CREST
CENTER FOR RESEARCH IN ECONOMICS AND STATISTICS

WORKING PAPER SERIES

Iterated Function Systems driven by non independent sequences: structure and inference

Baye Matar Kandji

n°3 / January 2022

CREST
Center for Research in Economics and Statistics
UMR 9194

5 Avenue Henry Le Chatelier
TSA 96642
91764 Palaiseau Cedex
FRANCE

Phone: +33 (0)1 70 26 67 00
Email: info@crest.science
https://crest.science/

Iterated Function Systems driven by non independent sequences: structure and inference

Baye Matar Kandji*[1]

[1]CREST, ENSAE, Institut Polytechnique de Paris.

January 26, 2022

Abstract

The paper investigates the existence of a strictly stationary solution to an Iterated Function System (IFS) driven by a stationary and ergodic sequence. When the driving sequence is not independent, the strictly stationary solution may admit no moment but we show an exponential control of the trajectories. We exploit these results to prove, under mild conditions, the consistency of the quasi-maximum likelihood estimator of GARCH models with non independent innovations.

Keywords: Stochastic Recurrence Equation, Semi-strong GARCH, Quasi Maximum Likelihood, inference without moments.

*5 Avenue Henri Le Chatelier, 91120 Palaiseau, France. E-Mail: bayematar.kandji@ensae.fr

1 Introduction

Since Kesten (1973), the study of the theoretical properties of the Stochastic Recurrence Equation (SRE) $\boldsymbol{X}_t = \boldsymbol{A}_t \boldsymbol{X}_{t-1} + \boldsymbol{B}_t$ has received much attention. This equation gathers a large class of classical econometric processes such as the GARCH and ARMA models, and their numerous variants. Brandt (1986) proposes a sufficient condition of existence and uniqueness of a strictly stationary solution in the case where $(\boldsymbol{A}_t, \boldsymbol{B}_t)_t$ is stationary and ergodic. Under a certain irreducibility condition, Bougerol and Picard (1992a) establish that this condition is necessary in the independent and identically distributed (iid) case and deduce a necessary and sufficient condition for the existence of a unique stationary solution of a general GARCH(p,q) model (see Bougerol and Picard (1992b)). The probabilistic properties of the stationary solution of SRE model in the iid case are well known. In the scalar case, Kesten (1973) shows that $\mathbb{P}(\pm \boldsymbol{X}_1 > x) \sim c_{\pm} x^{-a}, \quad x \to \infty$. A thorough study of these SRE models, in particular their tail behavior, is presented in Buraczewski et al. (2016) and the references therein. The SRE model is the affine mapping particular case of the so-called Stochastic Iterated Function Systems (IFS) $\boldsymbol{X}_t = \boldsymbol{\Psi}(\boldsymbol{\theta}_t, \boldsymbol{X}_{t-1})$. Most of the theoretical properties established for SRE models (stationary, tail properties) can be extended to IFS equations.

In recent years, the iid assumption on the innovations of the econometric models is often replaced by a less restrictive martingale difference assumption. See Escanciano (2009) for the classical GARCH(p,q) model or Francq and Thieu (2019) and Han and Kristensen (2014) for GARCH-X models. This amounts to study an IFS equation driven by non iid innovations. To our knowledge, all existing works on the inference of IFS models assume the existence of a small-order moment of the observed process. However, stationary IFS equations with non-iid innovations may not admit any finite moment. The aim of this paper is to establish that the stationary trajectories of the IFS equations enjoy an exponential control property. We also show that this properties is sufficient to establish the consistency of the Quasi-Maximum Likelihood estimator (QMLE) of semi-strong GARCH models.

The rest of the paper is organized as follows. In Section 2 we present our main result. Section 3 is devoted to proofs and Section 4 investigates the estimation of the semi-strong GARCH(p,q) model.

2 Stochastic IFS without moments

Let $(E, \mathcal{E})$ be a measurable space and (F, d) a complete and separable metric space (Polish space). Let $(\theta_t)_{t\in\mathbb{Z}}$ be a stationary and ergodic process valued in E, and let $\mathbf{\Psi} : E \times F \to F$ a function such that $x \mapsto \mathbf{\Psi}(\theta, x)$ is Lipschitz continuous for all $\theta \in E$. Let

$$\mathbf{\Lambda}_t = \mathbf{\Lambda}(\mathbf{\Psi}_t) = \sup_{x_1, x_2 \in F, x_1 \neq x_2} \frac{d\left(\mathbf{\Psi}_t\left(x_1\right), \mathbf{\Psi}_t\left(x_2\right)\right)}{d\left(x_1, x_2\right)}$$

where $\mathbf{\Psi}_t = \mathbf{\Psi}(\theta_t, \cdot)$. Let $\mathbf{\Lambda}_t^{(0)} = 1$ and $\mathbf{\Lambda}_t^{(r)} = \mathbf{\Lambda}(\mathbf{\Psi}_t \circ \cdots \circ \mathbf{\Psi}_{t-r+1})$ for all $r > 0$.

Consider the IFS model,

$$\boldsymbol{X}_t = \mathbf{\Psi}\left(\theta_t, \boldsymbol{X}_{t-1}\right) = \mathbf{\Psi}_t\left(\boldsymbol{X}_{t-1}\right), \qquad \text{for all } t \in \mathbb{Z}. \tag{2.1}$$

The following result is due to Bougerol (1993, Theorem 2.8), see also Straumann and Mikosch (2006, Theorem 2.8).

Theorem 2.1. *Assume the following conditions hold: (i) there exists a constant $c \in F$ such that $\mathbb{E} \ln^+ d\left(\mathbf{\Psi}_0(c), c\right) < \infty$, (ii) $\mathbb{E} \ln^+ \mathbf{\Lambda}_0 < \infty$ and (iii) $\lim_{r\to\infty} \frac{1}{r} \ln \mathbf{\Lambda}_0^{(r)} < 0$ a.s. Then there exists a unique stationary (and ergodic) solution $(\boldsymbol{X}_t)_{t\in\mathbb{Z}}$ to the equation (2.1).*

Moreover we have:

$$\textit{for all } \ t \in \mathbb{Z}, \quad d(\boldsymbol{X}_t, c) \leq \sum_{n=0}^{\infty} \mathbf{\Lambda}_t^{(n)} d\left(\mathbf{\Psi}_{t-n}(c), c\right) < \infty, \quad a.s \tag{2.2}$$

Note that $(\ln \mathbf{\Lambda}_0^{(r)})_{r>1}$ is a sub-additive sequence. Therefore, by the subadditive ergodic theorem of Kingman (1973), the limit in assumption (iii) exists.

For the reader's convenience and because we have not been able to find Equation (2.2) exactly under this form, we provide a proof for Theorem 2.1 in the appendix.

Remark 2.1. If (θ_t) is iid, it is possible to prove that $d(\boldsymbol{X}_1, c)$ has a power-law tail, see Buraczewski et al. (2016, Theorem 5.3.6). This implies that there exists $s > 0$ such that $\mathbb{E}d(\boldsymbol{X}_1, c)^s < \infty$. This small moment property is often used in the statistical inference of IFS models. For example, it is commonly used to prove the consistency of GARCH models and its derivatives (see Berkes et al. (2003) for GARCH model and Francq et al. (2018) for EGARCH and Log-GARCH model). If (θ_t) is not iid, the example below shows that the stationary solution may not admit any small-order moment.

Example 2.1. *Let $\delta \in (0,1)$ and $(z_t)_{t\in\mathbb{Z}}$ an iid non negative real process where $\mathbb{E}z_t = \frac{1-\delta}{2}$ and $\mathbb{E}z_t^2 = \infty$. The process (θ_t), defined by $\theta_t = \sum_{k=0}^{\infty} \delta^k z_{t-k}$ for all $t \in \mathbb{Z}$ satisfies $\mathbb{E}\theta_t = \frac{1}{2}$ and for all $t \in \mathbb{Z}$, $x_t = 1 + \sum_{k=1}^{\infty} \prod_{j=1}^{k} \theta_{t-j+1}$ exists a.s. Moreover (x_t) is the unique stationary solution of*

$$x_t = \theta_t x_{t-1} + 1. \tag{2.3}$$

Note that $x_t \geq \prod_{j=1}^{k} \theta_{t-j+1} \geq \delta^{\frac{k(k-1)}{2}} (z_{t-k+1})^k$ for all $k \in \mathbb{N}^*$. For all $s > 0$ we thus have $\mathbb{E}x_0^s \geq \mathbb{E}\delta^{\frac{sk(k-1)}{2}} (z_0)^{sk} = \infty$ for k such that $sk > 2$.

We now state our main result, which provides a way to circumvent the non existence of small order moments for models such as that of Example 2.1. Section 4 will be devoted to the statistical study of a class of econometric models where the existence of moments is not guaranteed.

Theorem 2.2. *Under conditions of Theorem 2.1 then: for all $t \in \mathbb{Z}$*

1. $\limsup_{n\to\infty} \frac{1}{n} \ln d(\boldsymbol{X}_{t+n}, c) \leq 0$ *and* *2.* $\limsup_{n\to\infty} \frac{1}{n} \ln d(\boldsymbol{X}_{t-n}, c) \leq 0$ *a.s*

Theorem 2.2 can be interpreted as an exponential control of the trajectory of the stationary solution. Note that the moment properties $\mathbb{E}d(\boldsymbol{X}_1, c)^s < \infty$ for some $s > 0$ implies the results of Theorem 2.2 but the converse is false.

3 Proof of the main result

To show Theorem 2.2, we first define a SRE which bounds the distance between $\boldsymbol{X}_t$ and some point $c \in F$.

Note that by Kingman (1973) $\lim_{r\to\infty} \frac{1}{r} \ln \boldsymbol{\Lambda}_0^{(r)} = \lim_{r\to\infty} \frac{1}{r} \mathbb{E} \ln \boldsymbol{\Lambda}_0^{(r)}$ *a.s.*, so by *iii)* of Theorem 2.1 there exists a positive integer r_0 such that $\mathbb{E} \ln \boldsymbol{\Lambda}_0^{(r_0)} < 0$. It can be shown that $\mathbb{E}\left[\ln\left((\boldsymbol{\Lambda}_0^{(r_0)} + u)\right)\right] \overset{u\to 0}{\to} \mathbb{E} \ln \boldsymbol{\Lambda}_0^{(r_0)}$, see Straumann and Mikosch (2006, proof of Theorem 2.10). Therefore $\exists\ u_0 > 0,\ \ln(u_0) \leq \gamma_0 := \mathbb{E}\left[\ln\left((\boldsymbol{\Lambda}_0^{(r_0)} + u_0)\right)\right] < 0$. We thus have, for all $v \in [\gamma_0, 0)$,

$$\mathbb{E}\left[\ln\left(\delta(v)(\boldsymbol{\Lambda}_0^{(r_0)} + u_0)\right)\right] = v \tag{3.1}$$

with $\delta(v) = \exp(v - \gamma_0) \geq 1$.

Now, for any integer $p \in [0, r_0 - 1]$, define $(\boldsymbol{a}_{p,t}(v), \boldsymbol{b}_{p,t})_{t\in\mathbb{Z}}$ by

$$\boldsymbol{a}_{p,t}(v) = \delta(v)(\boldsymbol{\Lambda}_{r_0 t+p}^{(r_0)} + u_0), \text{ and } \boldsymbol{b}_{p,t} = 1 + \sum_{k=0}^{r_0-1} \boldsymbol{\Lambda}_{r_0 t+p}^{(k)} d\left(\boldsymbol{\Psi}_{r_0 t+p-k}(c), c\right).$$

By Assumptions (i) and (ii) of Theorem 2.1 and by the elementary inequality $\ln\left(\sum_{i=1}^n a_i\right) \leq \ln n + \sum_{i=1}^n \ln^+ a_i$ for non-negative $\{a_i\}_{i=1}^n$, we have $E\ln^+ \boldsymbol{a}_{p,t}(v) < \infty$ and $E\ln^+ \boldsymbol{b}_{p,t}(v) < \infty$. Therefore, in view of (3.1), there exists a unique stationary solution $(\boldsymbol{z}_{p,t}(v))_t$ to the equation

$$\boldsymbol{z}_{p,t}(v) = \boldsymbol{a}_{p,t}(v)\boldsymbol{z}_{p,t-1}(v) + \boldsymbol{b}_{p,t}. \tag{3.2}$$

Note that by Brandt (1986)

$$\boldsymbol{z}_{p,t}(v) = \sum_{q=0}^{\infty} \left(\prod_{i=0}^{q-1} \boldsymbol{a}_{p,t-i}(v) \right) \boldsymbol{b}_{p,t-q}. \tag{3.3}$$

By iterating Equation (3.2) we have

$$\boldsymbol{z}_{p,t}(v) = \sum_{q=0}^{n} \left(\prod_{i=0}^{q-1} \boldsymbol{a}_{p,t-i}(v) \right) \boldsymbol{b}_{p,t-q} + \left(\prod_{i=0}^{n} \boldsymbol{a}_{p,t-i}(v) \right) \boldsymbol{z}_{p,t-(n+1)}(v), \quad \forall n \geq 1. \tag{3.4}$$

By (3.3) and (3.4), $\left(\prod_{i=0}^n \boldsymbol{a}_{p,t-i}(v)\right) \boldsymbol{z}_{p,t-(n+1)}(v)$ is the remainder of a convergent series, hence it almost surely converges to 0. i.e.

$$\left(\prod_{k=0}^{n-1} \boldsymbol{a}_{p,t-k}(v) \right) \boldsymbol{z}_{p,t-n}(v) \overset{n\to\infty}{\to} 0 \quad a.s. \tag{3.5}$$

We now give technical lemmas which make the link between the processes $(\boldsymbol{X}_t)$ and $(\boldsymbol{z}_{p,t}(v))_t$.

Lemma 3.1. *For all $v \in [\gamma_0, 0)$, $0 \leq p \leq r_0 - 1$, and $t \in \mathbb{Z}$, we have*

$$d(\boldsymbol{X}_{r_0 t+p}, c) \leq \boldsymbol{z}_{p,t}(v) \quad a.s \tag{3.6}$$

Proof. of Lemma 3.1:

For any integer n, let q and m the quotient and remainder of the Euclidean division of n by r_0: $n = qr_0 + m$. By sub-multiplicativity we have

$$\boldsymbol{\Lambda}_t^{(n)} \leq \left(\prod_{i=0}^{q-1} \boldsymbol{\Lambda}_{t-ir_0}^{(r_0)} \right) \boldsymbol{\Lambda}_{t-qr_0}^{(m)}, \quad \text{with} \quad \prod_{i=0}^{-1} \boldsymbol{\Lambda}_{t-ir_0}^{(r_0)} = 1.$$

For all $q \in \mathbb{N}$, we then obtain

$$\sum_{n=qr_0}^{(q+1)r_0-1} \boldsymbol{\Lambda}_t^{(n)} d\left(\boldsymbol{\Psi}_{t-n}(c), c\right) \leq \left(\prod_{i=0}^{q-1} \boldsymbol{\Lambda}_{t-ir_0}^{(r_0)} \right) \sum_{m=0}^{r_0-1} \boldsymbol{\Lambda}_{t-qr_0}^{(m)} d\left(\boldsymbol{\Psi}_{t-qr_0-m}(c), c\right).$$

It follows that

$$\sum_{n=0}^{\infty} \boldsymbol{\Lambda}_t^{(n)} d\left(\boldsymbol{\Psi}_{t-n}(c), c\right) = \sum_{q=0}^{\infty} \sum_{n=qr_0}^{(q+1)r_0-1} \boldsymbol{\Lambda}_t^{(n)} d\left(\boldsymbol{\Psi}_{t-n}(c), c\right)$$
$$\leq \sum_{q=0}^{\infty} \left(\prod_{i=0}^{q-1} \boldsymbol{\Lambda}_{t-ir_0}^{(r_0)} \right) \sum_{m=0}^{r_0-1} \boldsymbol{\Lambda}_{t-qr_0}^{(m)} d\left(\boldsymbol{\Psi}_{t-qr_0-m}(c), c\right).$$

Since $\delta(v) \geq 1$ and $u_0 > 0$, we obtain

$$\left(\prod_{i=0}^{q-1} \boldsymbol{a}_{p,t-i}(v) \right) \boldsymbol{b}_{p,t-q} \geq \left(\prod_{i=0}^{q-1} \boldsymbol{\Lambda}_{(r_0t+p)-ir_0}^{(r_0)} \right) \sum_{m=0}^{r_0-1} \boldsymbol{\Lambda}_{(r_0t+p)-qr_0}^{(m)} d\left(\boldsymbol{\Psi}_{(r_0t+p)-qr_0-m}(c), c\right).$$

In view of the last two inequalities, (3.4) and (2.2), we have

$$\boldsymbol{z}_{p,t}(v) \geq \sum_{n=0}^{\infty} \boldsymbol{\Lambda}_{r_0t+p}^{(n)} d\left(\boldsymbol{\Psi}_{r_0t+p-n}(c), c\right) \geq d(\boldsymbol{X}_{r_0t+p}, c),$$

which proofs (3.6), which concludes the proof. □

Let **Aff** denote the set of affine maps from $\mathbb{R}$ into $\mathbb{R}$. Such a map $\boldsymbol{f}_{a,b}$ can be written in

$$\boldsymbol{f}_{a,b}(x) = ax + b, \quad x \in \mathbb{R}, \quad \text{where } (a,b) \in \mathbb{R}^2$$

Note that $(\mathbf{Aff}, \circ)$ is a topological semigroup.

Lemma 3.2. *Let define the function Φ from **Aff** to $\mathbb{R}_+$ by $\Phi(\boldsymbol{f}_{a,b}) = |a| + |b|$.*

1. *For any x, $|x| \geq 1$, $|\boldsymbol{f}_{a,b}(x)| \leq \Phi(\boldsymbol{f}_{a,b})|x|$*
2. *If $|d| \geq 1$ then $\Phi(\boldsymbol{f}_{a,b} \circ \boldsymbol{f}_{c,d}) \leq \Phi(\boldsymbol{f}_{a,b})\Phi(\boldsymbol{f}_{c,d})$*

Since Lemma 3.2 is elementary, its proof is skipped. Note that Φ is the 1-norm in the vector space of affine maps.

Lemma 3.3. *For all $p \in \{0, \ldots, r_0 - 1\}$ and $t \in \mathbb{Z}$, letting $Q_p(t) = r_0t + p$, we have*

1. $\limsup_{n\to\infty} \frac{1}{n} \ln d(\boldsymbol{X}_{Q_p(t+n)}, c) \leq 0$, *2.* $\limsup_{n\to\infty} \frac{1}{n} \ln d(\boldsymbol{X}_{Q_p(t-n)}, c) \leq 0$ *a.s.*

In the previous lemma we distinguished cases 1. and 2. because their proofs are different.

Proof. of Lemma 3.3: Let $\boldsymbol{f}_t$ be the random affine map defined by

$$\boldsymbol{f}_t(x) = \boldsymbol{a}_{p,t}(v)x + \boldsymbol{b}_{p,t}$$

for all $x \in \mathbb{R}$. Define also the backward and forward maps

$$\boldsymbol{\gamma}^B_{t,n} = \boldsymbol{f}_t \circ \boldsymbol{f}_{t-1} \cdots \circ \boldsymbol{f}_{t-n+1} \quad \text{and} \quad \boldsymbol{\gamma}^F_{t,n} = \boldsymbol{f}_{t+n} \circ \boldsymbol{f}_{t+n-1} \cdots \circ \boldsymbol{f}_{t+1}$$

for all $(t,n) \in \mathbb{Z} \times \mathbb{N}^*$. Note that, almost surly

$$\boldsymbol{\gamma}^F_{t,n} = \boldsymbol{\gamma}^B_{t+n,n}, \; \boldsymbol{z}_{p,t}(v) = \boldsymbol{\gamma}^B_{t,n}(\boldsymbol{z}_{p,t-n}(v)) \text{ and } \boldsymbol{z}_{p,t+n}(v) = \boldsymbol{\gamma}^F_{t,n}(\boldsymbol{z}_{p,t}(v)). \quad (3.7)$$

Since $\boldsymbol{b}_{p,t} \geq 1$, by *2.)* of Lemma 3.2

$$(\boldsymbol{u}^B_{t,n})_n := (\ln \Phi(\boldsymbol{\gamma}^B_{t,n}))_n \quad \text{and} \quad (\boldsymbol{u}^F_{t,n})_n := (\ln \Phi(\boldsymbol{\gamma}^F_{t,n}))_n \quad a.s \qquad (3.8)$$

are sub-additive sequences. By argument already used, we have $\mathbb{E}|\ln \Phi(\boldsymbol{\gamma}^B_{t,1})| = \mathbb{E}|\ln \Phi(\boldsymbol{\gamma}^F_{t,1})| = \mathbb{E}|\ln \Phi(\boldsymbol{f}_t)| < \infty$. In view of (3.7) and *1.* of Lemma 3.2,

$$\limsup_{n\to\infty} \frac{1}{n} \ln \boldsymbol{z}_{p,t+n}(v) \leq \limsup_{n\to\infty} \frac{1}{n} \boldsymbol{u}^F_{t,n} + \limsup_{n\to\infty} \frac{1}{n} \ln \boldsymbol{z}_{p,t}(v) \quad a.s.$$

Because $\boldsymbol{z}_{p,t}(v)$ does not depend on n, we have $\limsup\limits_{n\to\infty} \frac{1}{n} \ln \boldsymbol{z}_{p,t}(v) = 0$ $a.s.$ Therefore

$$\limsup_{n\to\infty} \frac{1}{n} \ln \boldsymbol{z}_{p,t+n}(v) \leq \limsup_{n\to\infty} \frac{1}{n} \boldsymbol{u}^F_{t,n} \quad a.s. \qquad (3.9)$$

Since for any $n \in \mathbb{N}^*$, $\boldsymbol{u}^B_{t,n}$ and $\boldsymbol{u}^F_{t,n}$ have the same law, by (3.8) and Kingman sub-additive ergodic theorem,

$$\limsup_{n\to\infty} \frac{1}{n} \boldsymbol{u}^F_{t,n} = \limsup_{n\to\infty} \frac{1}{n} \mathbb{E}\boldsymbol{u}^B_{t,n} = \limsup_{n\to\infty} \frac{1}{n} \boldsymbol{u}^B_{t,n} \quad a.s. \qquad (3.10)$$

On the other hand, in view of (3.4), we have by positivity of the coefficients,

$$\Phi(\boldsymbol{\gamma}^B_{t,n+1}) = \sum_{q=0}^{n} \left(\prod_{i=0}^{q-1} \boldsymbol{a}_{p,t-i}(v) \right) \boldsymbol{b}_{p,t-q} + \left(\prod_{i=0}^{n} \boldsymbol{a}_{p,t-i}(v) \right) \overset{n\to\infty}{\to} \boldsymbol{z}_{p,t}(v) \quad a.s.$$

Therefore

$$\lim_{n\to\infty} \boldsymbol{u}^B_{t,n} = \ln \boldsymbol{z}_{p,t}(v) \quad a.s,$$

which entails

$$\limsup_{n\to\infty} \frac{1}{n} \boldsymbol{u}^B_{t,n} = 0 \quad a.s. \qquad (3.11)$$

By (3.9), (3.10) and (3.11) we get

$$\limsup_{n\to\infty} \frac{1}{n} \ln \boldsymbol{z}_{p,t+n}(v) \le 0 \quad a.s$$

which implies, by Equation (3.6), Point 1. of the lemma

For the second point, by (3.6), (3.5), (3.1) and the ergodic theorem, we have

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \ln d(\boldsymbol{X}_{Q_p(t-n)}, c) &\le \limsup_{n\to\infty} \frac{1}{n} \ln \boldsymbol{z}_{p,t-n}(v) \\
&\le \limsup_{n\to\infty} \frac{1}{n} \ln \left(\prod_{i=0}^{n-1} \boldsymbol{a}_{p,t-i}(v) \right) \boldsymbol{z}_{p,t-n}(v) \\
&\quad - \liminf_{n\to\infty} \frac{1}{n} \ln \left(\prod_{i=0}^{n-1} \boldsymbol{a}_{p,t-i}(v) \right) \\
&\le -\, v \ a.s.
\end{aligned}$$

for all $v \in [\gamma_0, 0)$. Letting $v \to 0^-$ we get the result. □

We are now ready to prove Theorem 2.2

Proof of Theorem 2.2. for all $t \in \mathbb{Z}$, let $t' \in \mathbb{Z}$ and p', $0 \le p' \le r_0 - 1$ such that $t = r_0 t' + p'$. Noting that

$$\{t + k, k \in \mathbb{N}\} \subset \bigcup_{0 \le p \le r_0 - 1} \{r_0(t' + k) + p, k \in \mathbb{N}\}$$

so by the previous relation and the first point of Lemma 3.3 we have

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \ln d(\boldsymbol{X}_{t+n}, c) &\le \max_{0 \le p \le r_0 - 1} \left(\limsup_{n\to\infty} \frac{1}{Q_p(t' + n)} \ln d(\boldsymbol{X}_{Q_p(t'+n)}, c) \right) \\
&\le C \max_{0 \le p \le r_0 - 1} \left(\limsup_{n\to\infty} \frac{1}{n} \ln d(\boldsymbol{X}_{Q_p(t'+n)}, c) \right) \le 0,
\end{aligned}$$

for $C = \max_{0 \le p \le r_0 - 1} \left(\sup_{n \ge 0} \frac{n}{Q_p(t'+n)} \right)$, which gives the first point of the theorem. The second point follows by the same arguments. □

4 Inference for semi-strong GARCH(p,q)

Consider the GARCH (p, q) model

$$\begin{aligned} \epsilon_t &= \sqrt{h_t}\eta_t, \\ h_t &= \omega_0 + \textstyle\sum_{i=1}^{q}\alpha_{0i}\epsilon_{t-i}^2 + \sum_{j=1}^{p}\beta_{0j}h_{t-j}, \quad \forall t \in \mathbb{Z} \end{aligned} \tag{4.1}$$

where $\omega_0 > 0, \alpha_{0i} \geqslant 0$ $(i = 1, \ldots, q)$ and $\beta_{0j} \geqslant 0$ $(j = 1, \ldots, p)$. When (η_t) is iid, Model (4.1) is a standard GARCH, for which the statistical inference has been thoroughly studied (see Berkes et al. (2003) and Francq and Zakoïan (2004)). Escanciano (2009) succeeded in estimating the GARCH model without the assumption that (η_t) is iid, but had to assume that $E|\epsilon_t|^s < \infty$ for some small $s > 0$. The aim of this section is to relax this extra moment assumption.

Let

$$A_t = \begin{pmatrix} \alpha_{01}\eta_t^2 & \cdots & \alpha_{0q}\eta_t^2 & \beta_{01}\eta_t^2 & \cdots & \beta_{0p}\eta_t^2 \\ & I_{q-1} & & & 0_{(q-1)\times p} & \\ \alpha_{01} & \cdots & \alpha_{0q} & \beta_{01} & \cdots & \beta_{0p} \\ & 0_{(p-1)\times q} & & & I_{p-1} & \end{pmatrix} \text{ and } b_t = \begin{pmatrix} \omega_0 \\ 0_{p+q-1} \end{pmatrix}$$

with standard notations.

Model (4.1) is a special case of (2.1) where we use the notations $\boldsymbol{X}_t = \left(\epsilon_t^2, \ldots, \epsilon_{t-q+1}^2, h_t^2, \ldots, h_{t-p+1}^2\right)$, $\theta_t = (A_t, b_t)$, $\boldsymbol{\Psi}(\theta, x) = Ax + b$, and $d(x, y) = \|x - y\|$ for any norm $\|\cdot\| \in \mathbb{R}^{p+q}$. Remark that $\boldsymbol{\Lambda}_t^{(r)} = \|A_t A_{t-1} \ldots A_{t-r+1}\|$.

In the sequel, we do not assume that (η_t) is iid but only stationary and ergodic. If $\mathbb{E}\ln^+ \eta_1^2 < \infty$, Theorem 2.1 applies with $c = 0_{p+q}$. Therefore there exists a unique non-anticipative strictly stationary solution (ϵ_t) to model (4.1) if

$$\begin{aligned} \gamma(\mathbf{A}) &:= \inf_{r \in \mathbb{N}^*} \frac{1}{r}\mathbb{E}\left(\ln \|A_0 A_{-1} \ldots A_{-r+1}\|\right) \\ &= \lim_{r\to\infty} \frac{1}{r} \ln \|A_0 A_{-1} \ldots A_{-r+1}\| < 0 \text{ a.s.} \end{aligned}$$

By Theorem 2.2, it follows that the strictly stationary solution of (4.1) satisfies

$$\limsup_{n\to\infty} \frac{1}{n} \ln \epsilon_{t+n}^2 \leq 0, \quad \limsup_{n\to\infty} \frac{1}{n} \ln \epsilon_{t-n}^2 \leq 0 \quad \text{a.s.} \tag{4.2}$$

for all $t \in \mathbb{Z}$.

4.1 QMLE estimator

Let $\{\epsilon_t\}_{t=1}^n$ be a sample of size n of the unique non-anticipative strictly stationary solution of model (4.1). The vector of the parameters

$$\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{p+q+1})^{\mathrm{T}} = (\omega, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)^{\mathrm{T}}$$

belongs to a parameter space $\boldsymbol{\Theta} \subset]0,+\infty[\times [0,\infty[^{p+q}$. The true value of the parameter is unknown and is denoted by $\boldsymbol{\theta}_0 = (\omega_0, \alpha_{01}, \ldots, \alpha_{0q}, \beta_{01}, \ldots, \beta_{0p})^{\mathrm{T}}$. Conditionally on initial values $\epsilon_0, \ldots, \epsilon_{1-q}$, $\tilde{\sigma}_0^2, \ldots, \tilde{\sigma}_{1-p}^2$, the Gaussian quasi-likelihood is defined by

$$L_n(\boldsymbol{\theta}) = L_n(\boldsymbol{\theta}; \epsilon_1, \ldots, \epsilon_n) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\tilde{\sigma}_t^2}} \exp\left(-\frac{\epsilon_t^2}{2\tilde{\sigma}_t^2}\right),$$

where the $\tilde{\sigma}_t^2$ are defined recursively, for $t \geqslant 1$, by

$$\tilde{\sigma}_t^2 = \tilde{\sigma}_t^2(\boldsymbol{\theta}) = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \tilde{\sigma}_{t-j}^2.$$

For instance, the initial values can be chosen as

$$\epsilon_0^2 = \ldots = \epsilon_{1-q}^2 = \tilde{\sigma}_0^2 = \ldots = \tilde{\sigma}_{1-p}^2 = c$$

with $c = \omega$ or ϵ_1^2. The standard estimator of the GARCH parameter $\boldsymbol{\theta}_0$ is the QMLE defined as any measurable solution $\hat{\boldsymbol{\theta}}_n$ of

$$\hat{\boldsymbol{\theta}}_n = \arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} L_n(\boldsymbol{\theta}) = \operatorname*{argmin}_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \tilde{\mathbf{I}}_n(\boldsymbol{\theta}) \tag{4.3}$$

where $\tilde{\mathbf{I}}_n(\boldsymbol{\theta}) = n^{-1}\sum_{t=1}^{n} \tilde{\ell}_t$ and $\tilde{\ell}_t = \tilde{\ell}_t(\boldsymbol{\theta}) = \frac{\epsilon_t^2}{\tilde{\sigma}_t^2} + \ln \tilde{\sigma}_t^2$.

Let $\mathcal{A}_{\boldsymbol{\theta}}(z) = \sum_{i=1}^{q} \alpha_i z^i$ and $\mathcal{B}_{\boldsymbol{\theta}}(z) = 1 - \sum_{j=1}^{p} \beta_j z^j$. It is not restrictive to assume that $q \geq 1$. By convention $\mathcal{B}_{\boldsymbol{\theta}}(z) = 1$ if $p = 0$. Let $\mathcal{F}_{t-1}$ be the σ-field generated by $(\epsilon_{t-1}, \epsilon_{t-2}, \ldots)$. To show the strong consistency, the following assumptions will be made.

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

A2 $\gamma(\mathbf{A}_0) < 0$ and $\forall \boldsymbol{\theta} \in \boldsymbol{\Theta}$, $\sum_{j=1}^{p} \beta_j < 1$.

A3 (η_t) is stationary and ergodic, η_t^2 has a non-degenerate distribution with *i)* $\mathbb{E}[\eta_t^2 \mid \mathcal{F}_{t-1}] = 1$ a.s. and *ii)* $E \ln \eta_t^2 > -\infty$.

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

Remark 4.1. Assumptions **A1**, **A2** and **A4** are standard (see Francq and Zakoïan (2004) for comments on these assumptions). Condition *i)* in **A3** is obviously less restrictive than the iid assumption with finite second-order moments (see Example 2.1 of Francq and Zakoïan (2020)). This assumption was first used by Lee and Hansen (1994) for inference of GARCH models.

Escanciano (2009) established the consistency of the QMLE under this assumption, with a small-order moment condition of the observed process instead of our condition *ii)* of assumption **A3**. Note that this later condition precludes densities with too much mass around zero, but is satisfied by most commonly used distributions. It is also weaker than the regularity condition on the η_t law ($\lim_{t\to 0} t^{-\mu} P\{\eta_0^2 \leqslant t\} = 0$, for some $\mu > 0$) used by Berkes et al. (2003)[1].

Assumption **A2** implies that the roots of $\mathcal{B}_{\boldsymbol{\theta}}(z)$ are outside the unit disc. Therefore, by the second inequality of (4.2), we can define $(\sigma_t^2) = \{\sigma_t^2(\boldsymbol{\theta})\}$ as the (unique) strictly stationary, ergodic and non-anticipative solution[2] of

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \quad \forall t. \tag{4.4}$$

Note that $\sigma_t^2(\boldsymbol{\theta}_0) = h_t$. Let

$$\mathbf{I}_n(\boldsymbol{\theta}) = \mathbf{I}_n(\boldsymbol{\theta}; \epsilon_n, \epsilon_{n-1}\ldots,) = n^{-1}\sum_{t=1}^{n} \ell_t, \quad \ell_t = \ell_t(\boldsymbol{\theta}) = \frac{\epsilon_t^2}{\sigma_t^2} + \ln \sigma_t^2.$$

[1] Knowing that $\mathbb{E}\left(\ln^+(\eta_1^2)\right) < \infty$ by *i)* **A3**, to establish *ii)* **A3** it is therefore sufficient to prove that $\mathbb{E}\left(\ln^-(\eta_1^2)\right) < \infty$. Using $\mathbb{E}\left(\ln^-(\eta_1^2)\right) = \int_0^\infty \mathbb{P}(\ln^+(\frac{1}{\eta_1^2}) \geq s)ds = \int_0^\infty \mathbb{P}(\ln(\frac{1}{\eta_1^2}) \geq s)ds = \int_0^\infty \mathbb{P}(\frac{1}{\eta_1^2} \geq \exp(s))ds = \int_0^\infty \mathbb{P}(\eta_1^2 \leq \exp(-s))ds$, we have under the condition of Berkes et al. (2003) that $\mathbb{P}(\eta_1^2 \leq \exp(-s)) = o(\exp(-\mu s))$ when $s \to \infty$, which gives the result.

[2] Rewrite (4.4) in vector form as

$$\underline{\sigma}_t^2 = \underline{c}_t + B\underline{\sigma}_{t-1}^2,$$

where

$$\underline{\sigma}_t^2 = \begin{pmatrix} \sigma_t^2 \\ \sigma_{t-1}^2 \\ \vdots \\ \sigma_{t-p+1}^2 \end{pmatrix}, \quad \underline{c}_t = \begin{pmatrix} \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_p \\ 1 & 0 & \cdots & 0 \\ \vdots & & & \\ 0 & \cdots & 1 & 0 \end{pmatrix},$$

we have by the second inequality of (4.2) that $\limsup_{n\to\infty} \frac{1}{n} \ln \|\underline{c}_n\| \leq 0$. By Assumption **A2**, we deduct that $\limsup_{n\to\infty} \frac{1}{n} \ln \|B^n \underline{c}_{n-1}^2\| \leq \limsup_{n\to\infty} \frac{1}{n} \ln \|B^n\| + \limsup_{n\to\infty} \frac{1}{n} \ln \|\underline{c}_n\| < 0$. From which we deduce by the Cauchy rule that the series $\hat{\sigma}_t^2 := \sum_{n=0}^{\infty} B^n \underline{c}_{t-n}^2$ converges almost surly. We note that $(\hat{\sigma}_t^2)$ is a strictly stationary, ergodic and non-anticipative solution of (4.4). For unicity, assume that there exists another stationary process $\left(\underline{\sigma}_t^2 *\right)$ of (4.4). For all $n \geq 0$, we have $\|\underline{\sigma}_t^2 * - \hat{\sigma}_t^2\| = \|B^n \underline{\sigma}_{t-n}^2 * - B^n \hat{\sigma}_{t-n}^2\| \leq \|B^n\| \|\underline{\sigma}_{t-n}^2 *\| + \|B^n\| \|\hat{\sigma}_{t-n}^2\|$. Since $\|B^n\| \to 0$ a.s. as $n \to \infty$ and $\|\underline{\sigma}_{t-n}^2 *\|$ and $\|\hat{\sigma}_{t-n}^2\|$ converges in law by stationary, Slutsky's theorem entails that $\|\underline{\sigma}_t^2 * - \hat{\sigma}_t^2\|$ converges in law to 0 a.s. as $n \to \infty$. Since $\|\underline{\sigma}_t^2 * - \hat{\sigma}_t^2\|$ does not depend on n, we conclude that $\|\underline{\sigma}_t^2 * - \hat{\sigma}_t^2\| = 0$ $a.s.$

We are now able to establish the strong consistency of the QMLE.

Theorem 4.1. *Let* $\left(\hat{\boldsymbol{\theta}}_n\right)$ *be a sequence of QMLE satisfying* (4.3)*, with any initial condition c, under* ***A1-A4****, almost surely* $\hat{\boldsymbol{\theta}}_n \to \boldsymbol{\theta}_0$*, as* $n \to \infty$.

Remark 4.2. *Escanciano (2009) establishes the asymptotic normality of the QMLE under the assumption that a small-order moment exists. This moment condition is mainly used to justify the existence of the asymptotic covariance of the QMLE. To the best of our knowledge, the asymptotic normality has never been shown without a hypothesis that implies the existence of a small-order moment. In some cases, the asymptotic covariance matrix may not exist without a finite moment of sufficiently high order (see Francq and Zakoian (2007, paragraph 3.1)). In our framework, the study of asymptotic distribution of the semi-strong GARCH without moment condition remains difficult and is left for future work.*

Proof. of Theorem 4.1.
The proof relies on the following intermediate results.

$i)$ $\lim_{n\to\infty} \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} |\mathbf{l}_n(\boldsymbol{\theta}) - \tilde{\mathbf{l}}_n(\boldsymbol{\theta})| = 0, \quad a.s.$

$ii)$ if $\sigma_t^2(\boldsymbol{\theta}) = \sigma_t^2(\boldsymbol{\theta}_0) \quad a.s.,$ then $\boldsymbol{\theta} = \boldsymbol{\theta}_0$,

$iii)$ if $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, then $\mathbb{E}\{\ell_1(\boldsymbol{\theta}) - \ell_1(\boldsymbol{\theta}_0)\} > 0$,

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

$$\liminf_{n\to\infty} \left(\inf_{\boldsymbol{\theta}^*\in V(\boldsymbol{\theta})\cap\boldsymbol{\Theta}} \tilde{\mathbf{l}}_n(\boldsymbol{\theta}^*) - \tilde{\mathbf{l}}_n(\boldsymbol{\theta}_0) \right) > 0 \quad a.s.$$

An analysis of the proof of *i)* and *ii)* in Francq and Zakoïan (2004) shows that the authors only use their small-order moment result on ϵ_n^2 [see Berkes et al. (2003, Lemma 2.3)] to prove that $\lim_{n\to\infty} \delta^n \epsilon_n^2 = 0 \; a.s, \; \forall \; \delta \in [0,1)$. Since the first inequality of (4.2) implies the latter result, the proofs of Points *i)* and *ii)* follow.

Now let us turn of to the proof of *iii)*. Let $W_t(\boldsymbol{\theta}) = \sigma_t^2(\boldsymbol{\theta}_0)/\sigma_t^2(\boldsymbol{\theta})$ and, for $K > 0$, $A_K = [K^{-1}, K]$, write

$$\ell_t(\boldsymbol{\theta}) - \ell_t(\boldsymbol{\theta}_0) = g(W_t(\boldsymbol{\theta}), \eta_t^2)\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K} + g(W_t(\boldsymbol{\theta}), \eta_t^2)\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K^c}$$

where, for $x > 0, y \geq 0$, $g(x, y) = -\log x + y(x-1)$. Introducing the negative part $x^- = \max(-x, 0)$ of any real number x, we thus have

$$\ell_t(\boldsymbol{\theta}) - \ell_t(\boldsymbol{\theta}_0) \geq g(W_t(\boldsymbol{\theta}), \eta_t^2)\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K} - \left\{g(W_t(\boldsymbol{\theta}), \eta_t^2)\right\}^- \mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K^c} \quad (4.5)$$

The expectation of the first term in the r.h.s. is well-defined and satisfies

$$E[g(W_t(\boldsymbol{\theta}), \eta_t^2)\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K}] = E[g(W_t(\boldsymbol{\theta}), 1)\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K}] \geq 0$$

since $g(x, 1) \geq 0$ for any $x \geq 0$, with equality only if $x = 1$. By *ii)* we have that $W_t(\boldsymbol{\theta}) = 1$ a.s. if and only if $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. We thus have, by Beppo-Levi's theorem,

$$\begin{aligned} \lim_{K\to\infty} E[g(W_t(\boldsymbol{\theta}), \eta_t^2)\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K}] &= E[g(W_t(\boldsymbol{\theta}), 1) \lim_{K\to\infty} \mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K}] \\ &= E[g(W_t(\boldsymbol{\theta}), 1)] > 0 \quad \text{for} \quad \boldsymbol{\theta} \neq \boldsymbol{\theta}_0. \end{aligned}$$

To deal with the expectation of the second term in the r.h.s. of (4.5) we use the fact that for $y > 0$, $g(x, y) \geq g(1/y, y)$. It follows that

$$-E\left[\left\{g(W_t(\boldsymbol{\theta}), \eta_t^2)\right\}^- \mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K^c}\right] \geq -E\left[\left\{g(1/\eta_t^2, \eta_t^2)\right\}^- \mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K^c}\right] \to 0 \quad \text{as } K \to \infty,$$

because, by *ii)* **A3**, $E\left[\{g(1/\eta_t^2, \eta_t^2)\}^-\right] < \infty$ and thus the convergence holds by Lebesgue's dominated convergence theorem. This completes the proof of Step *iii)*.

Now we prove *iv)*. For any $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ we have

$$\tilde{\mathbf{l}}_n(\boldsymbol{\theta}) - \tilde{\mathbf{l}}_n(\boldsymbol{\theta}_0) \geq \mathbf{l}_n(\boldsymbol{\theta}) - \mathbf{l}_n(\boldsymbol{\theta}_0) - |\tilde{\mathbf{l}}_n(\boldsymbol{\theta}) - \mathbf{l}_n(\boldsymbol{\theta})| - |\tilde{\mathbf{l}}_n(\boldsymbol{\theta}_0) - \mathbf{l}_n(\boldsymbol{\theta}_0)|.$$

Hence, using *i)*

$$\begin{aligned} &\liminf_{n\to\infty} \left(\inf_{\boldsymbol{\theta}^*\in V(\boldsymbol{\theta})\cap\boldsymbol{\Theta}} \tilde{\mathbf{l}}_n(\boldsymbol{\theta}^*) - \tilde{\mathbf{l}}_n(\boldsymbol{\theta}_0) \right) \\ &\geq \liminf_{n\to\infty} \left(\inf_{\boldsymbol{\theta}^*\in V(\boldsymbol{\theta})\cap\boldsymbol{\Theta}} \mathbf{l}_n(\boldsymbol{\theta}^*) - \mathbf{l}_n(\boldsymbol{\theta}_0) \right) - 2\limsup_{n\to\infty} \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} |\tilde{\mathbf{l}}_n(\boldsymbol{\theta}) - \mathbf{l}_n(\boldsymbol{\theta})| \\ &= \liminf_{n\to\infty} \left(\inf_{\boldsymbol{\theta}^*\in V(\boldsymbol{\theta})\cap\boldsymbol{\Theta}} \mathbf{l}_n(\boldsymbol{\theta}^*) - \mathbf{l}_n(\boldsymbol{\theta}_0) \right). \end{aligned} \tag{4.6}$$

For any $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and any positive integer k, let $V_k(\boldsymbol{\theta})$ the open ball of center $\boldsymbol{\theta}$ and radius $1/k$. We have

$$\liminf_{n\to\infty} \left(\inf_{\boldsymbol{\theta}^*\in V_k(\boldsymbol{\theta})\cap\boldsymbol{\Theta}} \mathbf{l}_n(\boldsymbol{\theta}^*) - \mathbf{l}_n(\boldsymbol{\theta}_0) \right) \geq \liminf_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \inf_{\boldsymbol{\theta}^*\in V_k(\boldsymbol{\theta})\cap\boldsymbol{\Theta}} \ell_t(\boldsymbol{\theta}^*) - \ell_t(\boldsymbol{\theta}_0). \tag{4.7}$$

By arguments already given, under *ii)* **A3**,

$$E\left(\inf_{\boldsymbol{\theta}^*\in V_k(\boldsymbol{\theta})\cap\boldsymbol{\Theta}} \ell_t(\boldsymbol{\theta}^*) - \ell_t(\boldsymbol{\theta}_0)\right)^- \leq E\left(g(1/\eta_t^2, \eta_t^2))\right)^- < \infty.$$

Therefore $E\left(\inf_{\boldsymbol{\theta}^*\in V_k(\boldsymbol{\theta})\cap\boldsymbol{\Theta}} \ell_t(\boldsymbol{\theta}^*) - \ell_t(\boldsymbol{\theta}_0)\right)$ exists in $\mathbb{R}\cup\{+\infty\}$, and the ergodic theorem applies (see **?**, Exercises 7.3 and 7.4). From (4.7) we obtain

$$\liminf_{n\to\infty}\left(\inf_{\boldsymbol{\theta}^*\in V_k(\boldsymbol{\theta})\cap\boldsymbol{\Theta}} \mathbf{l}_n(\boldsymbol{\theta}^*) - \mathbf{l}_n(\boldsymbol{\theta}_0)\right) \geq E\left(\inf_{\boldsymbol{\theta}^*\in V_k(\boldsymbol{\theta})\cap\boldsymbol{\Theta}} \ell_t(\boldsymbol{\theta}^*) - \ell_t(\boldsymbol{\theta}_0)\right).$$

The latter term into parentheses converges to $\ell_t(\boldsymbol{\theta}) - \ell_t(\boldsymbol{\theta}_0)$ as $k\to\infty$, and, by standard arguments using the positive and negative parts of $\inf_{\boldsymbol{\theta}^*\in V_k(\boldsymbol{\theta})\cap\boldsymbol{\Theta}} \ell_t(\boldsymbol{\theta}^*) - \ell_t(\boldsymbol{\theta}_0)$, we have that

$$\lim_{k\to\infty} E\left(\inf_{\boldsymbol{\theta}^*\in V_k(\boldsymbol{\theta})\cap\boldsymbol{\Theta}} \ell_t(\boldsymbol{\theta}^*) - \ell_t(\boldsymbol{\theta}_0)\right) = E\left\{\ell_t(\boldsymbol{\theta}) - \ell_t(\boldsymbol{\theta}_0)\right\},$$

which by *i)* is strictly positive. In view of (4.6), the proof of *iv)* is complete.

Now we complete the proof of the theorem. The set $\boldsymbol{\Theta}$ is covered by the union of an arbitrary neighborhood $V(\boldsymbol{\theta}_0)$ of $\boldsymbol{\theta}_0$ and, for any $\boldsymbol{\theta}\neq\boldsymbol{\theta}_0$, by neighborhoods $V(\boldsymbol{\theta})$ satisfying *iv)*. Obviously, $\inf_{\boldsymbol{\theta}^*\in V(\boldsymbol{\theta}_0)\cap\boldsymbol{\Theta}} \tilde{\mathbf{l}}_n(\boldsymbol{\theta}^*) \leq \tilde{\mathbf{l}}_n(\boldsymbol{\theta}_0)$, *a.s.* Moreover, by compactness of $\boldsymbol{\Theta}$, there exists a finite subcover of the form $V(\boldsymbol{\theta}_0), V(\boldsymbol{\theta}_1), \ldots, V(\boldsymbol{\theta}_M)$. By *iv)*, for $i = 1,\ldots,M$, there exists n_i such that for $n \geq n_i$,

$$\inf_{\boldsymbol{\theta}^*\in V(\boldsymbol{\theta}_i)\cap\boldsymbol{\Theta}} \tilde{\mathbf{l}}_n(\boldsymbol{\theta}^*) > \tilde{\mathbf{l}}_n(\boldsymbol{\theta}_0), \quad a.s.$$

Thus for $n \geq \max_{i=1,\ldots,M}(n_i)$,

$$\inf_{\boldsymbol{\theta}^*\in \bigcup_{i=1,\ldots,M} V(\boldsymbol{\theta}_i)\cap\boldsymbol{\Theta}} \tilde{\mathbf{l}}_n(\boldsymbol{\theta}^*) > \tilde{\mathbf{l}}_n(\boldsymbol{\theta}_0), \quad a.s.$$

from which we deduce that $\widehat{\boldsymbol{\theta}}_n$ belongs to $V(\boldsymbol{\theta}_0)$ for sufficiently large n. □

Acknowledgements

I am most thankful to Christian Francq and Jean-Michel Zakoïan for their guidance and feedback.

References

I. Berkes, L. Horváth, and P. Kokoszka. GARCH processes: structure and estimation. *Bernoulli*, 9(2):201 – 227, 2003.

P. Bougerol. Kalman filtering with random coefficients and contractions. *Siam Journal on Control and Optimization*, 31:942–959, 1993.

P. Bougerol and N. Picard. Stationarity of garch processes and of some nonnegative time series. *Journal of Econometrics*, 52(1):115–127, 1992a.

P. Bougerol and N. Picard. Strict Stationarity of Generalized Autoregressive Processes. *The Annals of Probability*, 20(4):1714 – 1730, 1992b.

A. Brandt. The stochastic equation yn 1=anyn bn with stationary coefficients. *Advances in Applied Probability*, 18:211–220, 1986.

D. Buraczewski, E. Damek, T. Mikosch, et al. Stochastic models with power-law tails. Springer, 2016.

J. C. Escanciano. Quasi-maximum likelihood estimation of semi-strong garch models. *Econometric Theory*, 25(2):561–570, 2009.

C. Francq and L. Q. Thieu. Qml inference for volatility models with covariates. *Econometric Theory*, 35(1):37–72, 2019.

C. Francq and J.-M. Zakoian. Quasi-maximum likelihood estimation in garch processes when some coefficients are equal to zero. *Stochastic Processes and their Applications*, 117(9):1265–1284, 2007.

C. Francq and J.-M. Zakoïan. Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10(4):605 – 637, 2004.

C. Francq and J.-M. Zakoïan. Virtual historical simulation for estimating the conditional var of large portfolios. *Journal of Econometrics*, 217(2): 356–380, 2020.

C. Francq, O. Wintenberger, and J.-M. Zakoïan. Goodness-of-fit tests for log-garch and egarch models. *Test*, 27(1):27–51, 2018.

H. Han and D. Kristensen. Asymptotic theory for the qmle in garch-x models with stationary and nonstationary covariates. *Journal of Business & Economic Statistics*, 32(3):416–429, 2014.

H. Kesten. Random difference equations and renewal theory for products of random matrices. *Acta Mathematica*, 131:207–248, 1973.

J. F. C. Kingman. Subadditive Ergodic Theory. *The Annals of Probability*, 1(6):883 – 899, 1973.

D. Straumann and T. Mikosch. Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. *The Annals of Statistics*, 34(5):2449 – 2495, 2006.

A Complementary proofs

Proof of Theorem 2.1. For all $t \in \mathbb{Z}$ and $n \in \mathbb{N}$, let

$$\boldsymbol{X}_{t,n} = \boldsymbol{\Psi}\left(\theta_t, \boldsymbol{X}_{t-1,n-1}\right) \tag{A.1}$$

with $\boldsymbol{X}_{t,0} = c$. Note that

$$\boldsymbol{X}_{t,n} = \psi_n\left(\theta_t, \theta_{t-1}, \ldots, \theta_{t-n+1}\right)$$

for some measurable function $\psi_n : (E^n, \mathcal{B}_{E^n}) \to (F, \mathcal{B}_F)$, with the usual notation. For all n, the sequence $\left(\boldsymbol{X}_{t,n}\right)_{t\in\mathbb{Z}}$ is thus stationary and ergodic. If for all t, the limit $\boldsymbol{X}_t = \lim_{n\to\infty} \boldsymbol{X}_{t,n}$ exists a.s., then by taking the limit of both sides of Equation (A.1), it can be seen that the process $(\boldsymbol{X}_t)$ is solution of Equation (2.1). When it exists, the limit is a measurable function of the form $\boldsymbol{X}_t = \psi\left(\theta_t, \theta_{t-1}, \ldots\right)$[3] and is therefore stationary and ergodic. To prove the existence of $\lim_{n\to\infty} \boldsymbol{X}_{t,n}$, we will show that, a.s., $\left(\boldsymbol{X}_{t,n}\right)_{n\in\mathbb{N}}$ is a Cauchy sequence in the complete space F. By iterating equation (A.1) we have

$$\boldsymbol{X}_{t,n} = \boldsymbol{\Psi}_t \circ \cdots \circ \boldsymbol{\Psi}_{t-n+1}\left(c\right).$$

If follows that

$$d\left(\boldsymbol{X}_{t,n}, \boldsymbol{X}_{t,n-1}\right) \leq \boldsymbol{\Lambda}_t^{(n-1)} d\left(\boldsymbol{\Psi}_{t-n+1}\left(c\right), c\right).$$

For $n < m$, we thus have

$$\begin{aligned} d\left(\boldsymbol{X}_{t,m}, \boldsymbol{X}_{t,n}\right) &\leq \sum_{k=0}^{m-n-1} d\left(\boldsymbol{X}_{t,m-k}, \boldsymbol{X}_{t,m-k-1}\right) \\ &\leq \sum_{k=0}^{m-n-1} \boldsymbol{\Lambda}_t^{(m-k-1)} d\left(\boldsymbol{\Psi}_{t-m+k+1}\left(c\right), c\right) \qquad \text{(A.2)} \\ &\leq \sum_{j=n}^{\infty} \boldsymbol{\Lambda}_t^{(j)} d\left(\boldsymbol{\Psi}_{t-j}\left(c\right), c\right). \end{aligned}$$

[3]For the measurability of $\boldsymbol{X}_t$, one can consider $\boldsymbol{X}_{t,n}$ as functions of $(\theta_t, \theta_{t-1}, \cdots)$ and argue that in metric space, a limit of measurable functions is measurable.

Note that

$$\limsup_{j\to\infty} \ln\left(\boldsymbol{\Lambda}_t^{(j)} d\left(\boldsymbol{\Psi}_{t-j}(c), c\right)\right)^{1/j} = \limsup_{j\to\infty} \frac{1}{j}\left(\ln \boldsymbol{\Lambda}_t^{(j)} + \ln d\left(\boldsymbol{\Psi}_{t-j}(c), c\right)\right) < 0$$

under (i) and (ii), by using Kingman's sub-additive ergodic theorem (see Kingman (1973)) and **?**, Exercise 4.12. We conclude, from the Cauchy criterion for the convergence of series with positive terms, that

$$\sum_{j=1}^{\infty} \boldsymbol{\Lambda}_t^{(j)} d\left(\boldsymbol{\Psi}_{t-j}(c), c\right)$$

is a.s. finite, under (i) and (ii). It follows that $(\boldsymbol{X}_{t,n})_{n\in\mathbb{N}}$ is a.s. a Cauchy sequence in F. The existence of a stationary and ergodic solution to Equation (2.1) follows.

Assume that there exists another stationary process $(\boldsymbol{X}_t^*)$ such that $\boldsymbol{X}_t^* = \boldsymbol{\Psi}_t\left(\boldsymbol{X}_{t-1}^*\right)$. For all $N \geq 0$, we have

$$d\left(\boldsymbol{X}_t, \boldsymbol{X}_t^*\right) \leq \boldsymbol{\Lambda}_t^{(N+1)} d\left(\boldsymbol{X}_{t-N}, \boldsymbol{X}_{t-N}^*\right). \tag{A.3}$$

Since $\boldsymbol{\Lambda}_t^{(N+1)} \to 0$ a.s. as $N \to \infty$, and $d\left(\boldsymbol{X}_{t-N}, \boldsymbol{X}_{t-N}^*\right) = O_P(1)$ by stationarity, the right-hand side of Equation (A.3) tends to zero in probability. Since the left-hand side does not depend on N, we have $P\left(|\boldsymbol{X}_t - \boldsymbol{X}_t^*| > \epsilon\right) = 0$ for all $\epsilon > 0$, and thus $P\left(\boldsymbol{X}_t = \boldsymbol{X}_t^*\right) = 1$, which establishes the uniqueness. In view of Equation (A.2), we have

$$d\left(\boldsymbol{X}_t, c\right) \leq \sum_{j=0}^{\infty} \boldsymbol{\Lambda}_t^{(j)} d\left(\boldsymbol{\Psi}_{t-j}(c), c\right)$$

and Equation (2.2) follows. □

CREST
CENTER FOR RESEARCH
IN ECONOMICS AND STATISTICS

CREST
Center for Research in Economics and Statistics
UMR 9194

5 Avenue Henry Le Chatelier
TSA 96642
91764 Palaiseau Cedex
FRANCE

Phone: +33 (0)1 70 26 67 00
Email: info@crest.science
https://crest.science/