields the strong consistency of the IQMLE. The second assertion follows from the fact that asymptotically INV($\hat{\theta}_n^c$) is satisfied even if $\theta_0 \notin \Theta^c$.

Let us prove that $\|n^{-1}\log\Lambda_n(\theta) - \mathbb{E}\log\Lambda_0(\theta)\|_\Theta \xrightarrow{a.s.} 0$ when $\beta \neq 0$ only. Notice that $\log\Lambda_n(\theta)$ is a random element in the Banach space $\mathcal{C}(\Theta)$. The desired result is a consequence of the ergodic theorem if $\mathbb{E}\|\log\Lambda_0\|_\Theta < \infty$. Since

$$\|\log\Lambda_0\|_\Theta \leq \max\{\|\log|\beta|\|_\Theta, \|\log|2^{-1}(\gamma X_0 + \delta|X_0|)\exp(-2^{-1}\alpha/(1-\beta)) - \beta|\|_\Theta\}$$

and Θ is a compact set, the desired result follows by continuity and the dominated convergence theorem as $\mathbb{E}\log|X_0| = \mathbb{E}\log\sigma_0 + \mathbb{E}\log|Z_0| < \infty$. □

The advantage of the IQMLE $\hat{\theta}_n^c$ is that the corresponding volatility forecasting is stable with respect to the choice of the initial value, and that it is strongly consistent when the QMLE is strongly consistent.

5. The asymptotic normality of the (I)QMLE in the EGARCH(1,1) model

In this section we extend the result of Theorem 5.7.9 of [30] to non-degenerate cases when $\beta_0 \neq 0$. We obtain the asymptotic normality of the QMLE and IQMLE without assuming uniform moment on the compact set Θ for the likelihood and its derivatives. However, we assume additionally that

(MX): (σ_t^2) is geometrically ergodic.

It is the case if Z_0 is nonsingular w.r.t. the Lebesgue measure on $\mathbb{R}$, see [1]. Let us assume the finite moments assumption

(MM): $\mathbb{E}[Z_0^4] < \infty$ and $\mathbb{E}[(\beta_0 - 2^{-1}(\gamma_0 Z_0 + \delta_0 |Z_0|))^2] < 1$.

Theorem 6. *Assume that assumptions of Theorem 4 are satisfied, that* $\theta_0 \in \overset{\circ}{\Theta}$ *and that* **(MX)** *and* **(MM)** *hold. Then* $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathcal{N}(0, \mathbf{\Sigma})$ *and* $\sqrt{n}(\hat{\theta}_n^c - \theta_0) \xrightarrow{d} \mathcal{N}(0, \mathbf{\Sigma})$ *where* $\mathbf{\Sigma}$ *is an invertible matrix.*

Proof. As the IQMLE is asymptotically equivalent to the QMLE under the assumptions of Theorem 4 we will prove the result only for this last estimator. We first prove that Assumption **(MM)** yields the existence of the asymptotic variance

$$\mathbf{\Sigma} = \mathbf{J}^{-1}\mathbf{I}\mathbf{J}^{-1}$$

with $\mathbf{J} = \mathbb{E}[\mathbb{H}\ell_0(\theta_0)]$ and $\mathbf{I} = \mathbb{E}[\nabla\ell_0(\theta_0)\nabla\ell_0(\theta_0)^T]$, where $\mathbb{H}\ell_0(\theta_0)$ and $\nabla\ell_0(\theta_0)$ are the Hessian matrix and the gradient vector of ℓ_0 at the point $\theta_0 \in \overset{\circ}{\Theta}$.

Lemma 1. *Under condition* **(MM)** *the moments* $\mathbf{\Sigma}$ *exist and form an invertible matrix.*

The proof of this lemma is given in Section 6. Then we prove that the functions $\hat{g}_t$ and g_t are two times continuously differentiable refining the arguments developed in [31]

Lemma 2. *Under the assumptions of Theorem 4 the functions* $\hat{g}_t$ *and* g_t *are two times continuously differentiable on a compact neighborhood* $\mathcal{V}(\theta_0)$ *of* $\theta_0 \in \overset{\circ}{\Theta}$ *and* $\|\nabla\hat{g}_t - \nabla g_t\|_{\mathcal{V}(\theta_0)} + \|\mathbb{H}\hat{g}_t - \mathbb{H}g_t\|_{\mathcal{V}(\theta_0)} \xrightarrow{e.a.s.} 0$.

The proof of this lemma is given in Section 6. The asymptotic normality follows from the Taylor development used in Section 5 of Bardet and Wintenberger [2] on the partial derivatives ∇_i of the real valued function L_n:

$$\nabla_i L_n(\hat{\theta}_n) - \nabla_i L_n(\theta_0) = \mathbb{H}L_n(\tilde{\theta}_{n,i})(\hat{\theta}_n - \theta_0) \quad \text{for all } 1 \le i \le d.$$

Then the asymptotic normality follows from the following sufficient conditions:

(a) $n^{-1/2}\nabla L_n(\theta_0) \to \mathcal{N}(0, \mathbf{I})$,
(b) $\|n^{-1}\mathbb{H}L_n(\tilde{\theta}_n) - \mathbf{J}\| \xrightarrow{\mathbb{P}} 0$ for any sequence $(\tilde{\theta}_n)$ converging a.s. to θ_0 and $\mathbf{J}$ is invertible,
(c) $n^{-1/2}\|\nabla\hat{L}_n(\hat{\theta}_n) - \nabla L_n(\hat{\theta}_n)\|$ converges a.s. to 0.

Note that

$$\nabla L_n(\theta_0) = \sum_{i=t}^{n} 4^{-1}\nabla g_t(\theta_0)(1 - \mathbb{Z}_t^2)$$

is a martingale as $\nabla g_t(\theta_0)$ is independent of Z_t for all $t \ge 1$ and $\mathbb{E}[1 - \mathbb{Z}_t^2)] = 0$ by assumption. Under **(MM)**, this martingale has finite moments of order 2 due to Lemma 1. An application of the CLT for differences of martingales of [6] yields (a).

The following Lemma is used to prove (b) without uniform moment assumption on the compact set Θ for the likelihood and its derivatives

Lemma 3. *Under the assumptions of Theorem 6 we have* $\|\mathbb{H}L_n(\tilde{\theta}_n) - \mathbb{H}L_n(\theta_0)\| \xrightarrow{a.s.} 0$ *for any sequence* $(\tilde{\theta}_n)$ *converging a.s. to* θ_0 *when* $n \to \infty$.

The proof of this lemma is given in Section 6. To prove that $\|n^{-1}\mathbb{H}L_n(\tilde{\theta}_n) - \mathbf{J}\| \xrightarrow{a.s.} 0$, using hte result of Lemma 3 it is sufficient to prove that $\|n^{-1}\mathbb{H}L_n(\theta_0) - \mathbf{J}\| \xrightarrow{a.s.} 0$. An application of the ergodic Theorem on the process $(\mathbb{H}l_t(\theta_0))$ that is integrable under **(MM)** yields $\|n^{-1}\mathbb{H}L_n(\tilde{\theta}_n) - \mathbf{J}\| \xrightarrow{a.s.} 0$. Thus the first assertion of (b) is proved. The fact that $\mathbf{J}$ is an invertible matrix is already known, see Lemma 1.

Finally (c) is obtained by using the exponential decrease of the approximation given in (15) and (**??**) uniformly on some compact neighborhood $\mathcal{V}$ of θ_0 and the identity

$$\nabla L_n = \sum_{i=t}^{n} 4^{-1}\nabla g_t(1 - X_t^2 \exp(-g_t)).$$

Note that we use again the fact that $\exp(-x)$ is a Lipschitz function on $[c, \infty)$. □

6. Proofs of the technical lemmas

6.1. **Proof of Lemma 1.** Let us denote $U_t = (1, \log\sigma_t^2, Z_t, |Z_t|)$ and $V_t = \beta_0 - 2^{-1}(\gamma_0 Z_t + \delta_0 |Z_t|)$. Then $(\nabla g_t(\theta_0))$ is the solution of the linear SRE

$$\nabla g_{t+1}(\theta_0) = U_t + V_t \nabla g_t(\theta_0), \qquad \forall t \in \mathbb{Z}.$$

Let us consider the process $Y_t = (\nabla g_t(\theta_0), \log(\sigma_t^2))' \in \mathbb{R}^5$. It satisfies the relation

$$Y_{t+1} = \begin{pmatrix} V_t & 0 & 0 & 0 & 0 \\ 0 & V_t & 0 & 0 & 1 \\ 0 & 0 & V_t & 0 & 0 \\ 0 & 0 & 0 & V_t & 0 \\ 0 & 0 & 0 & 0 & \beta_0 \end{pmatrix} Y_t + \begin{pmatrix} 1 \\ 0 \\ Z_t \\ |Z_t| \\ \alpha_0 + \gamma_0 Z_t + \delta_0 |Z_t| \end{pmatrix} =: \Gamma_t Y_t + R_t, \qquad \forall t \in \mathbb{Z}.$$

Thus (Y_t) is a random coefficients autoregressive model satisfying the assumptions of Theorem 4 (a) of [34] iff $\mathbb{E}V_t^2 < 1$. By a direct application of the Theorem 4 (a) of [34] we obtain that the process (Y_t) is second order stationary and thus the existence of the matrix $\mathbf{B} = \mathbb{E}[\nabla g_t(\theta_0)(\nabla g_t(\theta_0))^T]$.

Let us prove that $\mathbf{B}$ is invertible. By classical arguments, it is sufficient to prove that the components of the vector $\nabla g_0(\theta_0)$ are linearly independent. It is the case in the AGARCH(1,1) model as soon as the density of Z_0 is not concentrated on two points, se Lemma 8.2 of [31]. Thus $\mathbf{B}$ is an invertible matrix.

Finally, we have the identity $\mathbf{I} = 2^{-1}\mathbf{B}$ as

$$\mathbf{I} = 2^{-1}\mathbb{E}\left[(\nabla g_t(\theta_0)(\nabla g_t(\theta_0))^T Z_0^2 + \mathbb{H}g_t(\theta_0)(1 - Z_0^2)\right] = 2^{-1}\mathbb{E}[\nabla g_t(\theta_0)(\nabla g_t(\theta_0))^T] = 2^{-1}\mathbf{B}.$$

We also have the identity $\mathbf{J} = 4^{-1}(\mathbb{E}Z_0^4 - 1)\mathbf{B}$ because

$$\mathbf{J} = \mathbb{E}\left[4^{-1}\mathbb{E}\left[\nabla g_t(\theta_0)(\nabla g_t(\theta_0))^T(1 - \mathbb{Z}_t^2)^2\right] |\mathcal{F}_{t-1}\right]$$
$$= 4^{-1}\mathbb{E}[(1 - \mathbb{Z}_0^2)^2]\mathbb{E}[\nabla g_t(\theta_0)(\nabla g_t(\theta_0))^T] = 4^{-1}(\mathbb{E}Z_0^4 - 1)\mathbf{B}.$$

Thus by definition $\mathbf{\Sigma} = (\mathbb{E}Z_0^4 - 1)\mathbf{B}^{-1}$ exists and is an invertible matrix. The lemma is proved.

6.2. **Proof of Lemma 2.** The proof of the existence of the derivatives of the process (g_t) requires a refinement of Theorem 3.1 of [8]. We give this new result in full generality when the SRE is on a Polish space (E, d). A map $f : E \to E$ is a Lipschitz map if $\Lambda(f) = \sup_{(x,y)\in E^2} d(f(x), f(y))/d(x, y)$ is finite. Note that the proof of the existence of the stationary solution is due to Elton [13]. That the approximation scheme is e.a.s. convergent is due to Bougerol [8]. Both results hold under the assumption that the Lipschitz coefficients have a finite log-moment of order 1. Here we assume

additionally a finite log-moment of order $p > 2$ and a strongly mixing condition with geometric rate to assert some regularity on the solution of the SRE. The regularity property we study is the following one: let $(Y_t)_{t\geq 0}$ be a sequence of non negative r.v.

(EAS): For any non negative sequence $W_t \xrightarrow{e.a.s.} 0$ the series $(W_t Y_t)$ converges a.s.

Remark that **(EAS)** is implied for stationary sequences by a condition of log-moment of order 1 as in [8] by an application of the Borel Cantelli Lemma as in [5, 31]. More generally, the same reasoning yields that **(EAS)** is satisfied for any sequence (Y_t) such as the series $(\mathbb{P}(Y_t \geq \rho^t)) = (P(\log^+ Y_t \geq t\varepsilon))$ is convergent for any $0 < \rho < 1$ and $\varepsilon > 0$. It is also satisfied for any solution of a convergent SRE:

Theorem 7. *Let (Ψ_t) be a stationary ergodic sequence of Lipschitz maps from E to E that is also strongly mixing with geometric rate. Assume that $(d(\Psi_t(x), x))$ satisfies* **(EAS)** *for some $x \in E$, $\mathbb{E}[(\log^+ \Lambda(\Psi_0))^p] < \infty$ for some $p > 2$ and*

$$\mathbb{E}[\log \Lambda(\Psi_0^{(r)})] = \mathbb{E}[\log \Lambda(\Psi_0 \circ \cdots \circ \Psi_{-r+1})] < 0 \text{ for some } r \geq 1. \tag{17}$$

Then the SRE $Y_{t+1} = \Psi_t(Y_t)$ for all $t \in \mathbb{Z}$ is convergent: it admits a unique stationary solution $(Y_t)_{t\in\mathbb{Z}}$ which is ergodic and for any $y \in E$

$$Y_t = \lim_{m\to\infty} \Psi_t \circ \cdots \circ \Psi_{t-m}(y), \quad t \in \mathbb{Z}.$$

The Y_t are measurable with respect to the $\sigma(\Psi_{t-k}, k \geq 0)$ and

$$d(\hat{Y}_t, Y_t) \xrightarrow{e.a.s.} 0, \quad t \to \infty$$

for $(\tilde{Y}_t)$ satisfying $\tilde{Y}_{t+1} = \Psi_t(\hat{Y}_t)$ for all $t > 0$ and $\hat{Y}_0 = y$ for any $y \in E$. Moreover $d(Y_t, y)$ satisfies **(EAS)** *for any $y \in E$.*

Proof. A careful look at the proof of Theorem 3.1 of [8] shows that the condition $\mathbb{E}[\log^+ d(\Psi_0(x), x)] < \infty$ is only used to assert **(EAS)** on $(d(\Psi_t(x), x))$. Thus we only prove the last assertion as the other ones follows from the proof of Theorem 3.1 of [8].

First notice that $(d(\hat{Y}_t, Y_t))$ satisfies **(EAS)** because $d(\hat{Y}_t, Y_t) \xrightarrow{e.a.s.} 0$. Let us show that $d(\hat{Y}_t, y)$ satisfies **(EAS)** when $\hat{Y}_0 = y$. Fix $K < 0$ such that $\mathbb{E}[\log \Lambda(\Psi_0^{(r)}) \vee K] \leq \log a$ for some $0 < a < 1$. Then

$$d(\hat{Y}_t, y) \leq \sum_{j=1}^{t} d(\hat{Y}_j, \hat{Y}_{j-1}) \leq \sum_{j=1}^{t} \Lambda(\Psi_j^{(j-1)}) d(\Psi_1(y), y) \leq \frac{\sup_{1\leq j\leq t} \Lambda(\Psi_j^{(j-1)}) a^{-j/r}}{1 - a^{1/r}} d(\Psi_1(y), y).$$

Let us prove that $\sup_{1\leq j\leq t} \Lambda(\Psi_j^{(j-1)}) a^{-j/r}$ satisfies **(EAS)**. It is implied by the Borel-Cantelli Lemma from the convergence of the series

$$\Big(\mathbb{P}\Big(\sup_{1\leq j\leq t} \log \Lambda(\Psi_j^{(j-1)}) - j \log(a^{1/r}) \geq t\varepsilon\Big)\Big).$$

By subadditivity, assuming that $t/r \in \mathbb{N}$ for convenience, we have

$$\mathbb{P}(\sup_{1\leq j\leq t} \log \Lambda(\Psi_j^{(j-1)}) - j \log(a^{1/r}) \geq t\varepsilon) \leq \mathbb{P}\Big(\sup_{1\leq j\leq t/r} \sum_{j=1}^{t/r} \log \Lambda(\Psi_{jr}^{(r)}) - \frac{t}{r} \log(a) \geq t\varepsilon\Big).$$

This series is convergent by an application of Theorem 1 in [28] as $\log(\Lambda(\Psi_t))$ has a finite moment of order $p > 2$ and is strongly mixing with geometric rate. Finally, for any non negative sequence $W_t \xrightarrow{e.a.s.} 0$ we have that the series $(W_t \sup_{1\leq j\leq t} \Lambda(\Psi_j^{(j-1)}) a^{-j/r})$ is convergent and the desired result follows. □

We are now ready to state the following refinement of the theorem 2.10 of [31] used to prove that the functions $\hat{g}_t$ and g_t are two times continuously differentiable in a neighborhood of θ_0:

Theorem 8. *Let B be a separable Banach space and (Ψ_t) be a stationary ergodic sequence of Lipschitz maps from B into B that is strongly mixing with geometric rate. Assume that*

S: *$(\|\Psi_t(0)\|)$ satisfies* **(EAS)**, *$\mathbb{E}[(\log^+ \Lambda(\Psi_0))^p] < \infty$ for some $p > 2$ and $\mathbb{E}[\log \Lambda(\Psi_0^{(r)})] < 0$ for some $r \geq 1$.*

Let $(\hat{\Psi}_t)_{t\in\mathbb{N}}$ be a sequence of Lipschitz maps such that

S': *$\|\hat{\Psi}_t(0) - \Psi_t(0)\| \xrightarrow{e.a.s.} 0$ and $\Lambda(\hat{\Psi}_t - \Psi_t) \xrightarrow{e.a.s.} 0$ as $t \to \infty$.*

Then the unique stationary solution Y_t of the SRE $Y_{t+1} = \Psi_t(Y_t)$ for all $t \in \mathbb{Z}$ exists, $(\|Y_t\|)$ satisfies **(EAS)** *and for every solution $(\hat{Y}_t)_{t\in\mathbb{N}}$ of the perturbed SRE $\hat{Y}_{t+1} = \hat{\Psi}_t(\hat{Y}_t)$ for all $t \geq 1$, one has that $\|\hat{Y}_t - Y_t\| \xrightarrow{e.a.s.} 0$.*

Proof of Theorem 8. We assume the same conditions than in the theorem 2.10 of [31] except that the conditions $\mathbb{E}[\log^+ \|\Psi_0(0)\|] < \infty$ and $\mathbb{E}[\log^+ \|Y_0\|] < \infty$ do not hold. A careful look at the proof of Theorem 2.10 in [31] shows that these conditions are assumed such that the series $(W_t\|\Psi_t(0)\|)$ and $(W_t\|Y_t\|)$ are convergent. The first series is convergent by assumption, the second one is convergent as $(\|Y_t\|)$ satisfies **(EAS)** from the last assertion of Theorem 7. □

Let us come back to the proof of Lemma 2. First note that as (σ_t^2) is strongly mixing with geometric rate it is also the case of the process

$$(\phi_t(x, \theta)) = (\alpha + \beta x + (\gamma X_t + \delta|X_t|) \exp(-x/2)).$$

Deriving the SRE (3), we obtain new linear SREs satisfied by the first derivates of $\hat{g}_t$:

$$\nabla \hat{g}_{t+1}(\theta) = \phi_t'(\hat{g}_t(\theta), \theta)\nabla \hat{g}_t(\theta) + \nabla_\theta \phi_t(\hat{g}_t(\theta), \theta) \quad \forall t \geq 0 \tag{18}$$

where

$$\phi_t'(x, \theta) = \beta - 2^{-1}(\gamma X_t + \delta|X_t|) \exp(-x/2),$$
$$\nabla_\theta \phi_t(x, \theta) = (1, x, X_t \exp(-x/2), |X_t| \exp(-x/2))'.$$

Note that both functions $\phi_t'(x, \theta)$ and $\nabla_\theta \phi_t(x, \theta)$ are continuous in (x, θ). Under **(CI)**, we assert from an application of Theorem 2 that there exists a compact neighborhood $\mathcal{V}(\theta_0) \subset \overset{\circ}{\Theta}$ of θ_0 such that $\mathbb{E}[\log \|\Lambda(\phi_0^{(r)})\|_{\mathcal{V}(\theta_0)}] < 0$ for some $r \geq 1$. The SRE (18) can be written as a linear perturbed SRE $\nabla \hat{g}_{t+1} = \hat{\Phi}_t'(\nabla \hat{g}_t)$ on the Banach space of continuous functions on $\mathcal{V}(\theta_0)$. The sequence $(\hat{\Phi}_t')$ is not stationary but it is well approximated by the stationary sequences (Φ_t') defined as, for all $t \geq 0$:

$$\Phi_t'(h) := \phi_t'(g_t, \cdot)h + \nabla_\theta \phi_t(g_t, \cdot).$$

We will apply Theorem 8 on the SREs driven by $(\hat{\Phi}_t')$ and (Φ_t') to assert the existence of the firs derivatives (∇g_t). Let us first check the conditions **S** of Theorem 8 on Φ_t'. The series $(\|\Phi_t'(0)\|_{\mathcal{V}(\theta_0)}) = (\|\nabla_\theta \phi_t(g_t, \cdot)\|_{\mathcal{V}(\theta_0)})$ satisfies **(EAS)** from $\mathbb{E}\log^+ |X_t| < \infty$ and because $(\|\hat{g}_t\|_{\mathcal{V}(\theta_0)})$ satisfies **(EAS)** by an application of Theorem 7 on (g_t). By definition of $\mathcal{V}(\theta_0)$, as $\phi_t'(x, \theta) \leq \Lambda(\phi_0, \theta)$, the last two conditions of **S** are implied by the continuous invertibility of the model and by the fact that $\mathbb{E}(\log^+ X_0)^p < \infty$ for $p = 8$ as $\mathbb{E}Z_0^4 < \infty$.

Let us check Assumption **S'** of the theorem 8 on $\hat{\Phi}_t' - \Phi_t'$. As ϕ_t is two times continuously differentiable, we have

$$\|\Phi_t'(0) - \hat{\Phi}_t'(0)\| = \|\nabla_\theta \phi_t(g_t, \cdot) - \nabla_\theta \phi_t(\hat{g}_t, \cdot)\|_{\mathcal{V}(\theta_0)} \leq \Big\| \sup_{x\geq c} \nabla_\theta \phi_t' \Big\|_{\mathcal{V}(\theta_0)} \|g_t - \hat{g}_t\|_{\mathcal{V}(\theta_0)}$$

where $\nabla_\theta \phi_t'(x,\theta) = (0, 1, X_t/2\exp(-x/2), |X_t|/2\exp(-x/2))'$ and $c := \min_{\mathcal{V}(\theta_0)} \alpha/(1-\beta)$. Remark that $\mathbb{E}[\log^+ \| \sup_{x\geq c} \nabla_\theta \phi_t'\|_{\mathcal{V}(\theta_0)}] < \infty$ because $\mathbb{E}[\log^+ |X_t|] < \infty$. Thus

$$\sum_{t=0}^{\infty} \mathbb{P}\Big(\Big\| \sup_{x\geq c} \nabla_\theta \phi_t \Big\|_{\mathcal{V}(\theta_0)} \geq \rho^{-t} \Big) < \infty$$

for any $\rho > 1$ and then $\|\Phi_t'(0) - \hat{\Phi}_t'(0)\| \xrightarrow{e.a.s.} 0$ by an application of the Borel-Cantelli Lemma because $\|g_t - \hat{g}_t\|_{\mathcal{V}(\theta_0)} \xrightarrow{e.a.s.} 0$.

To check the second condition of **S'**, we remark that as ϕ_t is two times continuously differentiable, we also have

$$\Lambda(\Phi_t' - \hat{\Phi}_t') \leq \|\phi_t'(\hat{g}_t, \cdot) - \phi_t'(g_t, \cdot)\|_{\mathcal{V}(\theta_0)} \leq \| \sup_{x\geq c} \phi_t''\|_{\mathcal{V}(\theta_0)} \|g_t - \hat{g}_t\|_{\mathcal{V}(\theta_0)}$$

where $\phi_t''(x,\theta) = 4^{-1}(\gamma X_t + \delta |X_t|)\exp(-x/2)$. That $\|\Phi_t'(0) - \hat{\Phi}_t'(0)\| \xrightarrow{e.a.s.} 0$ follows again by an application of the Borel-Cantelli Lemma as $\|g_t - \hat{g}_t\|_{\mathcal{V}(\theta_0)} \xrightarrow{e.a.s.} 0$ and $\sum_{t=0}^{\infty} \mathbb{P}(\| \sup_{x\geq c} \phi_t''\|_{\mathcal{V}(\theta_0)} \geq \rho^{-t}) < \infty$ because $\mathbb{E}[\log^+ \| \sup_{x\geq c} \phi_t''\|_{\mathcal{V}(\theta_0)}] < \infty$ due to $\mathbb{E}[\log^+ |X_t|] < \infty$. An application of Theorem 8 yields to the existence of the stationary solution (∇g_t) of the SRE driven by (Φ_t') such that $(\|\nabla g_t\|_{\mathcal{V}(\theta_0)})$ satisfies **(EAS)**.

Deriving a second time and keeping the same notation than above we obtain another linear SRE satisfied by $(\mathbb{H}\hat{g}_{t+1}(\theta))$ for any $1 \leq i, j \leq d$

$$\begin{aligned} \mathbb{H}\hat{g}_{t+1}(\theta) = {} & \phi_t'(\hat{g}_t(\theta),\theta)\mathbb{H}\hat{g}_t(\theta) + \phi_t''(\hat{g}_t(\theta),\theta)\nabla\hat{g}_t(\theta)\nabla\hat{g}_t(\theta)^T \\ & + \nabla_\theta\phi_t'(\hat{g}_t(\theta),\theta)\nabla\hat{g}_t(\theta)^T + \nabla\hat{g}_t(\theta)\nabla_\theta\phi_t'(\hat{g}_t(\theta),\theta)^T + \mathbb{H}_\theta\phi_t(\hat{g}_t(\theta),\theta). \end{aligned} \tag{19}$$

For the EGARCH(1,1) model $\mathbb{H}_\theta\phi_t(x,\theta)$ is identically null. Thus we consider the perturbed SRE

$$\hat{\Phi}_t''(h) := \phi_t''(\hat{g}_t,\cdot)h + \phi_t''(\hat{g}_t,\cdot)\nabla\hat{g}_t\nabla g_t^T + \nabla_\theta\phi_t'(\hat{g}_t,\cdot)\nabla g_t^T + \nabla\hat{g}_t\nabla_\theta\phi_t'(\hat{g}_t,\cdot)^T.$$

We can apply Theorem 8 on this perturbed SRE and the corresponding stationary SRE

$$\Phi_t''(h) := \phi_t''(g_t,\cdot)h + \phi_t''(g_t,\cdot)\nabla g_t\nabla g_t^T + \nabla_\theta\phi_t'(g_t,\cdot)\nabla g_t^T + \nabla g_t\nabla_\theta\phi_t'(g_t,\cdot)^T.$$

The details of the proof are omitted as they are similar than those use above for the first derivative. Note that $(\|\nabla g_t\|_{\mathcal{V}(\theta_0)})$ satisfies **(EAS)** is required to apply Theorems 7 and 8 in order that the first condition in **S** is satisfied for $\Phi_t''(0)$ as it is a function of ∇g_t for any $t \geq 1$.

6.3. **Proof of Lemma 3.** Let us fix $\mathcal{V}(\theta_0)$ as in the proof of Lemma 2 such that g_t is two times continuously invertible on $\mathcal{V}(\theta_0)$ and such that the uniform log moments on the derivatives exist. Because $\|\tilde{\theta}_t - \theta_0\| \xrightarrow{a.s.} 0$ there exists some random integer $M \geq 1$ such that $\tilde{\theta}_t \in \mathcal{V}(\theta_0)$ for any $t \geq M$. Consider the SRE

$$\nabla g_{t+1}(\theta) = \phi_t'(g_t(\theta),\theta)\nabla g_t(\theta) + \nabla_\theta\phi_t(g_t(\theta),\theta) \quad \forall t \in \mathbb{Z}$$

where

$$\begin{aligned} \phi_t'(x,\theta) &= \beta - 2^{-1}(\gamma Z_t + \delta|Z_t|)\exp(-(x - g_t(\theta_0))/2), \\ \nabla_\theta\phi_t(x,\theta) &= (1, x, Z_t\exp(-(x - g_t(\theta_0))/2), |Z_t|\exp(-(x - g_t(\theta_0))/2))'. \end{aligned}$$

As $x \to \exp(-x/2)$ is a Lipschitz continuous function for $x \geq c$, as $\mathcal{V}(\theta_0)$ is a compact set there exists some $C > 0$ such that for all $x \in \mathcal{K}$, all $\theta \in \mathcal{V}(\theta_0)$ we have

$$|\phi_t'(x,\theta) - \phi_t'(g_t(\theta_0),\theta_0)| + \|\nabla_\theta\phi_t(x,\theta) - \nabla_\theta\phi_t(g_t(\theta_0),\theta_0)\| \leq C|Z_t|(\|\theta - \theta_0\| + |x - g_t(\theta_0)|).$$

Thus, for any $n \geq t \geq M$, denoting $v_t(\tilde{\theta}_n) = \|\nabla g_{t+1}(\hat{\theta}_n) - \nabla g_{t+1}(\theta_0)\|$ we obtain

$$\begin{aligned} v_{t+1}(\tilde{\theta}_n) \leq & |\phi_t'(g_t(\theta_0),\theta_0)| v_t(\tilde{\theta}_n) + \|\nabla g_t(\hat{\theta}_n)\| |\phi_t'(g_t(\theta_0),\theta_0) - \phi_t'(\hat{g}_t(\hat{\theta}_n),\hat{\theta}_n)| \\ & + \|\nabla_\theta \phi_t(\hat{g}_t(\hat{\theta}_n),\hat{\theta}_n) - \nabla_\theta \phi_t(g_t(\theta_0),\theta_0)\| \\ \leq & |\phi_t'(g_t(\theta_0),\theta_0)| v_t(\tilde{\theta}_n) + (\|\nabla g_t\|_{\mathcal{V}(\theta_0)} + 1) C |Z_t| (\|\tilde{\theta}_n - \theta_0\| + |g_t(\tilde{\theta}_n) - g_t(\theta_0)|). \end{aligned}$$

By a recursive argument, we obtain for all $n \geq t \geq M$

$$\begin{aligned} v_t(\tilde{\theta}_n) \leq \sum_{j=M}^{t-1} \prod_{i=j+1}^{t-1} |\phi_i'(g_i(\theta_0),\theta_0)| (\|\nabla g_t\|_{\mathcal{V}(\theta_0)} + 1) C |Z_t| (\|\tilde{\theta}_n - \theta_0\| + |g_t(\tilde{\theta}_n) - g_t(\theta_0)|) \\ + \prod_{i=M}^{t-1} |\phi_i'(g_i(\theta_0),\theta_0)| v_M. \end{aligned}$$

That $\prod_{i=j+1}^{t-1} |\phi_i'(g_i(\theta_0),\theta_0)| \xrightarrow{e.a.s.} 0$ follows from the assumption **(MM)** $\mathbb{E}\phi_i'(g_i(\theta_0),\theta_0)^2 < 1$ and by using the subadditive ergodic theorem of [22] on the logarithms. Thus the last term of the sum converges e.a.s to 0 and the corresponding Cesaro mean $c_n = n^{-1} \sum_{t=M}^{n} \prod_{i=M}^{t-1} |\phi_i'(g_i(\theta_0),\theta_0)| v_M$ also converges e.a.s. to 0.

Let us treat the term

$$\sum_{j=M}^{t-1} \prod_{i=j+1}^{t-1} |\phi_i'(g_i(\theta_0),\theta_0)| (\|\nabla g_t\|_{\mathcal{V}(\theta_0)} + 1) C |Z_t| \|\tilde{\theta}_n - \theta_0\|.$$

It is a.s. smaller than

$$\|\tilde{\theta}_n - \theta_0\| \sum_{j=M}^{\infty} \prod_{i=j+1}^{t-1} |\phi_i'(g_i(\theta_0),\theta_0)| (\|\nabla g_t\|_{\mathcal{V}(\theta_0)} + 1) C |Z_t|.$$

Indeed, this series is a.s. convergent because $(\|\nabla g_t\|_{\mathcal{V}(\theta_0)})$ satisfies **(EAS)** from Theorem 8. Thus we obtain that there exist some random variable $a > 0$ such that

$$n^{-1} \sum_{t=M}^{n} \sum_{j=M}^{t-1} \prod_{i=j+1}^{t-1} |\phi_i'(g_i(\theta_0),\theta_0)| (\|\nabla g_t\|_{\mathcal{V}(\theta_0)} + 1) C |Z_t| \|\tilde{\theta}_n - \theta_0\| \leq a \|\tilde{\theta}_n - \theta_0\|.$$

Finally, the reminding term of the upper bound is

$$\sum_{j=M}^{t-1} \prod_{i=j+1}^{t-1} |\phi_i'(g_i(\theta_0),\theta_0)| (\|\nabla g_t\|_{\mathcal{V}(\theta_0)} + 1) C |Z_t| |g_t(\tilde{\theta}_n) - g_t(\theta_0)|.$$

We treat it as in the proof of Theorem 3. Indeed, its uniform norm converges a.s. by an application of similar arguments than above. Thus there exists a random continuous function b satisfying $b(\theta_0) = 0$ and

$$\begin{aligned} n^{-1} \sum_{t=M}^{n} \sum_{j=M}^{t-1} \prod_{i=j+1}^{t-1} & |\phi_i'(g_i(\theta_0),\theta_0)| (\|\nabla g_t\|_{\mathcal{V}(\theta_0)} + 1) C |Z_t| |g_t(\tilde{\theta}_n) - g_t(\theta_0)| \\ & \leq \sum_{j=M}^{\infty} \prod_{i=j+1}^{t-1} |\phi_i'(g_i(\theta_0),\theta_0)| (\|\nabla g_t\|_{\mathcal{V}(\theta_0)} + 1) C |Z_t| |g_t(\tilde{\theta}_n) - g_t(\theta_0)| \\ & \leq b(\tilde{\theta}_n). \end{aligned}$$

Finally we obtain that $n^{-1}\sum_{t=M}^n v_t(\tilde\theta_n) \le a\|\tilde\theta_n-\theta_0\| + b(\tilde\theta_n) + c_n$. Using this bound and the SRE satisfied by the differences $\mathbb{H}g_t(\tilde\theta_n) - \mathbb{H}g_t(\theta_0)$ we obtain following similar reasoning that

$$\frac{1}{n}\sum_{t=M}^n \|\mathbb{H}g_t(\tilde\theta_n) - \mathbb{H}g_t(\theta_0)\| \le a'\|\tilde\theta_n-\theta_0\| + b'(\tilde\theta_n) + c_n',$$

where a' is a positive r.v., b' is a random continuous function satisfying $b'(\theta_0)=0$ and $c_n' \xrightarrow{e.a.s.} 0$. We then conclude that for any possible value of M we have that

$$n^{-1}\sum_{t=M}^n \|\mathbb{H}g_t(\tilde\theta_n) - \mathbb{H}g_t(\theta_0)\| \xrightarrow{a.s.} 0. \tag{20}$$

Let us fix $\delta>0$. It remains to estimate

$$\mathbb{P}\Big(\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^M \|\mathbb{H}g_t(\tilde\theta_n) - \mathbb{H}g_t(\theta_0)\| = 0\Big) = \sum_{k=1}^\infty \mathbb{P}\Big(\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^k \|\mathbb{H}g_t(\tilde\theta_n) - \mathbb{H}g_t(\theta_0)\| = 0\Big)\mathbb{P}(M=k).$$

By continuity of the second derivative $\mathbb{H}g_t$ we have that $\mathbb{H}g_t(\tilde\theta_n)\to\mathbb{H}g_t(\theta_0)$ a.s. We deduce that for any $k\ge 1$ we have

$$\mathbb{P}\Big(\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^k \|\mathbb{H}g_t(\tilde\theta_n) - \mathbb{H}g_t(\theta_0)\| = 0\Big) = 1.$$

The desired result follows easily.

Acknowledgments. I would like to thank C. Francq and J.-M. Zakoïan for helpful discussions and for pointing out some mistakes in an earlier version.

References

[1] Alsmeyer, G. (2003) On the Harris recurrence and iterated random Lipschitz fonctions and related convergence rate results. *J. Theor. Probab.* **16**, 217–247.

[2] Bardet, J. and Wintenberger, O. (2009) Asymptotic normality of the quasi maximum likelihood estimator for multidimensional causal processes, *Ann. Statist.* **37**, 2730–2759.

[3] Basrak, B., Davis, R.A. and Mikosch. T. (2002) Regular variation of GARCH processes. *Stoch. Proc. Appl.* **99**, 95–116.

[4] Berkes, I. and Horvth, L. (2004) The efficiency of the estimators of the parameters in GARCH processes.', Ann. Statist. 32(2), 633–655.

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

[6] Billingsley, P. (1999) *Convergence of Probability Measures*, John Wiley and Sons Inc., New York.

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

[8] Bougerol, P. (1993) Kalman filtering with random coefficients and contractions. *SIAM J. Control and Optimization* **31**, 942–959.

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

[10] Brandt, M. W. and Jones, C. S. (2006) Volatility forecasting with range-based EGARCH models *Journal of Business and Economic Statistics* **24**, 470–486.

[11] Cont, R. (2001) Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance* **1**, 223–236.

[12] Demos, A. and Kyriakopoulou, D. (2010) Bias correction of ML and QML estimators in the EGARCH(1,1) model. Preprint.

[13] Elton, J. H. (1990) A multiplicative ergodic theorem for Lipschitz maps. *Stoch. Proc. Appl.* **34**, 39–47.

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

[15] Francq, C., Wintenberger, O. et Zakoïan, J.-M. (2012) GARCH models without positivity constraints: Exponential or Log GARCH? Preprint.
[16] Francq, C. et Zakoïan, J.-M. (2010) *GARCH Models - Structure, Estimation and Finance Applications*. John Wiley.
[17] Francq, C. and Zakoïan, J. (2004), Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* **10**, 605–637.
[18] Granger, C. and Andersen, A. (1978) On the invertibility on time series models. *Stoch Proc. and their Appl.* **8**, 87–92.
[19] Harvey, A. C. and Chakravarty, T. (2008) Beta-t-(E)GARCH. Preprint.
[20] He, C.; Terasvirta, T. and Malmsten, H. (2002) Moment structure of a family of first-order exponential GARCH models, *Econom. Theory* **18(04)**, 868–885.
[21] Jeantheau, T. (1998) Strong consistency of estimation for multivariate ARCH models. *Econom. Theory* **14**, 70–86.
[22] Kingman, J.F.C. (1973) Subadditive ergodic theory. *Ann. Probab.*, **6**, 883–909.
[23] Krengel, U. (1985) *Ergodic Theorems*. de Gruyter, Berlin.
[24] Letac, G. (1986) A contraction principle for certain Markov chains and its applications. random matrices and their applications. *Proceedings of AMS-IMS-SIAM Joint Summer Research Conference 1984* **50**, 263–273.
[25] Mikosch, T. and Rezapour, M. (2012) Stochastic volatility models with possible extremal clustering. *Bernoulli* To appear.
[26] Nelson, D. B. (1990) Stationarity and persistence in the GARCH(1,1) model. *Econom. Theory* **6**, 318-334.
[27] Nelson, D. B. (1991), Conditional Heteroskedasticity in Asset Returns : A New Approach. *Econometrica* **59**, 347–370.
[28] Shao, Q-M. (1993). Complete convergence for α-mixing sequences. *Statist. Probab. Letters* **16**, 279–287.
[29] Sorokin, A. (2011) Non-invertibility in some heteroscedastic models. Arxiv preprint 1104.3318.
[30] Straumann, D. (2005) *Estimation in Conditionally Heteroscedastic Time Series Models*. Springer, Vol. 181, New York.
[31] Straumann, D. and Mikosch, T. (2006) Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: a stochastic recurrence equation approach *Ann. Statist.* **34**, 2449–2495.
[32] Sucarrat, G. and Escribano, A. (2010) The power log-GARCH model. Preprint.
[33] Tong, H. (1993) *Non-Linear Time Series, A Dynamical System Approach* Oxford Science Publications, Oxford.
[34] Tweedie, R.L. (1998) Invariant measures for Markov chains with no irreducibility Assumptions. *Journal of Applied Probability* **25**, 275–285.
[35] Vrontos, I. D.; Dellaportas, P. and Politis, D. N. (2000) Full Bayesian Inference for GARCH and EGARCH Models. *Journal of Business and Economic Statistics* **18(2)**, 187–98.
[36] Zaffaroni, P. (2009) Whittle estimation of EGARCH and other exponential volatility models *J. Econometrics* **151**, 190–200.

Olivier Wintenberger, Université de Paris-Dauphine, Centre De Recherche en Mathématiques de la Décision UMR CNRS 7534, Place du Maréchal De Lattre De Tassigny, 75775 Paris Cedex 16, France
E-mail address: owintenb@ceremade.dauphine.fr