COMMENTS ON THE PRESENCE OF SERIAL CORRELATION IN THE RANDOM COEFFICIENTS OF AN AUTOREGRESSIVE PROCESS

FRÉDÉRIC PROÏA AND MARIUS SOLTANE

ABSTRACT. We consider an RCAR(p) process and we establish that the standard estimation lacks consistency as soon as there exists a nonzero serial correlation in the coefficients. We give the correct asymptotic behavior and some simulations come to illustrate the results.

1. INTRODUCTION AND MOTIVATIONS

This note is devoted to the estimation issue in the random coefficients non-linear generalization of the autoregressive process or order p, that is usually defined as

$$X_t = \sum_{k=1}^{p} (\theta_k + b_{k,t}) \, X_{t-k} + \varepsilon_t$$

where, for each $1 \leqslant k \leqslant p$, $(b_{k,t})_{t \in \mathbb{Z}}$ is a white noise included in the coefficients, independent of $(\varepsilon_t)_{t \in \mathbb{Z}}$. We refer the reader to the monograph of [Nicholls and Quinn, 1982] for a rich introduction to the topic. The conditions of existence of a stationary solution to this equation have been widely studied since the seminal works of [Anděl, 1976] or [Nicholls and Quinn, 1981b], see also the more recent paper of [Aue et al., 2006]. The estimation of θ has been extensively developed to this day as well. [Nicholls and Quinn, 1981a] suggest to make use of the OLS which turns out to be the same, whether there is randomness or not in the coefficients. Under the stationarity conditions, the OLS is known to be strongly consistent and asymptotically normal in both cases (but with different variances). Looking for a unified theory, that is, irrespective of stationarity issues, [Aue and Horváth, 2011] and [Berkes et al., 2009] show that the QMLE is consistent and asymptotically normal as soon as there is some randomness in the coefficients, the variability of which allows to circumvent the well-known unit root issues. Later, [Hill and Peng, 2014] develop an empirical likelihood estimation which is asymptotically normal even when the coefficients are non-random. Let us also mention the WLS approach of [Schick, 1996] and the M-estimation of [Koul and Schick, 1996] adapted to stationary processes, which provide asymptotically normal estimators with optimal variance. Yet, the OLS estimation, easy to compute, that does not require numerical optimization or additional parametrization, is still very popular. We shall note that all these works, except those of Nicholls and Quinn, are related to the first-order process. By contrast, studies on general multivariate RCAR processes do not seem widespread in the literature, to the best of our knowledge. In this paper, we are interested in the implications of serial correlation in the random coefficients of a p-order RCAR process. Among the fundamental hypotheses of the usual theory, the coefficients are considered as white noises. However, our main statement is that, in a chronological model, it is

unlikely that the coefficients are serially uncorrelated, from the moment they are considered as random. From that point of view, we intend to show that the OLS method may lead to inappropriate conclusions when there is some serial correlation in the random coefficients. This work can be seen as a partial generalization of [Proïa and Soltane, 2018] where in particular, for the first-order RCAR process, the lack of consistency is established together with the correct behavior of the statistic. For this purpose, consider the RCAR process $(X_t)_{t \in \mathbb{Z}}$ of order p generated by the autoregression

$$(1.1) \qquad X_t = \sum_{k=1}^{p} (\theta_k + \alpha_k \, \eta_{k,\,t-1} + \eta_{k,\,t}) \, X_{t-k} + \varepsilon_t$$

where $(\varepsilon_t)_{t \in \mathbb{Z}}$ and $(\eta_{k,\,t})_{t \in \mathbb{Z}}$ are uncorrelated strong white noises (that is, i.i.d. sequences) with qth-order moments σ_q and $\tau_{k,\,q}$ respectively. We also assume that $(\eta_{k,\,t})_{t \in \mathbb{Z}}$ and $(\eta_{\ell,\,t})_{t \in \mathbb{Z}}$ are uncorrelated, for $k \neq \ell$. The main fact to give credit to this model is Prop. 3.2.1 of [Brockwell and Davis, 1991] which states that any stationary process with finite memory can be expressed by a moving average structure. Thus, by means of an MA(1) model, we introduce a lag-one memory in the random coefficients as a simple case study meant to illustrate the effect on the estimation. Note that, although much more cumbersome, the same study may lead to the same conclusions for any stationary coefficients with finite memory, *via* general MA structures. However, for infinite memory (*e.g.* stationary random AR(1) coefficients), the problem is probably much more complicated. In Section 2, we introduce our hypotheses and we detail the asymptotic behavior of the standard OLS in presence of serial correlation, under the stationarity conditions. In particular, we brought to light a disturbing consequence when testing the significance of θ. Some simulations come to illustrate our results in Section 3 while Section 4 contains the proofs. Finally, we postpone to the Appendix the purely computational steps, for the sake or readability.

2. Influence of serial correlation in the coefficients

It will be convenient to write (1.1) in the vector form

$$(2.1) \qquad \Phi_t = (C_\theta + N_{t-1} \, D_\alpha + N_t) \, \Phi_{t-1} + E_t$$

where $\Phi_t^T = (X_t, \ldots, X_{t-p+1})$, $E_t^T = (\varepsilon_t, 0, \ldots, 0)$, N_t is a matrix with $\eta_t = (\eta_{1,\,t}, \ldots, \eta_{p,\,t})$ in the first row and 0 elsewhere, $D_\alpha = \text{diag}(\alpha_1, \ldots, \alpha_p)$ and C_θ is the companion matrix of the underlying AR(p) process. We already make the assumption that any odd moment of ε_0 and $\eta_{k,0}$ is zero as soon as it exists ($\sigma_{2q+1} = \tau_{k,\,2q+1} = 0$), and that the distribution of the noises guarantees

$$(2.2) \qquad \mathbb{E}[\ln \| C_\theta + N_0 \, D_\alpha + N_1 \|] < 0 \quad \text{and} \quad \mathbb{E}[\ln^+ |\varepsilon_0|] < +\infty.$$

Those moments conditions are assumed to hold throughout the study. As for the moments of the process, the hypotheses are related to the space Θ where the parameters live. Define

$$\Theta_m = \{ \{\theta, \alpha, \sigma_q, \tau_{k,\,q'}\} \mid \mathbb{E}[\eta_{k,\,t}^a X_t^m] < +\infty \text{ for } a \in \{0, \ldots, m\} \text{ and } k \in \{1, \ldots, p\}\}$$

for $m = 2, 4$, where σ_q and $\tau_{k,\,q'}$ name in a generic way the highest-order moments of (ε_t) and $(\eta_{k,\,t})$ necessary to obtain $\mathbb{E}[\eta_{k,\,t}^a X_t^m] < +\infty$. We give in the Appendix some precise facts about Θ_2 and Θ_4, in particular we can see that $q = m$ and $q' = 2m$ are needed. The pathological cases are put together in a set called Θ^* that is built step by step during the

reasonings. First, we have the following causal representation showing in particular that the process is adapted to the filtration

(2.3) $$\mathcal{F}_t = \sigma((\varepsilon_s, \eta_{1,s}, \ldots, \eta_{p,s}), \ s \leqslant t).$$

The reader will find the proofs of our results in Section 4.

Proposition 2.1. *For all $t \in \mathbb{Z}$,*

(2.4) $$\Phi_t = E_t + \sum_{k=1}^{\infty} E_{t-k} \prod_{\ell=0}^{k-1} (C_\theta + N_{t-\ell-1} D_\alpha + N_{t-\ell}) \quad \text{a.s.}$$

In consequence, (X_t) is strictly stationary and ergodic.

Suppose now that $(X_{-p+1}, \ldots, X_n)$ is an available trajectory from which we deduce the standard OLS estimator of θ, the mean value of the coefficients,

(2.5) $$\widehat{\theta}_n = S_{n-1}^{-1} \sum_{t=1}^{n} \Phi_{t-1} X_t \quad \text{where} \quad S_n = \sum_{t=0}^{n} \Phi_t \Phi_t^T.$$

To simplify, the initial vector Φ_0 is assumed to follow the stationary distribution of the process. It is important to note that (2.5) is the OLS of θ with respect to $\sum_t (X_t - \theta^T \Phi_{t-1})^2$ even when $\alpha \neq 0$, as it is done in Sec. 3 of [Nicholls and Quinn, 1981a], because our interest is precisely to show that the usual estimation (*i.e.* assuming $\alpha = 0$) may lead to inappropriate conclusions in case of misspecification of the coefficients.

Remark 2.1. From the causal representation, ergodicity implies that

(2.6) $$\frac{1}{n} \sum_{t=1}^{n} Y_t \xrightarrow{\text{a.s.}} \mathbb{E}[Y_0] < +\infty \quad \text{for any process of the form} \quad Y_t = \prod_{i=1}^{m} X_{t-d_i} \eta_{k_i, t-d_i'}^{a_i}$$

where $0 \leqslant d_i' \leqslant d_i$, $a_i \in \{0, 1\}$ and $1 \leqslant k_i \leqslant p$, provided that $\Theta \subset \Theta_m$ ($m = 2, 4$). In all the study, any term taking the form of Y_t is called a *second-order* (or *fourth-order*) *isolated term* when $m = 2$ (or $m = 4$). According to (2.6), isolated terms satisfy $Y_t = o(n)$ and thus $|Y_t|^{1/2} = o(\sqrt{n})$ a.s. Such arguments are frequently used in the proofs.

The autocovariances of the stationary process are denoted by $\ell_i = \mathbb{E}[X_i X_0]$ and, by ergodicity, the sample covariances are strongly consistent estimators, that is

(2.7) $$\frac{1}{n} \sum_{t=1}^{n} X_t X_{t-i} \xrightarrow{\text{a.s.}} \ell_i.$$

Based on the autocovariances, we build

(2.8) $$\Lambda_0 = \begin{pmatrix} \ell_0 & \ell_1 & \cdots & \ell_{p-1} \\ \ell_1 & \ell_0 & \cdots & \ell_{p-2} \\ \vdots & \vdots & & \vdots \\ \ell_{p-1} & \ell_{p-2} & \cdots & \ell_0 \end{pmatrix} \quad \text{and} \quad L_1 = \begin{pmatrix} \ell_1 \\ \ell_2 \\ \vdots \\ \ell_p \end{pmatrix}.$$

The matrix Λ_0 is clearly positive semi-definite but the case where it would be non-invertible is part of Θ^*. This is the multivariate extension of $2\alpha\tau_2 \neq 1$ in [Proïa and Soltane, 2018], this can also be compared to assumption (v) in [Nicholls and Quinn, 1981a]. The asymptotic behavior of (2.5) is now going to be studied in terms of convergence, normality and rate. The result below is immediate from the ergodic theorem, provided that $\Theta \subset \Theta_2$.

3

Theorem 2.1. *Assume that $\Theta \subset \Theta_2 \backslash \Theta^*$. Then, we have the almost sure convergence*

$$\widehat{\theta}_n \xrightarrow{\text{a.s.}} \theta^* = \Lambda_0^{-1} L_1.$$

We give in Remark B.1 of the Appendix another expression of θ^* which directly shows that $\theta^* = \theta$ as soon as $\alpha = 0$, as it is established in Thm. 4.1 of [Nicholls and Quinn, 1981a]. However, except for $p = 1$ (see [Proïa and Soltane, 2018]), we generally do not have $\theta^* = 0$ when $\theta = 0$, and we will use this disturbing fact in the short example of Section 3.

Theorem 2.2. *Assume that $\Theta \subset \Theta_4 \backslash \Theta^*$. Then, there exists a limit matrix L such that we have the asymptotic normality*

$$\sqrt{n} \, (\widehat{\theta}_n - \theta^*) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Lambda_0^{-1} L \Lambda_0^{-1}).$$

Theorem 2.3. *Assume that $\Theta \subset \Theta_4 \backslash \Theta^*$. Then, we have the rate of convergence*

$$\limsup_{n \to +\infty} \frac{n}{2 \ln \ln n} \, \|\widehat{\theta}_n - \theta^*\|^2 < +\infty \quad \text{a.s.}$$

In particular, we can see that despite the correlation in the noise of the coefficients, the hypotheses are sufficient to ensure that the estimator achieves the usual rate of convergence in stable autoregressions, *i.e.*

$$\|\widehat{\theta}_n - \theta^*\| = O\left(\sqrt{\frac{\ln \ln n}{n}}\right) \quad \text{a.s.}$$

Note that σ_2 and $\tau_{k,4}$ are involved in Θ_2 to obtain the almost sure convergence, whereas σ_4 and $\tau_{k,8}$ are involved in Θ_4 to get the asymptotic normality and the rate of convergence.

Remark 2.2. Assuming that $(\eta_{k,t})$ and $(\eta_{\ell,t})$ are uncorrelated for $k \neq \ell$ is a matter of simplification of the calculations. We might as well consider a covariance structure for $(\eta_{1,t}, \ldots, \eta_{p,t})$ and obtain the same kind of results, at the cost of additional parameters and refined hypotheses (see *e.g.* assumption (iii) in [Nicholls and Quinn, 1981a]).

3. Illustration and Perspectives

To conclude this short note, let us illustrate a consequence of the presence of correlation in the coefficients when testing the significance of θ. For the sake of simplicity, take $p = 2$ and consider the test of $\mathcal{H}_0 : ``\theta_2 = 0"$ against $\mathcal{H}_1 : ``\theta_2 \neq 0"$. Thanks to Remark B.1, it is possible to show that, under $\mathcal{H}_0$,

$$\theta_1^* = \frac{\theta_1 \, (\theta_1^2 + \beta_2 - (1 - \beta_1)^2 - \beta_1 \, (\beta_1 + \beta_2 - 1))}{(\theta_1 - 2\beta_1 - \beta_2 + 1) \, (\theta_1 + 2\beta_1 + \beta_2 - 1)}$$

and

$$\theta_2^* = \frac{\theta_1^2 \, \beta_2 - \beta_1 \, ((1 - \beta_2)^2 + 4\beta_1 \, (\beta_1 + \beta_2 - 1))}{(\theta_1 - 2\beta_1 - \beta_2 + 1) \, (\theta_1 + 2\beta_1 + \beta_2 - 1)}$$

where $\beta_1 = \alpha_1 \tau_{1,2}$ and $\beta_2 = \alpha_2 \tau_{2,2}$. As a consequence, $\sqrt{n} \, |\widehat{\theta}_{2,n}|$ is almost surely divergent and even if $\theta_2 = 0$, we may detect a non-zero-mean coefficient where there is in fact a zero-mean autocorrelated one. Worse, suppose also that $\alpha_2 = \tau_{2,2} = 0$ so that there is no direct influence of X_{t-2} on X_t. Then, we still generally have $\theta_2^* \neq 0$. In other words, a correlation in the random coefficient associated with X_{t-1}, *i.e.* $\alpha_1 \neq 0$, generates a spurious detection of a direct influence of X_{t-2} on X_t. This phenomenon can be observed on some

simple but representative examples. To test $\mathcal{H}_0$, we infer two statistics from Thm. 4.1 of [Nicholls and Quinn, 1981a] and the procedure of estimation given by the authors,

$$Z_{1,n} = \frac{\sqrt{n}\,\widehat{\theta}_{2,n}}{v_{1,n}} \quad \text{and} \quad Z_{2,n} = \frac{\sqrt{n}\,\widehat{\theta}_{2,n}}{v_{2,n}}.$$

The first one takes into account random coefficients (which means that $\tau_{1,2}$ and $\tau_{2,2}$ are estimated to get $v_{1,n}$) and the second one is built assuming fixed coefficients ($\tau_{1,2}$ and $\tau_{2,2}$ are not estimated but set to 0). Note that we make sure that $\Theta \subset \Theta_4$ when we vary the settings. Unsurprisingly, $Z_{2,n}$ is less reliable since it does not model the random effect owing to $\tau_{1,2} > 0$ and the corrected statistic $Z_{1,n}$ behaves as expected when $\alpha_1 = 0$.
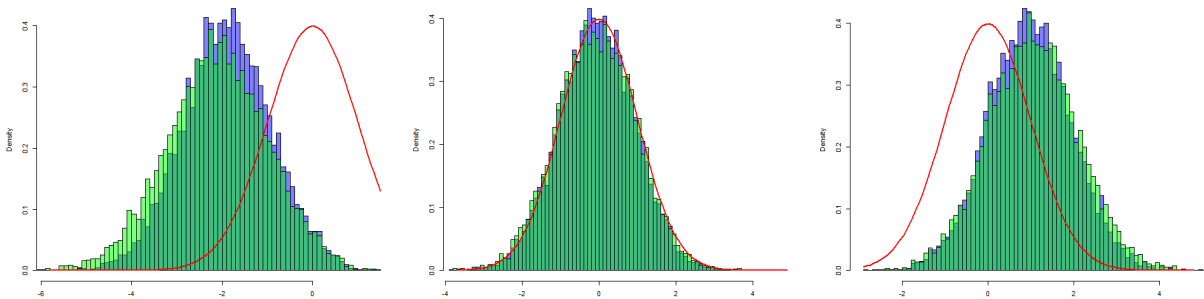

FIGURE 1. Empirical distribution of $Z_{1,n}$ (blue) and $Z_{2,n}$ (green) for Gaussian noises $\eta_{1,0} \sim \mathcal{N}(0, 0.2)$ and $\varepsilon_0 \sim \mathcal{N}(0, 1)$, $\theta_1 = 0.3$, $\theta_2 = \alpha_2 = \tau_{2,2} = 0$, and $\alpha_1 \in \{-0.5, 0, 0.3\}$ from left to right, obtained with $N = 10000$ repetitions of size $n = 500$. The red curve is the theoretical $\mathcal{N}(0, 1)$.

However, as it is visible on Figure 1 and confirmed by Figure 2, when $\alpha_1 \neq 0$ both tests reject $\mathcal{H}_0$ with a rate growing well above the 5% threshold that we used in this experiment even if, for the same reason as before, $Z_{1,n}$ appears slightly more robust. This is indeed what theory predicts. Theorem 2.2 also confirms that, once correctly recentered, $Z_{1,n}$ and $Z_{2,n}$ must remain asymptotically normal.
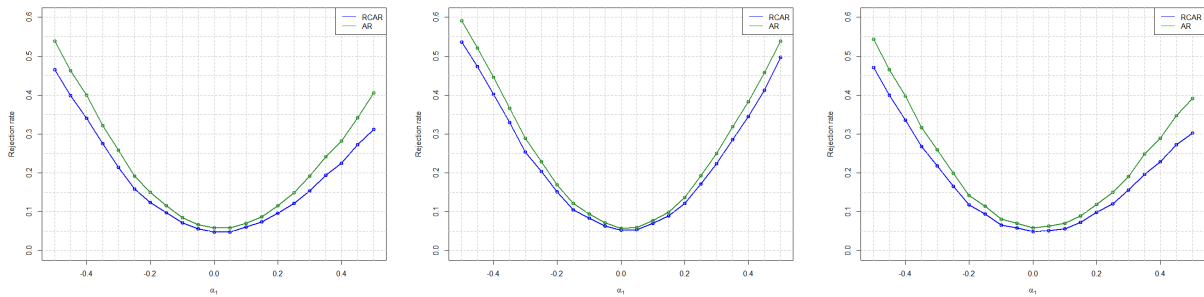

FIGURE 2. Rate of rejection of $\mathcal{H}_0$ for numerous values of $\alpha_1 \in [-0.5, 0.5]$ at the 5% threshold. The simulations are conducted with the setting of Figure 1 and, from left to right, with $\theta_1 = -0.3$, $\theta_1 = 0$ and $\theta_1 = 0.3$.

This example clearly highlights the question of whether we can build a consistent estimate for θ of course, but also for the covariance in the coefficients. In [Proïa and Soltane, 2018], this is done (see Sec. 4), mainly due to the fact that calculations are feasible when $p = 1$,

and a test for serial correlation followed. We can see through this note that this is not as easy in the general case and that this must be a reflection for future studies. Besides, we only considered the OLS but it would be worth working either with a QMLE or with a two-stage procedure to be able to exhibit a reliable estimate despite a possible serial correlation. We conclude by emphasizing the importance of developing statistical testing procedures for correlation in the random coefficients in the future works dedicated to RCAR processes. For practitioners, testing for serial correlation should appear as logic consequence of the tests for randomness in the coefficients, especially when the main hypothesis of the model is the existence of temporal correlations in the phenomenon.

4. TECHNICAL PROOFS

4.1. **Proof of Proposition 2.1.** From our hypotheses on the noises $(\eta_{k,t})$, the matrix-valued process $(C_\theta + N_{t-1} D_\alpha + N_t)$ is strictly stationary and ergodic. Thus, it follows from (2.2) that we can find $\delta < 0$ and a random k_0 such that, as soon as $k \geqslant k_0$,

$$\frac{1}{k} \sum_{\ell=0}^{k-1} \ln \|C_\theta + N_{t-\ell-1} D_\alpha + N_{t-\ell}\| < \delta \quad \text{a.s.}$$

See also Lem. 1.1 of [Brandt, 1986] for a similar reasoning. For $n \geqslant 1$, consider the truncation

$$\Phi_{t,n} = E_t + \sum_{k=1}^{n} E_{t-k} \prod_{\ell=0}^{k-1} (C_\theta + N_{t-\ell-1} D_\alpha + N_{t-\ell}).$$

Then, by the triangle inequality, for n large enough we have

$$\|\Phi_{t,n}\| \leqslant |\varepsilon_t| + \sum_{k=1}^{k_0-1} |\varepsilon_{t-k}| \prod_{\ell=0}^{k-1} \|C_\theta + N_{t-\ell-1} D_\alpha + N_{t-\ell}\| + \sum_{k=k_0}^{n} |\varepsilon_{t-k}| \, e^{\delta k}$$

and Lem. 2.2 of [Berkes et al., 2003] ensures the a.s. convergence of the last term under (2.2). Thus, the limit superior of $\|\Phi_{t,n}\|$ is finite and so is (2.4) with probability 1. Moreover, it is easy to check that this is a solution to the recurrence (2.1). Finally, the strict stationarity and ergodicity of (Φ_t) may be obtained following the same reasoning as in [Nicholls and Quinn, 1981a]. Indeed, one can see that there exists ϕ independent of t such that $\Phi_t = \phi((E_t, N_t), (E_{t-1}, N_{t-1}), \ldots)$ and the strict stationarity and ergodicity of the process (E_t, N_t) are passed to (Φ_t). That also implies the $\mathcal{F}_t$-measurability of (X_t). $\qquad\square$

4.2. **Proof of Theorem 2.2.** Let the filtration generated by $\mathcal{F}_0^* = \sigma(\Phi_0, \eta_{1,0}, \ldots, \eta_{p,0})$ and, for $n \geqslant 1$, by

(4.1) $\qquad \mathcal{F}_n^* = \sigma(\Phi_0, \eta_{1,0}, \ldots, \eta_{p,0}, (\varepsilon_1, \eta_{1,1}, \ldots, \eta_{p,1}), \ldots, (\varepsilon_n, \eta_{1,n}, \ldots, \eta_{p,n})).$

In the sequel, to avoid a huge amount of different notation, M_n and R_n will be generic terms, not necessarily identical from one line to another, designating vector $\mathcal{F}_n^*$-martingales (see *e.g.* [Duflo, 1997]) and isolated terms, respectively. We make use of the Appendix for some computational results and we start by two fundamental propositions showing that the

recentered empirical covariances are $\mathcal{F}_n^*$-martingales, up to residual terms. To this aim, we need to build the matrix $M_{\alpha,\beta}$ as follows. First, consider

$$
M_\theta = \begin{pmatrix} \theta_1 & 0 & \cdots & \cdots & 0 \\ \theta_2 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \theta_p & \cdots & \cdots & \theta_2 & \theta_1 \end{pmatrix} + \begin{pmatrix} 0 & \theta_2 & \cdots & \cdots & \theta_p \\ \vdots & \vdots & & \cdot^{\cdot} & 0 \\ \vdots & \vdots & \cdot^{\cdot} & & \vdots \\ \vdots & \theta_p & & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}
$$

and use it to define

$$
M_{\alpha,\beta} = \begin{pmatrix} \frac{1}{1-2\beta_1} & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \left[M_\theta + \begin{pmatrix} 0 & \beta_2 & 0 & \cdots & 0 \\ \beta_1 & 0 & 0 & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \beta_1 & 0 \end{pmatrix} \right]
$$

where $\beta_1 = \alpha_1 \tau_{1,2}$ and $\beta_2 = \alpha_2 \tau_{2,2} + \ldots + \alpha_p \tau_{p,2}$. Then, set U_0 as the first column of $M_{\alpha,\beta}$ and set K as the remaining part of $M_{\alpha,\beta}$ to which we add a zero vector on the right. Using the notation of the reasonings below, the pathological set Θ^* is enhanced with the situations where $2\beta_1 = 1$, $\alpha_1 \beta_1 = 1$, $\det(I_p - K) = 0$, $s_0 = 1$ or $U^T (I_p - K)^{-1} U_0 = 1 - s_0$.

Proposition 4.1. *Assume that $\Theta \subset \Theta_4 \backslash \Theta^*$. Then, we have the decomposition*

$$
\sum_{t=1}^n \Phi_{t-1} X_t - n L_1 \;=\; (I_p - K)^{-1} U_0 \left(\sum_{t=1}^n X_t^2 - n \ell_0 \right) + M_n + R_n
$$

where M_n is a vector $\mathcal{F}_n^*$-martingale and the remainder term satisfies $\|R_n\| = o(\sqrt{n})$ a.s.

Proof. The proposition is established through Lemma A.2. Indeed, this statement shows that, as soon as $\Theta \cap \Theta^* = \varnothing$, there is a decomposition of the form

$$
\frac{1}{n} \sum_{t=1}^n \Phi_{t-1} X_t \;=\; (I_p - K)^{-1} U_0 \frac{1}{n} \sum_{t=1}^n X_t^2 + \frac{M_n}{n} + \frac{R_n}{n}.
$$

By ergodicity, taking the limit on both sides gives $L_1 = (I_p - K)^{-1} U_0 \ell_0$. The remainder term R_n is made of second-order isolated terms so, by Remark 2.1 and because $\Theta \subset \Theta_4$, we must have $\|R_n\| = o(\sqrt{n})$ a.s. $\qquad\square$ $\qquad\square$

Proposition 4.2. *Assume that $\Theta \subset \Theta_4 \backslash \Theta^*$. Then, we have the decomposition*

$$
\sum_{t=1}^n X_t^2 - n \ell_0 \;=\; M_n + R_n
$$

where M_n is a scalar $\mathcal{F}_n^*$-martingale and the remainder term satisfies $|R_n| = o(\sqrt{n})$ a.s.

Proof. This is a consequence of Lemma A.4. As soon as $\Theta \cap \Theta^* = \varnothing$, it follows from this lemma and Proposition 4.1, using the same notation, that

$$
\left(\sum_{t=1}^n X_t^2 - n \ell_0 \right) \left[1 - \frac{U^T (I_p - K)^{-1} U_0}{1 - s_0} \right] \;=\; \frac{U^T M_n + R_n}{1 - s_0}.
$$

7

This concludes the proof since, likewise, the remainder term R_n is a linear combination of second-order isolated terms. $\square$ $\square$

We now come back to the proof of Theorem 2.2. The keystone of the reasoning consists in noting that there is a matrix A such that

$$\sum_{t=1}^{n} \Phi_{t-1} X_t - S_{n-1} \theta^* = A \left(\sum_{t=1}^{n} \Phi_{t-1} X_t - n L_1 \right) - \theta^* \left(\sum_{t=1}^{n} X_t^2 - n \ell_0 \right) + R_n$$

since, by ergodicity, $0 = L_1 - \Lambda_0 \theta^* = A L_1 - \theta^* \ell_0$. The combination of Propositions 4.1 and 4.2 enables to obtain the decomposition

$$(4.2) \qquad\qquad \sum_{t=1}^{n} \Phi_{t-1} X_t - S_{n-1} \theta^* = M_n + R_n$$

where, as in the previous proofs, M_n is a vector $\mathcal{F}_n^*$-martingale and the remainder term satisfies $\|R_n\| = o(\sqrt{n})$ a.s. Let us call m_n a generic element of M_n. As can be seen from the details of the Appendix, it always takes the form of

$$(4.3) \qquad\qquad m_n = \sum_{t=1}^{n} X_{t-d_1}^{a_1} X_{t-d_2}^{a_2} \eta_{k,t-d_3}^{a_3} \eta_{\ell,t-d_4}^{a_4} \nu_t$$

where $0 < d_3, d_4 \leqslant d_1, d_2$, $a_i \in \{0,1\}$ and where the zero-mean random variable ν_t is identically distributed and independent of $\mathcal{F}_{t-1}^*$. Provided that $\sigma_4 < +\infty$ and $\tau_{k,8} < +\infty$ (included in Θ_4), the ergodicity arguments (2.6) together with the fact that $\Theta \subset \Theta_4$ show that there exists a matrix L satisfying $\langle M \rangle_n / n \to L$ a.s. From the causal representation of the process given in Proposition 2.1 and the hypothesis on Φ_0, the increments of m_n are also strictly stationary and ergodic. Thus, for any $x > 0$,

$$\frac{1}{n} \sum_{t=1}^{n} \mathbb{E}[(\Delta m_t)^2 \, \mathbb{1}_{\{|\Delta m_t| \geqslant x\}} \, | \, \mathcal{F}_{t-1}^*] \xrightarrow{\text{a.s.}} \mathbb{E}[(\Delta m_1)^2 \, \mathbb{1}_{\{|\Delta m_1| \geqslant x\}}].$$

Since $\mathbb{E}[(\Delta m_1)^2] < +\infty$, the right-hand side can be made arbitrarily small for $x \to +\infty$. Once again generalizing to M_n, we obtain *via* the same arguments that the Lindeberg's condition is satisfied. We are now ready to apply the central limit theorem for vector martingales, given e.g. by Cor. 2.1.10 of [Duflo, 1997], and get that $M_n / \sqrt{n}$ is asymptotically normal with mean 0 and covariance L. Together with Remark 2.1, that leads to

$$\frac{1}{\sqrt{n}} \left(\sum_{t=1}^{n} \Phi_{t-1} X_t - S_{n-1} \theta^* \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, L).$$

Finally, the a.s. convergence of $n \, S_{n-1}^{-1}$ to Λ_0^{-1} and Slutsky's lemma conclude the proof. $\square$

4.3. **Proof of Theorem 2.3.** Let us come back to the $\mathcal{F}_n^*$-martingale m_n given in (4.3). By the Hartman-Wintner law of the iterated logarithm for martingales (see *e.g.* [Stout, 1970]),

$$\limsup_{n \to +\infty} \frac{m_n}{\sqrt{2 n \ln \ln n}} = v_m \quad \text{and} \quad \liminf_{n \to +\infty} \frac{m_n}{\sqrt{2 n \ln \ln n}} = -v_m \quad \text{a.s.}$$

where $v_m = \mathbb{E}[(\Delta m_1)^2] < +\infty$, provided that $\Theta \subset \Theta_4$. The limit inferior was obtained by replacing m_n by $-m_n$, which share the same variance and martingale properties. Exploiting

8

the latter bounds and generalizing to M_n, we can deduce that

$$(4.4) \qquad \limsup_{n \to +\infty} \frac{\|M_n\|}{\sqrt{2\,n\,\ln\ln n}} < +\infty \quad \text{a.s.}$$

Because (4.2) implies $\|\widehat{\theta}_n - \theta^*\| \leqslant \|S_{n-1}^{-1}\|\,\|M_n + R_n\|$, it is now easy to conclude the proof via (4.4). Indeed, we recall that $n\,S_{n-1}^{-1}$ is convergent and that $\|R_n\| = o(\sqrt{n})$, obviously implying that it is also $o(\sqrt{n\,\ln\ln n})$ a.s. $\qquad\square$

APPENDIX A. SOME MARTINGALES DECOMPOSITIONS

For an easier reading, let us gather in this Appendix the proofs only based on calculations. Like in Section 4, M_n, δ_n and R_n will be generic terms, not necessarily identical from one line to another, designating $\mathcal{F}_n^*$-martingales, differences of $\mathcal{F}_n^*$-martingales and isolated terms, respectively, where $\mathcal{F}_n^*$ is the filtration defined in (4.1). To save place, we deliberately skip the proofs for most of them because they only consist of tedious but straightforward calculations. We assume in all the Appendix that $\Theta \subset \Theta_4 \backslash \Theta^*$.

Lemma A.1. *We have the decomposition*

$$\sum_{t=1}^{n} X_t^2\,\eta_{1,t} \;=\; \frac{2\,\tau_{1,2}}{1 - 2\,\beta_1}\left[\sum_{k=1}^{p}\theta_k\sum_{t=1}^{n}X_t\,X_{t-k+1} + \beta_2\sum_{t=1}^{n}X_t\,X_{t-1}\right] + M_n + R_n$$

where $\beta_1 = \alpha_1\,\tau_{1,2}$ and $\beta_2 = \alpha_2\,\tau_{2,2} + \ldots + \alpha_p\,\tau_{p,2}$.

Proof. Develop X_t using (1.1) to get that, for all t and after a lot of simplifications,

$$X_t^2\,\eta_{1,t} = 2\,\tau_{1,2}\left(\sum_{k=1}^{p}\theta_k\,X_{t-1}\,X_{t-k} + \sum_{k=2}^{p}\alpha_k\,\tau_{k,2}\,X_{t-k}\,X_{t-k-1} + \alpha_1\,X_{t-1}^2\,\eta_{1,t-1}\right) + \delta_t.$$

It remains to sum over t and to gather all equivalent terms. Note that $\{2\,\beta_1 = 1\} \subset \Theta^*$. $\qquad\square$

Lemma A.2. *We have the decomposition*

$$\sum_{t=1}^{n} X_t\,X_{t-1} \;=\; \frac{1}{1 - 2\,\beta_1}\left[\sum_{k=1}^{p}\theta_k\sum_{t=1}^{n}X_t\,X_{t-k+1} + \beta_2\sum_{t=1}^{n}X_t\,X_{t-1}\right] + M_n + R_n$$

where $\beta_1 = \alpha_1\,\tau_{1,2}$ and $\beta_2 = \alpha_2\,\tau_{2,2} + \ldots + \alpha_p\,\tau_{p,2}$. In addition, for $i \geqslant 2$, we have

$$\sum_{t=1}^{n} X_t\,X_{t-i} \;=\; \sum_{k=1}^{p}\theta_k\sum_{t=1}^{n}X_t\,X_{t-|i-k|} + \beta_1\sum_{t=1}^{n}X_t\,X_{t-|i-2|} + M_n + R_n.$$

Proof. Develop X_t using (1.1) to get, after additional calculations, that for all t,

$$X_t\,X_{t-i} = \sum_{k=1}^{p}\theta_k\,X_{t-i}\,X_{t-k} + \alpha_1\,\tau_{1,2}\,X_{t-2}\,X_{t-i} + \delta_t$$

as soon as $i \geqslant 2$. For $i = 1$, we can show that for all t,

$$X_t\,X_{t-1} = \sum_{k=1}^{p}\theta_k\,X_{t-1}\,X_{t-k} + \alpha_1\,X_{t-1}^2\,\eta_{1,t-1} + \sum_{k=2}^{p}\alpha_k\,\tau_{k,2}\,X_{t-k}\,X_{t-k-1} + \delta_t.$$

Finally, it remains to sum over t and to exploit Lemma A.1. Note that $\{2\,\beta_1 = 1\} \subset \Theta^*$. $\qquad\square$

Lemma A.3. *There exists $\gamma_0, \ldots, \gamma_{p-1}$ such that we have the decomposition*

$$\sum_{t=1}^{n} X_t^2 \, \eta_{1,t}^2 \;=\; \frac{1}{1 - \alpha_1 \beta_1} \left[\sum_{k=1}^{p} \gamma_{k-1} \sum_{t=1}^{n} X_t \, X_{t-k+1} + \sigma_2 \, \tau_{1,2} \, n \right] + M_n + R_n$$

where $\beta_1 = \alpha_1 \tau_{1,2}$.

Proof. We first obtain, using (1.1), that for all t,

$$X_t^2 \, \eta_{1,t}^2 = X_t \, \eta_{1,t}^2 \left(\sum_{k=1}^{p} \theta_k \, X_{t-k} + \sum_{k=1}^{p} X_{t-k} \, \eta_{k,t} + \sum_{k=1}^{p} \alpha_k \, X_{t-k} \, \eta_{k,t-1} \right) + \eta_{1,t}^2 \, \varepsilon_t^2 + \delta_t.$$

Then, the statements above lead to the result. Note that $\{\alpha_1 \beta_1 = 1\} \subset \Theta^*$. $\qquad\square$

Lemma A.4. *There exists $U^T = (u_1, \ldots, u_{p-1}, 0)$ such that we have the decomposition*

$$\sum_{t=1}^{n} X_t^2 \;=\; \frac{U^T}{1 - s_0} \sum_{t=1}^{n} \Phi_{t-1} \, X_t + \left(\ell_0 - \frac{U^T L_1}{1 - s_0} \right) n + M_n + R_n$$

where

$$s_0 = \left(1 + \alpha_1 \theta_2 + \frac{2 \, \alpha_1 \, \theta_1^2}{1 - 2 \, \beta_1} + \frac{\sigma_2}{1 - \alpha_1 \beta_1} \right) \tau_{1,2} + \sum_{k=2}^{p} (1 + \alpha_k^2) \, \tau_{k,2}$$

and $\beta_1 = \alpha_1 \tau_{1,2}$.

Proof. As in the previous proofs, we start by developping X_t using (1.1) to get, for all t,

$$X_t^2 = X_t \left(\sum_{k=1}^{p} \theta_k \, X_{t-k} + \sum_{k=1}^{p} X_{t-k} \, \eta_{k,t} + \sum_{k=1}^{p} \alpha_k \, X_{t-k} \, \eta_{k,t-1} \right) + \varepsilon_t^2 + \delta_t.$$

Finally, after some additional steps of calculations, Lemmas A.1 and A.3 enable to simplify the summation. Note that $\{2\,\beta_1 = 1\} \cup \{\alpha_1 \beta_1 = 1\} \cup \{s_0 = 1\} \subset \Theta^*$. $\qquad\square$

APPENDIX B. MOMENTS OF THE PROCESS

In this section, we repeatedly need the well-known relations $\text{vec}(AXB) = (B^T \otimes A) \, \text{vec}(X)$ and $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$. First, let us give more details on the condition $\mathbb{E}[\eta_{k,t}^a \, X_t^2] < +\infty$ for $a \in \{0, 1, 2\}$. From the vector form (2.1) of the process together with its $\mathcal{F}_t$-measurability (2.4), we may obtain that for all t, developing and taking expectation,

$$
\begin{aligned}
\mathbb{E}[\text{vec}(\Phi_t \, \Phi_t^T)] \;=\; & (C_\theta \otimes C_\theta + \Gamma_0) \, \mathbb{E}[\text{vec}(\Phi_{t-1} \, \Phi_{t-1}^T)] \\
& + (I_p \otimes C_\theta) \, \mathbb{E}[((N_{t-1} \, D_\alpha) \otimes I_p) \, \text{vec}(\Phi_{t-1} \, \Phi_{t-1}^T)] \\
& + (C_\theta \otimes I_p) \, \mathbb{E}[(I_p \otimes (N_{t-1} \, D_\alpha)) \, \text{vec}(\Phi_{t-1} \, \Phi_{t-1}^T)] \\
& + \mathbb{E}[((N_{t-1} \, D_\alpha) \otimes (N_{t-1} \, D_\alpha)) \, \text{vec}(\Phi_{t-1} \, \Phi_{t-1}^T)] + \text{vec}(\Sigma_0)
\end{aligned}
$$

where $\Gamma_0 = \mathbb{E}[N_0 \otimes N_0]$ and $\Sigma_0 = \mathbb{E}[E_0 \, E_0^T]$. In a more compact form,

$$U_t = (C_\theta \otimes C_\theta + \Gamma_0) \, U_{t-1} + (I_p \otimes C_\theta) \, V_{1,t-1} + (C_\theta \otimes I_p) \, V_{2,t-1} + W_{t-1} + \text{vec}(\Sigma_0)$$

where U_t, $V_{1,t}$, $V_{2,t}$ and W_t are easily identifiable from the explicit relation above. Working similarly on the other components and stacking them into Ω_t, we get the linear system

(B.1)
$$\Omega_t = A_2 \, \Omega_{t-1} + B_2$$

where $B_2^T = (\text{vec}(\Sigma_0), 0, 0, \Gamma_{\alpha\alpha} \text{vec}(\Sigma_0))$,

$$A_2 = \begin{pmatrix} C_\theta \otimes C_\theta + \Gamma_0 & I_p \otimes C_\theta & C_\theta \otimes I_p & I_{p^2} \\ G_\alpha \otimes C_\theta + \Gamma_{\alpha c} & \Gamma_\alpha & G_\alpha \otimes I_p & 0 \\ C_\theta \otimes G_\alpha + \Gamma'_{\alpha c} & I_p \otimes G_\alpha & \Gamma'_\alpha & 0 \\ \Gamma_{\alpha\alpha} (C_\theta \otimes C_\theta) + \Lambda_{\alpha\alpha} & \Gamma_{\alpha\alpha} (I_p \otimes C_\theta) & \Gamma_{\alpha\alpha} (C_\theta \otimes I_p) & \Gamma_{\alpha\alpha} \end{pmatrix}$$

and where, additionally, $G_\alpha = \mathbb{E}[N_0 D_\alpha N_0]$, $\Gamma_\alpha = \mathbb{E}[(N_0 D_\alpha) \otimes N_0]$, $\Gamma'_\alpha = \mathbb{E}[N_0 \otimes (N_0 D_\alpha)]$, $\Gamma_{\alpha c} = \mathbb{E}[(N_0 D_\alpha C_\theta) \otimes N_0]$, $\Gamma'_{\alpha c} = \mathbb{E}[N_0 \otimes (N_0 D_\alpha C_\theta)]$, $\Gamma_{\alpha\alpha} = \mathbb{E}[(N_0 D_\alpha) \otimes (N_0 D_\alpha)]$ and $\Lambda_{\alpha\alpha} = \mathbb{E}[(N_0 D_\alpha N_0) \otimes (N_0 D_\alpha N_0)]$. In virtue of (B.1), the condition $\rho(A_2) < 1$ is necessary and sufficient for the existence of the second-order moments of the process, like in [Nicholls and Quinn, 1981b]. Note that the fourth-order moments of $(\eta_{k,t})$ and the second-order moments of (ε_t) are involved.

Remark B.1. Let $\Omega = (I_{4p^2} - A_2)^{-1} B_2$ be the steady state of (B.1) with vertical blocks U, V_1, V_2 and W. Then, by ergodicity,

$$\frac{1}{n} \sum_{t=1}^{n} \Phi_{t-1} \Phi_{t-1}^T \xrightarrow{\text{a.s.}} \text{vec}^{-1}(U) \quad \text{and} \quad \frac{1}{n} \sum_{t=1}^{n} \Phi_{t-1} \Phi_{t-1}^T D_\alpha N_{t-1} \xrightarrow{\text{a.s.}} \text{vec}^{-1}(V_1)$$

where $\text{vec}^{-1} : \mathbb{R}^{p^2} \to \mathbb{R}^{p \times p}$ stands for the inverse operator of $\text{vec} : \mathbb{R}^{p \times p} \to \mathbb{R}^{p^2}$. Thus, we deduce from (2.1) and (2.5) that $\theta^* = \theta + \Lambda_0^{-1}\lambda$, where $\Lambda_0 = \text{vec}^{-1}(U)$, given in (2.8), and λ is the first column of $\text{vec}^{-1}(V_1)$. In particular, $\theta^* = \theta \Leftrightarrow \lambda = 0 \Leftarrow \alpha = 0$.

The condition $\mathbb{E}[\eta_{k,t}^a X_t^4] < +\infty$ for $a \in \{0, \dots, 4\}$ can also be treated using the same strategy. However, due to the extent of calculations, we cannot afford it in this paper and we just give an outline. Instead of working on $\mathbb{E}[\text{vec}(\Phi_t \Phi_t^T)]$ like in the case of the second-order moments, we have to start with the treatment of $\mathbb{E}[\text{vec}(\text{vec}(\Phi_t \Phi_t^T) \text{vec}^T(\Phi_t \Phi_t^T))]$ which shall lead, after an extremely long development, to another system of the form

(B.2) $$\Pi_t = A_4 \Pi_{t-1} + B_4.$$

Here, the fourth-order moments of (ε_t) and the eighth-order moments of $(\eta_{k,t})$ are involved. Finally, the condition of existence of the fourth-order moments of the process is $\rho(A_4) < 1$, as it is done in [Proïa and Soltane, 2018] for $p = 1$.

References

[Anděl, 1976] Anděl, J. (1976). Autoregressive series with random parameters. *Math. Operationsforsch. Statist.*, 7(5):735–741.

[Aue and Horváth, 2011] Aue, A. and Horváth, L. (2011). Quasi-likelihood estimation in stationary and nonstationary autoregressive models with random coefficients. *Stat. Sinica.*, 21:973–999.

[Aue et al., 2006] Aue, A., Horváth, L., and Steinebach, J. (2006). Estimation in random coefficient autoregressive models. *J. Time. Ser. Anal.*, 27(1):61–76.

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

[Berkes et al., 2009] Berkes, I., Horváth, L., and Ling, S. (2009). Estimation in nonstationary random coefficient autoregressive models. *J. Time. Ser. Anal.*, 30-4:395–416.

[Brandt, 1986] Brandt, A. (1986). The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Adv. Appl. Probab.*, 18:211–220.

[Brockwell and Davis, 1991] Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods. Second Edition.* Springer Series in Statistics. Springer-Verlag, New-York.

[Duflo, 1997] Duflo, M. (1997). *Random iterative models*, volume 34 of *Applications of Mathematics, New York*. Springer-Verlag, Berlin.

[Hill and Peng, 2014] Hill, J. and Peng, L. (2014). Unified interval estimation for random coefficient autoregressive models. *J. Time. Ser. Anal.*, 35(3):282–297.

[Koul and Schick, 1996] Koul, H. and Schick, A. (1996). Adaptative estimation in a random coefficient autoregressive model. *Ann. Stat.*, 24(3):1025–1052.

[Nicholls and Quinn, 1981a] Nicholls, D. F. and Quinn, B. G. (1981a). The estimation of multivariate random coefficient autoregressive models. *J. Multivar. Anal.*, 11:544–555.

[Nicholls and Quinn, 1981b] Nicholls, D. F. and Quinn, B. G. (1981b). Multiple autoregressive models with random coefficients. *J. Multivar. Anal.*, 11:185–198.

[Nicholls and Quinn, 1982] Nicholls, D. F. and Quinn, B. G. (1982). *Random Coefficient Autoregressive Models: An Introduction*, volume 11 of *Lecture Notes in Statistics*. Springer-Verlag, New-York.

[Proïa and Soltane, 2018] Proïa, F. and Soltane, M. (2018). A test of correlation in the random coefficients of an autoregressive process. *Math. Meth. Stat.*, 26(2):119–144.

[Schick, 1996] Schick, A. (1996). $\sqrt{n}$–consistent estimation in a random coefficient autoregressive model. *Austral. J. Statist.*, 38-2:155–160.

[Stout, 1970] Stout, W. F. (1970). The Hartman-Wintner law of the iterated logarithm for martingales. *Ann. Math. Stat.*, 41(6):2158–2160.

Email address: frederic.proia@univ-angers.fr
Email address: marius.soltane.etu@univ-lemans.fr

LABORATOIRE ANGEVIN DE RECHERCHE EN MATHÉMATIQUES (LAREMA), CNRS, UNIVERSITÉ D'ANGERS, UNIVERSITÉ BRETAGNE LOIRE. 2 BOULEVARD LAVOISIER, 49045 ANGERS CEDEX 01, FRANCE.

LABORATOIRE MANCEAU DE MATHÉMATIQUES, LE MANS UNIVERSITÉ, AVENUE O. MESSIAEN, 72085 LE MANS CEDEX 9, FRANCE.