



# A new statistical procedure for testing the presence of a significant correlation in the residuals of stable autoregressive processes

Frédéric Proia

## ► To cite this version:

Frédéric Proia. A new statistical procedure for testing the presence of a significant correlation in the residuals of stable autoregressive processes. 2012. hal-00677600

**HAL Id: hal-00677600**

**<https://hal.science/hal-00677600>**

Preprint submitted on 8 Mar 2012

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

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

# A NEW STATISTICAL PROCEDURE FOR TESTING THE PRESENCE OF A SIGNIFICATIVE CORRELATION IN THE RESIDUALS OF STABLE AUTOREGRESSIVE PROCESSES

FRÉDÉRIC PROÏA

**ABSTRACT.** The purpose of this paper is to investigate the asymptotic behavior of the Durbin-Watson statistic for the general stable  $p$ -order autoregressive process when the driven noise is given by a first-order autoregressive process. We establish the almost sure convergence and the asymptotic normality for both the least squares estimator of the unknown vector parameter of the autoregressive process as well as for the serial correlation estimator associated with the driven noise. In addition, the almost sure rates of convergence of our estimates are also provided. Then, we prove the almost sure convergence and the asymptotic normality for the Durbin-Watson statistic. Finally, we propose a new bilateral statistical procedure for testing the presence of a significative first-order residual autocorrelation and we also explain how our procedure performs better than the commonly used Box-Pierce and Ljung-Box statistical tests for white noise applied to the stable autoregressive process, even on small-sized samples.

## 1. INTRODUCTION

The Durbin-Watson statistic was originally introduced by the eponymous econometricians Durbin and Watson [15], [16], [17] in the middle of last century, in order to detect the presence of a significative first-order autocorrelation in the residuals from a regression analysis. The statistical test worked pretty well in the independent framework of linear regression models, as it was specifically investigated by Tillman [33]. While the Durbin-Watson statistic started to become well-known in Econometry by being commonly used in the case of linear regression models containing lagged dependent random variables, Malinvaud [27] and Nerlove and Wallis [28] observed that its widespread use in inappropriate situations were leading to inadequate conclusions. More precisely, they noticed that the Durbin-Watson statistic was asymptotically biased in the dependent framework. To remedy this misuse, alternative compromises were suggested. In particular, Durbin [13] proposed a set of revisions of the original test, as the so-called *t-test* and *h-test*, and explained how to use them focusing on the first-order autoregressive process. It inspired a lot of works afterwards. More precisely, Maddala and Rao [26], Park [29] and then Inder [22], [23] and Durbin [14] looked into the approximation of the critical values and distributions under the null hypothesis, and showed by simulations that alternative tests significantly outperformed the inappropriate one, even on small-sized samples. Additional improvements were brought by King and Wu [24] and lately, Stocker [30]

---

*Key words and phrases.* Durbin-Watson statistic, Stable autoregressive process, Residual autocorrelation, Statistical test for serial correlation.

gave substantial contributions to the study of the asymptotic bias resulting from the presence of lagged dependent random variables. In most cases, the first-order autoregressive process was used as a reference for related research. This is the reason why the recent work of Bercu and Proïa [4] was focused on such a process in order to give a new light on the distribution of the Durbin-Watson statistic under the null hypothesis as well as under the alternative hypothesis. They provided a sharp theoretical analysis rather than Monte-Carlo approximations, and they proposed a statistical procedure derived from the Durbin-Watson statistic. They showed how, from a theoretical and a practical point of view, this procedure outperforms the commonly used Box-Pierce [7] and Ljung-Box [6] statistical tests, in the restrictive case of the first-order autoregressive process, even on small-sized samples. They also explained that such a procedure is asymptotically equivalent to the *h-test* of Durbin [13] for testing the significance of the first-order serial correlation. This work [4] had the ambition to bring the Durbin-Watson statistic back into light. It also inspired Bitseki Penda, Djellout and Proïa [5] who established moderate deviation principles on the least squares estimators and the Durbin-Watson statistic for the first-order autoregressive process where the driven noise is also given by a first-order autoregressive process.

Our goal is to extend of the previous results of Bercu and Proïa [4] to  $p$ -order autoregressive processes, contributing moreover to the investigation on several open questions left unanswered during four decades on the Durbin-Watson statistic [13], [14], [28]. One will observe that the multivariate framework is much more difficult to handle than the scalar case of [4]. We will focus our attention on the  $p$ -order autoregressive process given, for all  $n \geq 1$ , by

$$(1.1) \quad \begin{cases} X_n &= \theta_1 X_{n-1} + \dots + \theta_p X_{n-p} + \varepsilon_n \\ \varepsilon_n &= \rho \varepsilon_{n-1} + V_n \end{cases}$$

where the unknown parameter  $\theta = (\theta_1 \ \theta_2 \ \dots \ \theta_p)'$  is a nonzero vector such that  $\|\theta\|_1 < 1$ , and the unknown parameter  $|\rho| < 1$ . Via an extensive use of the theory of martingales [11], [20], we shall provide a sharp and rigorous analysis on the asymptotic behavior of the least squares estimators of  $\theta$  and  $\rho$ . The previous results of convergence were first established in probability [27], [28], and more recently almost surely [4] in the particular case where  $p = 1$ . We shall prove the almost sure convergence as well as the asymptotic normality of the least squares estimators of  $\theta$  and  $\rho$  in the more general multivariate framework, together with the almost sure rates of convergence of our estimates. We will deduce the almost sure convergence and the asymptotic normality for the Durbin-Watson statistic. Therefore, we shall be in the position to propose a new statistical procedure for testing the significance of the first-order serial correlation in the residuals and to explain why, on the basis of the empirical power, this test procedure outperforms Ljung-Box [6] and Box-Pierce [7] *portmanteau* tests for stable autoregressive processes. We will also show by simulation that it is equally powerful than the Breusch-Godfrey [8], [18] test and the *h-test* [13] on large samples, and better than all of them on small samples.

The paper is organized as follows. Section 2 is devoted to the estimation of the autoregressive parameter. We establish the almost sure convergence of the least squares vector estimator  $\hat{\theta}_n$  to the limiting value

$$(1.2) \quad \theta^* = \alpha (I_p - \theta_p \rho J_p) \beta$$

where  $I_p$  is the identity matrix of order  $p$ ,  $J_p$  is the exchange matrix of order  $p$ , and where  $\alpha$  and  $\beta$  will be calculated explicitly. The asymptotic normality of  $\hat{\theta}_n$  as well as the quadratic strong law and a set of results derived from the law of iterated logarithm are provided. Section 3 deals with the estimation of the serial correlation parameter. The almost sure convergence of the least squares estimator  $\hat{\rho}_n$  to

$$(1.3) \quad \rho^* = \theta_p \rho \theta_p^*$$

where  $\theta_p^*$  stands for the  $p$ -th component of  $\theta^*$  is also established along with the quadratic strong law, the law of iterated logarithm and the asymptotic normality. It enables us to establish in Section 4 the almost sure convergence of the Durbin-Watson statistic  $\hat{D}_n$  to

$$(1.4) \quad D^* = 2(1 - \rho^*)$$

together with its asymptotic normality. Our sharp analysis on the asymptotic behavior of the Durbin-Watson statistic remains true whatever the values of the parameters  $\theta$  and  $\rho$  as soon as  $\|\theta\|_1 < 1$  and  $|\rho| < 1$ , assumptions resulting from the stability of the model. Consequently, we are able in Section 4 to propose a new bilateral statistical test for the presence of a significative first-order residual autocorrelation. In Section 5, we compare the empirical power of the test procedure based on the Durbin-Watson statistic with the commonly used *portmanteau* tests of Box-Pierce [7] and Ljung-Box [6], to corroborate the theoretical aspects and explain the reasons why we provide a more powerful alternative for testing the significance of the one-period lagged residuals coefficient, even on small-sized samples. The test procedure is also compared with the Breusch-Godfrey [8], [18] test and the *h-test* of Durbin [13]. Finally, the proofs related to linear algebra calculations are postponed in Appendix A and all the technical proofs of Sections 2 and 3 are postponed in Appendices B and C, respectively.

**Remark 1.1.** *In the whole paper, for any matrix  $M$ ,  $M'$  is the transpose of  $M$ . For any square matrix  $M$ ,  $\text{tr}(M)$ ,  $\det(M)$ ,  $\|M\|_1$  and  $\rho(M)$  are the trace, the determinant, the 1-norm and the spectral radius of  $M$ , respectively. In addition,  $\lambda_{\min}(M)$  and  $\lambda_{\max}(M)$  denote the smallest and the largest eigenvalues of  $M$ , respectively. For any vector  $v$ ,  $\|v\|$  stands for the euclidean norm of  $v$  and  $\|v\|_1$  is the 1-norm of  $v$ .*

**Remark 1.2.** *Before starting, we denote by  $I_p$  be the identity matrix of order  $p$ ,  $J_p$  the exchange matrix of order  $p$  and  $e$  the  $p$ -dimensional vector given by*

$$I_p = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}, \quad J_p = \begin{pmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 1 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 1 & \dots & 0 & 0 \end{pmatrix}, \quad e = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

## 2. ON THE AUTOREGRESSIVE PARAMETER

Consider the  $p$ -order autoregressive process given by (1.1) where we shall suppose, to make calculations lighter without loss of generality, that the square-integrable initial values  $X_0 = \varepsilon_0$  and  $X_{-1}, X_{-2}, \dots, X_{-p} = 0$ . In all the sequel, we assume that  $(V_n)$  is a sequence of square-integrable, independent and identically distributed random variables with zero mean and variance  $\sigma^2 > 0$ . Let us start by introducing some notations. Let  $\Phi_n^p$  stand for the lag vector of order  $p$ , given for all  $n \geq 0$ , by

$$(2.1) \quad \Phi_n^p = \begin{pmatrix} X_n & X_{n-1} & \dots & X_{n-p+1} \end{pmatrix}'.$$

Denote by  $S_n$  the positive definite matrix defined, for all  $n \geq 0$ , as

$$(2.2) \quad S_n = \sum_{k=0}^n \Phi_k^p \Phi_k^{p'} + S$$

where the symmetric and positive definite matrix  $S$  is added in order to avoid an useless invertibility assumption. For the estimation of the unknown parameter  $\theta$ , it is natural to make use of the least squares estimator which minimizes

$$\nabla_n(\theta) = \sum_{k=1}^n (X_k - \theta' \Phi_{k-1}^p)^2.$$

A standard calculation leads, for all  $n \geq 1$ , to

$$(2.3) \quad \hat{\theta}_n = (S_{n-1})^{-1} \sum_{k=1}^n \Phi_{k-1}^p X_k.$$

Our first result is related to the almost sure convergence of  $\hat{\theta}_n$  to the limiting value  $\theta^* = \alpha (I_p - \theta_p \rho J_p) \beta$ , where

$$(2.4) \quad \alpha = \frac{1}{(1 - \theta_p \rho)(1 + \theta_p \rho)},$$

$$(2.5) \quad \beta = \begin{pmatrix} \theta_1 + \rho & \theta_2 - \theta_1 \rho & \dots & \theta_p - \theta_{p-1} \rho \end{pmatrix}'.$$

**Theorem 2.1.** *We have the almost sure convergence*

$$(2.6) \quad \lim_{n \rightarrow \infty} \hat{\theta}_n = \theta^* \quad \text{a.s.}$$

**Remark 2.1.** *In the particular case where  $\rho = 0$ , we obtain the strong consistency of the least squares estimate in a stable autoregressive model, already proved e.g. in [25], under the condition of stability  $\|\theta\|_1 < 1$ .*

Let us now introduce the square matrix  $B$  of order  $p + 2$ , partially made of the elements of  $\beta$  given by (2.5),

$$(2.7) \quad B = \begin{pmatrix} 1 & -\beta_1 & -\beta_2 & \dots & \dots & -\beta_{p-1} & -\beta_p & \theta_p \rho \\ -\beta_1 & 1 - \beta_2 & -\beta_3 & \dots & \dots & -\beta_p & \theta_p \rho & 0 \\ -\beta_2 & -\beta_1 - \beta_3 & 1 - \beta_4 & \dots & \dots & \theta_p \rho & 0 & 0 \\ \vdots & \vdots & \vdots & & & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & & \vdots & \vdots & \vdots \\ -\beta_p & -\beta_{p-1} + \theta_p \rho & -\beta_{p-2} & \dots & \dots & -\beta_1 & 1 & 0 \\ \theta_p \rho & -\beta_p & -\beta_{p-1} & \dots & \dots & -\beta_2 & -\beta_1 & 1 \end{pmatrix}.$$

Under our stability conditions, we are able to establish the invertibility of  $B$  in Lemma 2.1. The corollary that follows will be useful in the next section.

**Lemma 2.1.** *Under the stability conditions  $\|\theta\|_1 < 1$  and  $|\rho| < 1$ , the matrix  $B$  given by (2.7) is invertible.*

**Corollary 2.1.** *By virtue of Lemma 2.1, the submatrix  $C$  obtained by removing from  $B$  its first row and first column is invertible.*

From now on,  $\Lambda \in \mathbb{R}^{p+2}$  is the unique solution of the linear system  $B\Lambda = e$ , i.e.

$$(2.8) \quad \Lambda = B^{-1}e$$

where the vector  $e$  has already been defined in Remark 1.1, but in higher dimension. Denote by  $\lambda_0, \dots, \lambda_{p+1}$  the elements of  $\Lambda$  and let  $\Delta_p$  be the Toeplitz matrix of order  $p$  associated with the first  $p$  elements of  $\Lambda$ , that is

$$(2.9) \quad \Delta_p = \begin{pmatrix} \lambda_0 & \lambda_1 & \lambda_2 & \dots & \dots & \lambda_{p-1} \\ \lambda_1 & \lambda_0 & \lambda_1 & \dots & \dots & \lambda_{p-2} \\ \vdots & \vdots & \vdots & & & \vdots \\ \vdots & \vdots & \vdots & & & \vdots \\ \lambda_{p-1} & \lambda_{p-2} & \lambda_{p-3} & \dots & \dots & \lambda_0 \end{pmatrix}.$$

Via the same lines, we are able to establish the invertibility of  $\Delta_p$  in Lemma 2.2.

**Lemma 2.2.** *Under the stability conditions  $\|\theta\|_1 < 1$  and  $|\rho| < 1$ , for all  $p \geq 1$ , the matrix  $\Delta_p$  given by (2.9) is positive definite.*

In light of foregoing, our next result deals with the asymptotic normality of  $\hat{\theta}_n$ .

**Theorem 2.2.** *Assume that  $(V_n)$  has a finite moment of order 4. Then, we have the asymptotic normality*

$$(2.10) \quad \sqrt{n} \left( \hat{\theta}_n - \theta^* \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_\theta)$$

where the asymptotic covariance matrix is given by

$$(2.11) \quad \Sigma_\theta = \alpha^2 (I_p - \theta_p \rho J_p) \Delta_p^{-1} (I_p - \theta_p \rho J_p).$$

**Remark 2.2.** *The covariance matrix  $\Sigma_\theta$  is invertible under the stability conditions. Furthermore, due to the way it is constructed,  $\Sigma_\theta$  is bisymmetric.*

After establishing the almost sure convergence of the estimator  $\widehat{\theta}_n$  and its asymptotic normality, we focus our attention on the almost sure rates of convergence.

**Theorem 2.3.** *Assume that  $(V_n)$  has a finite moment of order 4. Then, we have the quadratic strong law*

$$(2.12) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \left( \widehat{\theta}_k - \theta^* \right) \left( \widehat{\theta}_k - \theta^* \right)' = \Sigma_\theta \quad \text{a.s.}$$

where  $\Sigma_\theta$  is given by (2.11). In addition, for all  $v \in \mathbb{R}^p$ , we also have the law of iterated logarithm

$$(2.13) \quad \begin{aligned} \limsup_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{1/2} v' \left( \widehat{\theta}_n - \theta^* \right) &= - \liminf_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{1/2} v' \left( \widehat{\theta}_n - \theta^* \right), \\ &= (v' \Sigma_\theta v)^{1/2} \quad \text{a.s.} \end{aligned}$$

Consequently,

$$(2.14) \quad \limsup_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right) \left( \widehat{\theta}_n - \theta^* \right) \left( \widehat{\theta}_n - \theta^* \right)' = \Sigma_\theta \quad \text{a.s.}$$

In particular,

$$(2.15) \quad \limsup_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right) \left\| \widehat{\theta}_n - \theta^* \right\|^2 = \text{tr}(\Sigma_\theta) \quad \text{a.s.}$$

**Remark 2.3.** *It clearly follows from (2.12) that*

$$(2.16) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \left\| \widehat{\theta}_k - \theta^* \right\|^2 = \text{tr}(\Sigma_\theta) \quad \text{a.s.}$$

Furthermore, from (2.15), we have the almost sure rate of convergence

$$(2.17) \quad \left\| \widehat{\theta}_n - \theta^* \right\|^2 = O \left( \frac{\log \log n}{n} \right) \quad \text{a.s.}$$

*Proof.* The proofs of Lemma 2.1 and Lemma 2.2 are given in Appendix A while those of Theorems 2.1 to 2.3 may be found in Appendix B.  $\square$

To conclude this section, let us draw a parallel between the results of [4] and the latter results for  $p = 1$ . In this particular case,  $\beta$  and  $\alpha$  reduce to  $(\theta + \rho)$  and  $(1 - \theta\rho)^{-1}(1 + \theta\rho)^{-1}$  respectively, and it is not hard to see that we obtain the almost sure convergence of our estimate to

$$\theta^* = \frac{\theta + \rho}{1 + \theta\rho}.$$

In addition, a straightforward calculation leads to

$$\Sigma_\theta = \frac{(1 - \theta^2)(1 - \theta\rho)(1 - \rho^2)}{(1 + \theta\rho)^3}.$$

One can verify that these results correspond to Theorem 2.1 and Theorem 2.2 of [4].

### 3. ON THE SERIAL CORRELATION PARAMETER

This section is devoted to the estimation of the serial correlation parameter  $\rho$ . First of all, it is necessary to evaluate, at stage  $n$ , the residual set  $(\widehat{\varepsilon}_n)$  resulting from the biased estimation of  $\theta$ . For all  $1 \leq k \leq n$ , let

$$(3.1) \quad \widehat{\varepsilon}_k = X_k - \widehat{\theta}_n' \Phi_{k-1}^p.$$

The initial value  $\widehat{\varepsilon}_0$  may be arbitrarily chosen and we take  $\widehat{\varepsilon}_0 = X_0$  for a matter of simplification. Then, a natural way to estimate  $\rho$  is to make use of the least squares estimator which minimizes

$$\nabla_n(\rho) = \sum_{k=1}^n (\widehat{\varepsilon}_k - \rho \widehat{\varepsilon}_{k-1})^2.$$

Hence, it clearly follows that, for all  $n \geq 1$ ,

$$(3.2) \quad \widehat{\rho}_n = \frac{\sum_{k=1}^n \widehat{\varepsilon}_k \widehat{\varepsilon}_{k-1}}{\sum_{k=1}^n \widehat{\varepsilon}_{k-1}^2}.$$

It is important to note that one deals here with a scalar problem, in contrast to the study of the estimator of  $\theta$  in Section 2. Our goal is to obtain the same asymptotic properties for the estimator of  $\rho$  as those obtained for each component of the one of  $\theta$ . However, one shall realize that the results of this section are much more tricky to establish than those of the previous one.

We first state the almost sure convergence of  $\widehat{\rho}_n$  to the limiting value  $\rho^* = \theta_p \rho \theta_p^*$ .

**Theorem 3.1.** *We have the almost sure convergence*

$$(3.3) \quad \lim_{n \rightarrow \infty} \widehat{\rho}_n = \rho^* \quad \text{a.s.}$$

Our next result deals with the joint asymptotic normality of  $\widehat{\theta}_n$  and  $\widehat{\rho}_n$ . For that purpose, it is necessary to introduce some additional notations. Denote by  $P$  the square matrix of order  $p+1$  given by

$$(3.4) \quad P = \begin{pmatrix} P_B & 0 \\ P_L' & \varphi \end{pmatrix}$$

where

$$\begin{aligned} P_B &= \alpha(I_p - \theta_p \rho J_p) \Delta_p^{-1}, \\ P_L &= J_p(I_p - \theta_p \rho J_p) (\alpha \theta_p \rho \Delta_p^{-1} e + \theta_p^* \beta), \\ \varphi &= -\alpha^{-1} \theta_p^*. \end{aligned}$$

Furthermore, let us introduce the Toeplitz matrix  $\Delta_{p+1}$  of order  $p+1$  which is the extension of  $\Delta_p$  given by (2.9) to the next dimension,

$$(3.5) \quad \Delta_{p+1} = \begin{pmatrix} \Delta_p & J_p \Lambda_p^1 \\ \Lambda_p^{1'} J_p & \lambda_0 \end{pmatrix}$$



with  $\Lambda_p^1 = (\lambda_1 \ \lambda_2 \ \dots \ \lambda_p)'$ , and the positive semidefinite covariance matrix  $\Gamma$  of order  $p+1$ , given by

$$(3.6) \quad \Gamma = P \Delta_{p+1} P'.$$

**Theorem 3.2.** *Assume that  $(V_n)$  has a finite moment of order 4. Then, we have the joint asymptotic normality*

$$(3.7) \quad \sqrt{n} \begin{pmatrix} \hat{\theta}_n - \theta^* \\ \hat{\rho}_n - \rho^* \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma).$$

In particular,

$$(3.8) \quad \sqrt{n}(\hat{\rho}_n - \rho^*) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\rho^2)$$

where  $\sigma_\rho^2 = \Gamma_{p+1, p+1}$  is the last diagonal element of  $\Gamma$ .

**Remark 3.1.** *The covariance matrix  $\Gamma$  has the following explicit expression,*

$$\Gamma = \begin{pmatrix} \Sigma_\theta & \theta_p \rho J_p \Sigma_\theta e \\ \theta_p \rho e' \Sigma_\theta J_p & \sigma_\rho^2 \end{pmatrix}$$

where

$$(3.9) \quad \sigma_\rho^2 = P_L' \Delta_p P_L - 2\alpha^{-1} \theta_p^* \Lambda_p^{1'} J_p P_L + (\alpha^{-1} \theta_p^*)^2 \lambda_0.$$

**Remark 3.2.** *The covariance matrix  $\Gamma$  is invertible under the stability conditions if and only if  $\theta_p^* \neq 0$  since, by a straightforward calculation,*

$$\det(\Gamma) = \alpha^{2(p-1)} (\theta_p^*)^2 \det(\Delta_{p+1}) \left( \frac{\det(I_p - \theta_p \rho J_p)}{\det(\Delta_p)} \right)^2,$$

according to Lemma 2.2 and noticing that  $(I_p - \theta_p \rho J_p)$  is strictly diagonally dominant, thus invertible. As a result, the joint asymptotic normality given by (3.7) is degenerate in any situation such that  $\theta_p^* = 0$ , that is

$$(3.10) \quad \theta_p - \theta_{p-1} \rho = \theta_p \rho (\theta_1 + \rho).$$

Moreover, (3.8) holds on  $\{\theta_p - \theta_{p-1} \rho \neq \theta_p \rho (\theta_1 + \rho)\} \cup \{\theta_p \neq 0, \rho \neq 0\}$ , otherwise the asymptotic normality associated with  $\hat{\rho}_n$  is degenerate. In fact, a more restrictive condition ensuring that (3.8) still holds may be  $\{\theta_p \neq 0\}$ , i.e. that one deals at least with a  $p$ -order autoregressive process. This restriction seems natural in the context of the study and can be compared to the assumption  $\{\theta \neq 0\}$  in [4]. Theorem 3.2 of [4] ensures that the joint asymptotic normality is degenerate under  $\{\theta = -\rho\}$ . One can note that such an assumption is equivalent to (3.10) in the case of the  $p$ -order process, since both of them mean that the last component of  $\theta^*$  has to be nonzero.

The almost sure rates of convergence for  $\hat{\rho}_n$  are as follows.

**Theorem 3.3.** *Assume that  $(V_n)$  has a finite moment of order 4. Then, we have the quadratic strong law*

$$(3.11) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \left( \hat{\rho}_k - \rho^* \right)^2 = \sigma_\rho^2 \quad \text{a.s.}$$

where  $\sigma_\rho^2$  is given by (3.9). In addition, we also have the law of iterated logarithm

$$(3.12) \quad \begin{aligned} \limsup_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{1/2} \left( \hat{\rho}_n - \rho^* \right) &= - \liminf_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{1/2} \left( \hat{\rho}_n - \rho^* \right), \\ &= \sigma_\rho \quad \text{a.s.} \end{aligned}$$

Consequently,

$$(3.13) \quad \limsup_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right) \left( \hat{\rho}_n - \rho^* \right)^2 = \sigma_\rho^2 \quad \text{a.s.}$$

**Remark 3.3.** *It clearly follows from (3.13) that we have the almost sure rate of convergence*

$$(3.14) \quad \left( \hat{\rho}_n - \rho^* \right)^2 = O \left( \frac{\log \log n}{n} \right) \quad \text{a.s.}$$

As before, let us also draw the parallel between the results of [4] and the latter results for  $p = 1$ . In this particular case, we immediately obtain  $\rho^* = \theta \rho \theta^*$ . Moreover, an additionnal step of calculation shows that

$$\sigma_\rho^2 = \frac{1 - \theta \rho}{(1 + \theta \rho)^3} \left( (\theta + \rho)^2 (1 + \theta \rho)^2 + (\theta \rho)^2 (1 - \theta^2) (1 - \rho^2) \right).$$

One can verify that these results correspond to Theorem 3.1 and Theorem 3.2 of [4]. Besides, the estimators  $\hat{\theta}_n$  and  $\hat{\rho}_n$  are self-normalized. Consequently, the asymptotic variances  $\Sigma_\theta$  and  $\sigma_\rho^2$  do not depend on the variance  $\sigma^2$  associated with the driven noise  $(V_n)$ . To be complete and provide an important statistical aspect, it seemed advisable to suggest an estimator of the true variance  $\sigma^2$  of the model, based on  $\hat{\theta}_n$  and  $\hat{\rho}_n$ . For that purpose, consider, for all  $n \geq 1$ , the estimator given by

$$(3.15) \quad \hat{\sigma}_n^2 = \left( 1 - \hat{\rho}_n^2 \hat{\theta}_{p,n}^{-2} \right) \frac{1}{n} \sum_{k=0}^n \hat{\varepsilon}_k^2$$

where  $\hat{\theta}_{p,n}$  stands for the  $p$ -th component of  $\hat{\theta}_n$ .

**Theorem 3.4.** *We have the almost sure convergence*

$$(3.16) \quad \lim_{n \rightarrow \infty} \hat{\sigma}_n^2 = \sigma^2 \quad \text{a.s.}$$

*Proof.* The proofs of Theorems 3.1 to 3.3 are given in Appendix C. The one of Theorem 3.4 is left to the reader as it directly follows from that of Theorem 3.1.  $\square$

## 4. ON THE DURBIN-WATSON STATISTIC

We shall now investigate the asymptotic behavior of the Durbin-Watson statistic for the general autoregressive process [15], [16], [17], given, for all  $n \geq 1$ , by

$$(4.1) \quad \widehat{D}_n = \frac{\sum_{k=1}^n (\widehat{\varepsilon}_k - \widehat{\varepsilon}_{k-1})^2}{\sum_{k=0}^n \widehat{\varepsilon}_k^2}.$$

As mentioned, the almost sure convergence and the asymptotic normality of the Durbin-Watson statistic have previously been investigated in [4] in the particular case where  $p = 1$ . It has enabled the authors to propose a new bilateral statistical test for the presence of a significative residual autocorrelation. They also explained how this statistical procedure outperformed the commonly used Ljung-Box [6] and Box-Pierce [7] *portmanteau* tests for white noise in the case of the first-order autoregressive process, and how it was asymptotically equivalent to the *h-test* of Durbin [13], on a theoretical basis and on simulated data. They went even deeper in the study, establishing the distribution of the statistic under the null hypothesis “ $\rho = \rho_0$ ”, with  $|\rho_0| < 1$ , as well as under the alternative hypothesis “ $\rho \neq \rho_0$ ”, and noticing the existence of a critical situation in the case where  $\theta = -\rho$ . This pathological case arises when the covariance matrix  $\Gamma$  given by (3.6) is singular, and can be compared in the multivariate framework to the content of Remark 3.2. Our goal is to obtain the same asymptotic results for all  $p \geq 1$  so as to build a new statistical procedure for testing serial correlation in the residuals. In this paper, we shall only focus our attention on the test “ $\rho = 0$ ” against “ $\rho \neq 0$ ”, of increased statistical interest. In the next section, we will explain how, from a theoretical and a practical point of view, the procedure proposed in Theorem 4.4 is more powerful than the *portmanteau* tests [6], [7], for testing the significance of the first-order serial correlation of the driven noise in a  $p$ -order autoregressive process.

First, one can observe that  $\widehat{D}_n$  and  $\widehat{\rho}_n$  are asymptotically linked together by an affine transformation. Consequently, the asymptotic behavior of  $\widehat{D}_n$  directly follows from the previous section. We start with the almost sure convergence of  $\widehat{D}_n$  to the limiting value  $D^* = 2(1 - \rho^*)$ .

**Theorem 4.1.** *We have the almost sure convergence*

$$(4.2) \quad \lim_{n \rightarrow \infty} \widehat{D}_n = D^* \quad \text{a.s.}$$

Our next result deals with the asymptotic normality of  $\widehat{D}_n$ . It will be the keystone of a new bilateral statistical procedure deciding in particular, for a given significance level, whether residuals have a significative first-order correlation or not. Denote

$$(4.3) \quad \sigma_D^2 = 4\sigma_\rho^2$$

where the variance  $\sigma_\rho^2$  is given by (3.9).

**Theorem 4.2.** *Assume that  $(V_n)$  has a finite moment of order 4. Then, we have the asymptotic normality*

$$(4.4) \quad \sqrt{n} \left( \widehat{D}_n - D^* \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_D^2).$$

**Remark 4.1.** We immediately deduce from (4.4) that

$$(4.5) \quad \frac{n}{\sigma_D^2} \left( \widehat{D}_n - D^* \right)^2 \xrightarrow{\mathcal{L}} \chi^2$$

where  $\chi^2$  has a Chi-square distribution with one degree of freedom.

Let us focus now on the almost sure rates of convergence of  $\widehat{D}_n$ .

**Theorem 4.3.** Assume that  $(V_n)$  has a finite moment of order 4. Then, we have the quadratic strong law

$$(4.6) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \left( \widehat{D}_k - D^* \right)^2 = \sigma_D^2 \quad \text{a.s.}$$

where  $\sigma_D^2$  is given by (4.3). In addition, we also have the law of iterated logarithm

$$(4.7) \quad \begin{aligned} \limsup_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{1/2} \left( \widehat{D}_n - D^* \right) &= - \liminf_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{1/2} \left( \widehat{D}_n - D^* \right), \\ &= \sigma_D \quad \text{a.s.} \end{aligned}$$

Consequently,

$$(4.8) \quad \limsup_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right) \left( \widehat{D}_n - D^* \right)^2 = \sigma_D^2 \quad \text{a.s.}$$

**Remark 4.2.** It clearly follows from (4.8) that we have the almost sure rate of convergence

$$(4.9) \quad \left( \widehat{D}_n - D^* \right)^2 = O \left( \frac{\log \log n}{n} \right) \quad \text{a.s.}$$

We are now in the position to propose our new bilateral statistical test built on the Durbin-Watson statistic  $\widehat{D}_n$ . First of all, we shall not investigate the particular case where  $\theta_p = 0$  since our procedure is of interest only for autoregressive processes of order  $p$ . One wishes to test the presence of a significative serial correlation, setting

$$\mathcal{H}_0 : \text{"}\rho = 0\text{"} \quad \text{against} \quad \mathcal{H}_1 : \text{"}\rho \neq 0\text{"}.$$

**Theorem 4.4.** Assume that  $(V_n)$  has a finite moment of order 4,  $\theta_p \neq 0$  and  $\theta_p^* \neq 0$ . Then, under the null hypothesis  $\mathcal{H}_0 : \text{"}\rho = 0\text{"}$ ,

$$(4.10) \quad \frac{n}{4\widehat{\theta}_{p,n}^2} \left( \widehat{D}_n - 2 \right)^2 \xrightarrow{\mathcal{L}} \chi^2$$

where  $\widehat{\theta}_{p,n}$  stands for the  $p$ -th component of  $\widehat{\theta}_n$ , and where  $\chi^2$  has a Chi-square distribution with one degree of freedom. In addition, under the alternative hypothesis  $\mathcal{H}_1 : \text{"}\rho \neq 0\text{"}$ ,

$$(4.11) \quad \lim_{n \rightarrow \infty} \frac{n}{4\widehat{\theta}_{p,n}^2} \left( \widehat{D}_n - 2 \right)^2 = +\infty \quad \text{a.s.}$$

From a practical point of view, for a significance level  $a$  where  $0 < a < 1$ , the acceptance and rejection regions are given by  $\mathcal{A} = [0, z_a]$  and  $\mathcal{R} = ]z_a, +\infty[$  where  $z_a$  stands for the  $(1 - a)$ -quantile of the Chi-square distribution with one degree of freedom. The null hypothesis  $\mathcal{H}_0$  will not be rejected if the empirical value

$$\frac{n}{4\hat{\theta}_{p,n}^2} \left( \hat{D}_n - 2 \right)^2 \leq z_a,$$

and will be rejected otherwise.

**Remark 4.3.** *In the particular case where  $\theta_p^* = 0$ , the test statistic do not respond under  $\mathcal{H}_1$  as described above. To avoid such situation, we suggest to make use of Theorem 2.2 for testing beforehand whether  $\hat{\theta}_{p,n}$  is significantly far from zero. Besides, testing  $\mathcal{H}_0 : \rho = 0$  with  $\theta_p^* = 0$  amounts to testing the significance of the  $p$ -th coefficient of the model, not rejected under  $\{\theta_p \neq 0\}$ . Roughly speaking, under  $\{\theta_p \neq 0\} \cap \{\theta_p^* = 0\}$ , we obviously have  $\rho \neq 0$  and the use of Theorem 4.4 would be irrelevant since  $\mathcal{H}_1$  is certainly true.*

*Proof.* The proofs of Theorems 4.1 to 4.3 are left to the reader as they follow essentially the same lines as those given in Appendix C of [4]. Theorem 4.4 is an immediate consequence of Theorem 4.2, noticing that  $\sigma_\rho^2$  reduces to  $\theta_p^2$  under  $\mathcal{H}_0$  and using the same methodology as in the proof of Theorem 3.1.  $\square$

## 5. EMPIRICAL POWER AND COMPARISONS

Following the same methodology as in Section 5 of [4] and also being inspired by the empirical work of Park [29], this section is devoted to the comparison of our statistical test procedure with the statistical tests commonly used in time series analysis to detect the presence of a significative first-order correlation in the residuals. We will distinguish two kinds of tests, depending on the fact that the serial dependence has been included in the proposed model structure, or not. The first class includes the alternative *h-test* of Durbin [13] and the procedure described in Theorem 4.4 whereas the second one includes the Box-Pierce [7] and Ljung-Box [6] *portmanteau* tests as well as the Breusch-Godfrey [8], [18] procedure. It seems natural to imagine that, with the aim of detecting a significative serial correlation in the residuals, the first class will outperform the second one since more information is taken into account to build the test statistics, as it is explained and observed in Section 5 of [4]. In fact, the Breusch-Godfrey procedure is clearly better than the *portmanteau* tests for the autoregressive modelling since the regression on the estimated residual set  $(\hat{\varepsilon}_n)$  is not only fitted by least squares on its past, but also includes the regressors of the original model, making a link between both structures. Our objective is to show by simulations that our procedure is asymptotically equivalent to the *h-test* and to the one of Breusch-Godfrey, and that it appears to be more powerful on small-sized samples, for testing the significance of the first-order serial correlation in

the residuals of a  $p$ -order autoregressive process. We shall assume in all the sequel that  $\theta_p \neq 0$  is a statistically significant parameter and we will test

$$\mathcal{H}_0 : “\rho = 0” \quad \text{against} \quad \mathcal{H}_1 : “\rho \neq 0”.$$

For a given  $|\rho| < 1$  and for each simulation, we compute the test statistics associated with all of these 5 procedures and we test  $\mathcal{H}_0$  against  $\mathcal{H}_1$ , at the 0.05 level of significance. We repeat the experiment 1000 times and the frequency with which  $\mathcal{H}_0$  is rejected provides the empirical power of each test, i.e. an estimator of

$$\mathbb{P}(\text{rejecting } \mathcal{H}_0 \mid \mathcal{H}_1 \text{ is true}).$$

We choose  $p = 3$ ,  $\theta = (0.1 \quad -0.2 \quad 0.6)'$  and we take  $(V_n)$  independent with standard  $\mathcal{N}(0, 1)$  distribution since the variance  $\sigma^2$  has no influence on the Durbin-Watson statistic. Finally, sample sizes  $n = 30$  and  $n = 500$  are used in order to evaluate the empirical power on small-sized samples as well as asymptotically. The first 50 observations are discarded and the remaining ones are taken as a sample to minimize the impact of the initial values.

$n = 30$	Values of $\rho$										
	-0.7	-0.5	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.5	0.7
Box-Pierce	0.01 (0.99)	0.01 (0.99)	0.01 (0.99)	0.01 (0.99)	0.01 (0.99)	- (0.99)	0.01 (0.99)	0.02 (0.98)	0.02 (0.98)	0.04 (0.96)	0.06 (0.94)
Ljung-Box	0.01 (0.99)	0.01 (0.99)	0.01 (0.99)	0.01 (0.99)	0.01 (0.99)	- (0.99)	0.02 (0.98)	0.03 (0.97)	0.03 (0.97)	0.05 (0.95)	0.07 (0.93)
Breusch-Godfrey	0.30 (0.70)	0.18 (0.82)	0.10 (0.90)	0.06 (0.94)	0.05 (0.95)	- (0.95)	0.06 (0.94)	0.10 (0.90)	0.14 (0.86)	0.31 (0.69)	0.46 (0.54)
H-Test	0.13 (0.87)	0.13 (0.87)	0.13 (0.87)	0.13 (0.87)	0.14 (0.86)	- (0.85)	0.18 (0.82)	0.21 (0.79)	0.25 (0.75)	0.32 (0.69)	0.35 (0.65)
Durbin-Watson	0.48 (0.52)	0.34 (0.66)	0.22 (0.79)	0.18 (0.82)	0.17 (0.83)	- (0.84)	0.22 (0.78)	0.25 (0.75)	0.33 (0.67)	0.46 (0.54)	0.61 (0.39)

Tab. 5.1. Figures give the power of the tests for  $n = 30$ , the frequencies of  $\mathcal{H}_0$  not rejected are in parentheses.

$n = 500$	Values of $\rho$										
	-0.7	-0.5	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.5	0.7
Box-Pierce	0.71 (0.29)	0.94 (0.06)	0.54 (0.47)	0.17 (0.83)	0.02 (0.98)	- (1.00)	0.03 (0.97)	0.32 (0.68)	0.78 (0.22)	1.00 (0.00)	1.00 (0.00)
Ljung-Box	0.71 (0.29)	0.94 (0.06)	0.54 (0.46)	0.17 (0.83)	0.02 (0.98)	- (1.00)	0.03 (0.97)	0.32 (0.68)	0.78 (0.22)	1.00 (0.00)	1.00 (0.00)
Breusch-Godfrey	1.00 (0.00)	1.00 (0.00)	0.98 (0.02)	0.75 (0.25)	0.25 (0.75)	- (0.95)	0.28 (0.72)	0.75 (0.26)	0.98 (0.02)	1.00 (0.00)	1.00 (0.00)
H-Test	0.99 (0.01)	1.00 (0.00)	0.96 (0.04)	0.72 (0.28)	0.22 (0.78)	- (0.95)	0.33 (0.67)	0.76 (0.24)	0.98 (0.02)	1.00 (0.00)	1.00 (0.00)
Durbin-Watson	1.00 (0.00)	1.00 (0.00)	0.96 (0.04)	0.71 (0.29)	0.22 (0.78)	- (0.96)	0.32 (0.68)	0.76 (0.24)	0.98 (0.02)	1.00 (0.00)	1.00 (0.00)

Tab. 5.2. Figures give the power of the tests for  $n = 500$ , the frequencies of  $\mathcal{H}_0$  not rejected are in parentheses.

The graphs below show the frequencies of  $\mathcal{H}_0$  not rejected when  $\rho$  varies from  $-0.7$  to  $0.7$ , for  $n = 30$  and then for  $n = 500$ . BP, LB, BG, HT and DW respectively stand for Box-Pierce, Ljung-Box, Breusch-Godfrey, H-Test and Durbin-Watson.

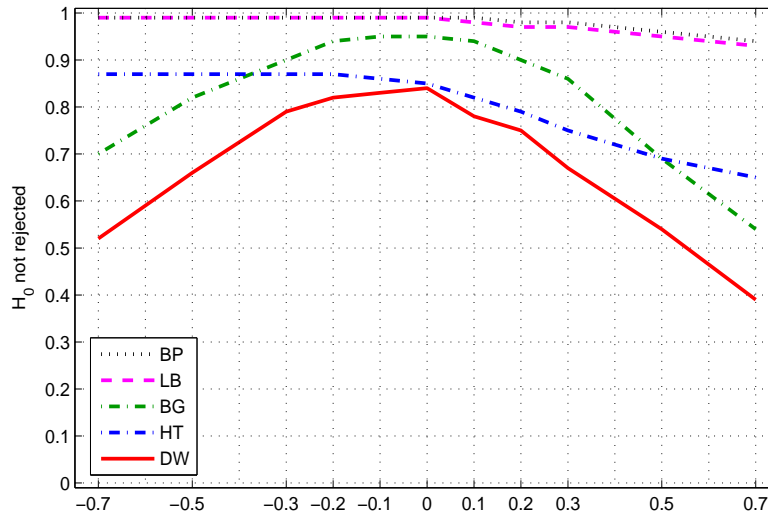


Fig. 5.1. Frequencies of  $\mathcal{H}_0$  not rejected for sample size  $n = 30$  and  $\rho$  varying from  $-0.7$  to  $0.7$  on the abscissa.

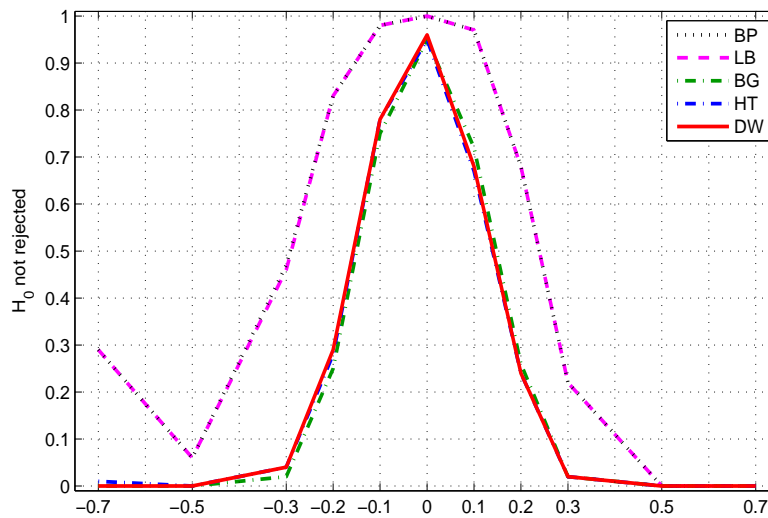


Fig. 5.2. Frequencies of  $\mathcal{H}_0$  not rejected for sample size  $n = 500$  and  $\rho$  varying from  $-0.7$  to  $0.7$  on the abscissa.

On small-sized samples, Fig 5.1 shows that our procedure outperforms all tests by always being more sensitive to the presence of correlation in the residuals, except under  $\mathcal{H}_0$  even if the 84% of non-rejection are quite satisfying. Indeed, the Box-Pierce and Ljung-Box test statistics are invariably smaller than they should have been in such a model, underestimating the alternative  $\mathcal{H}_1$  hypothesis, as it is explained in [4]. Fig. 5.2 clearly shows the asymptotic equivalence between the *h-test*, the Breusch-Godfrey test and our statistical procedure, as well as the superiority over the commonly used *portmanteau* tests, except under  $\mathcal{H}_0$  for the same reason. In conclusion, we suggest to make use of the statistical procedure of Theorem 4.4 on small-sized samples, to test the significance of the first-order serial correlation in the residuals, clearly more powerful. In addition, we note that this procedure is an

alternative to the Breusch-Godfrey test on large samples, and has to be advised instead of *portmanteau* tests in our particular framework. In order to keep this section brief, we have chosen  $p = 3$  and a particular value of  $\theta$ . However, all conclusions still remain the same when  $p$  and  $\theta$  vary, as long as the stability condition  $\|\theta\|_1 < 1$  and the significance of  $\theta_p \neq 0$  are satisfied. Nevertheless, one is supposed to have checked beforehand that he was at least dealing with a  $p$ -order autoregressive process. Moreover, our procedure is not much impacted by the critical situation and no unexpected peak occurs, contrary to the *portmanteau* tests. To conclude, note that when  $|\theta_p|$  is close to 1, all tests are asymptotically equivalent.

## APPENDIX A

### ON SOME LINEAR ALGEBRA CALCULATIONS

#### A.1. Proof of Lemma 2.1.

We start with the proof of Lemma 2.1. Our goal is to show that the matrix  $B$  given by (2.7) is invertible. Consider the decomposition  $B = B_1 + \rho B_2$ , where

$$B_1 = \begin{pmatrix} 1 & -\theta_1 & -\theta_2 & \dots & \dots & -\theta_{p-1} & -\theta_p & 0 \\ -\theta_1 & 1 - \theta_2 & -\theta_3 & \dots & \dots & -\theta_p & 0 & 0 \\ -\theta_2 & -\theta_1 - \theta_3 & 1 - \theta_4 & \dots & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & & \vdots & \vdots & \vdots \\ -\theta_p & -\theta_{p-1} & -\theta_{p-2} & \dots & \dots & -\theta_1 & 1 & 0 \\ 0 & -\theta_p & -\theta_{p-1} & \dots & \dots & -\theta_2 & -\theta_1 & 1 \end{pmatrix},$$

$$B_2 = \begin{pmatrix} 0 & -1 & \theta_1 & \dots & \dots & \theta_{p-2} & \theta_{p-1} & \theta_p \\ -1 & \theta_1 & \theta_2 & \dots & \dots & \theta_{p-1} & \theta_p & 0 \\ \theta_1 & -1 + \theta_2 & \theta_3 & \dots & \dots & \theta_p & 0 & 0 \\ \vdots & \vdots & \vdots & & & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & & \vdots & \vdots & \vdots \\ \theta_{p-1} & \theta_{p-2} + \theta_p & \theta_{p-3} & \dots & \dots & -1 & 0 & 0 \\ \theta_p & \theta_{p-1} & \theta_{p-2} & \dots & \dots & \theta_1 & -1 & 0 \end{pmatrix}.$$

It is trivial to see that  $|\theta_i + \theta_j| \leq |\theta_i| + |\theta_j|$  for all  $1 \leq i, j \leq p$ , and the same goes for  $1 - |\theta_i| \leq |1 - \theta_i|$ . These inequalities immediately imply that  $B_1$  is strictly diagonally dominant, and thus invertible by virtue of Levy-Desplanques' theorem 6.1.10 of [21]. Hence,  $B = (I_{p+2} + \rho B_2 B_1^{-1}) B_1$  and the invertibility of  $B$  only depends on the spectral radius of  $\rho B_2 B_1^{-1}$ , i.e. the supremum modulus of its eigenvalues. One



can explicitly obtain, by a straightforward calculation, that

$$B_2 B_1^{-1} = \begin{pmatrix} -\theta_1 & -1 - \theta_2 & \theta_1 - \theta_3 & \dots & \theta_{p-2} - \theta_p & \theta_{p-1} & \theta_p \\ -1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & -1 & 0 & \dots & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ 0 & \dots & \dots & 0 & -1 & 0 & 0 \\ 0 & \dots & \dots & \dots & 0 & -1 & 0 \end{pmatrix}.$$

The sum of the first row of  $B_2 B_1^{-1}$  is  $-1$ , involving *de facto* that  $-1$  is an eigenvalue of  $B_2 B_1^{-1}$  associated with the  $(p+2)$ -dimensional eigenvector  $(1 \ 1 \ \dots \ 1)'$ . By the same way, it is clear that  $1$  is an eigenvalue of  $B_2 B_1^{-1}$  associated with the eigenvector  $(1 \ -1 \ \dots \ (-1)^{p+1})'$ . Let  $P(\lambda) = \det(B_2 B_1^{-1} - \lambda I_{p+2})$  be the characteristic polynomial of  $B_2 B_1^{-1}$ . Then,  $P(\lambda)$  is recursively computable and explicitly given by

$$(A.1) \quad P(\lambda) = (-\lambda)^{p+2} + \sum_{k=1}^{p+2} b_k (-\lambda)^{p+2-k}$$

where  $(b_k)$  designates, for  $k \in \{1, \dots, p+2\}$ , the elements of the first line of  $B_2 B_1^{-1}$ . Since  $-1$  and  $1$  are zeroes of  $P(\lambda)$ , there exists a polynomial  $Q(\lambda)$  of degree  $p$  such that  $P(\lambda) = (\lambda^2 - 1)Q(\lambda)$ , and a direct calculation shows that  $Q$  is given by

$$(A.2) \quad Q(\lambda) = (-\lambda)^p - \sum_{k=1}^p \theta_k (-\lambda)^{p-k}.$$

Furthermore, let  $R(\lambda)$  be the polynomial of degree  $p$  defined as

$$(A.3) \quad R(\lambda) = \lambda^p - \sum_{k=1}^p |\theta_k| \lambda^{p-k},$$

and note that we clearly have  $R(|\lambda|) \leq |Q(\lambda)|$ , for all  $\lambda \in \mathbb{C}$ . Assume that  $\lambda_0 \in \mathbb{C}$  is an eigenvalue of  $B_2 B_1^{-1}$  such that  $|\lambda_0| > 1$ . Then,

$$\begin{aligned} R(|\lambda_0|) &= |\lambda_0|^p - \sum_{k=1}^p |\theta_k| |\lambda_0|^{p-k} = |\lambda_0|^p \left( 1 - \sum_{k=1}^p |\theta_k| |\lambda_0|^{-k} \right), \\ &\geq |\lambda_0|^p \left( 1 - \sum_{k=1}^p |\theta_k| \right) > 0 \end{aligned}$$

as soon as  $\|\theta\|_1 < 1$ . Consequently,  $|Q(\lambda_0)| > 0$ . This obviously contradicts the hypothesis that  $\lambda_0$  is an eigenvalue of  $B_2 B_1^{-1}$ . In conclusion, all the zeroes of  $Q(\lambda)$  lie in the unit circle, implying  $\rho(B_2 B_1^{-1}) \leq 1$ . Since  $1$  and  $-1$  are eigenvalues of  $B_2 B_1^{-1}$ , we have precisely  $\rho(B_2 B_1^{-1}) = 1$ , and therefore  $\rho(\rho B_2 B_1^{-1}) = |\rho| < 1$ . This guarantees the invertibility of  $B$  under the stability conditions, achieving the proof of Lemma 2.1. Finally, Corollary 2.1 immediately follows from Lemma 2.1. As a matter of fact, since  $B$  is invertible, we have  $\det(B) \neq 0$ . Denote by  $b$  the first

diagonal element of  $B^{-1}$ . Since  $\det(C)$  is the cofactor of the first diagonal element of  $B$ , we have

$$(A.4) \quad b = \frac{\det(C)}{\det(B)}.$$

However, it follows from (2.8) that  $b = \lambda_0$ . We shall prove in the next subsection that the matrix  $\Delta_p$  given by (2.9) is positive definite. It clearly implies that  $\lambda_0 > 0$  which means that  $b > 0$ , so  $\det(C) \neq 0$ , and the matrix  $C$  is invertible.  $\square$

## A.2. Proof of Lemma 2.2.

Let us start by proving that the spectral radius of the companion matrix associated with model (1.1) is strictly less than 1. By virtue of the fundamental autoregressive equation (B.8) detailed in the next section, the system (1.1) can be rewritten in the vectorial form, for all  $n \geq p + 1$ ,

$$(A.5) \quad \Phi_n^{p+1} = C_A \Phi_{n-1}^{p+1} + W_n$$

where  $\Phi_n^{p+1}$  stands for the extension of  $\Phi_n^p$  given by (2.1) to the next dimension,  $W_n = (V_n \ 0 \ \dots \ 0)'$  and where the companion matrix of order  $p + 1$

$$(A.6) \quad C_A = \begin{pmatrix} \theta_1 + \rho & \theta_2 - \theta_1 \rho & \dots & \theta_p - \theta_{p-1} \rho & -\theta_p \rho \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

Let  $P_A(\mu) = \det(C_A - \mu I_{p+1})$  be the characteristic polynomial of  $C_A$ . Then, it follows from Lemma 4.1.1 of [11] that

$$(A.7) \quad \begin{aligned} P_A(\mu) &= (-1)^p \left( \mu^{p+1} - (\theta_1 + \rho) \mu^p - \sum_{k=2}^p (\theta_k - \theta_{k-1} \rho) \mu^{p+1-k} + \theta_p \rho \right), \\ &= (-1)^p (\mu - \rho) \left( \mu^p - \sum_{k=1}^p \theta_k \mu^{p-k} \right) = (-1)^p (\mu - \rho) P(\mu) \end{aligned}$$

where the polynomial

$$P(\mu) = \mu^p - \sum_{k=1}^p \theta_k \mu^{p-k}.$$

Assume that  $\mu_0 \in \mathbb{C}$  is an eigenvalue of  $C_A$  such that  $|\mu_0| \geq 1$ . Then, under the stability condition  $|\rho| < 1$ , we obviously have  $\mu_0 \neq \rho$ . Consequently, we obtain that  $P(\mu_0) = 0$  which implies, since  $\mu_0 \neq 0$ , that

$$(A.8) \quad 1 - \sum_{k=1}^p \theta_k \mu_0^{-k} = 0.$$

Nevertheless,

$$\left| \sum_{k=1}^p \theta_k \mu_0^{-k} \right| \leq \sum_{k=1}^p |\theta_k| |\mu_0^{-k}| \leq \sum_{k=1}^p |\theta_k| < 1$$

as soon as  $\|\theta\|_1 < 1$  which contradicts (A.8). Hence,  $\rho(C_A) < 1$  under the stability conditions  $\|\theta\|_1 < 1$  and  $|\rho| < 1$ . Hereafter, let  $(Y_n)$  be the stationary autoregressive process satisfying, for all  $n \geq p+1$ ,

$$(A.9) \quad \Psi_n^{p+1} = C_A \Psi_{n-1}^{p+1} + W_n$$

where

$$\Psi_n^{p+1} = (Y_n \quad Y_{n-1} \quad \dots \quad Y_{n-p})'.$$

It follows from (A.9) that, for all  $n \geq p+1$ ,

$$Y_n = (\theta_1 + \rho)Y_{n-1} + \sum_{k=2}^p (\theta_k - \theta_{k-1}\rho)Y_{n-k} - \theta_p \rho Y_{n-p-1} + V_n.$$

By virtue of Theorem 4.4.2 of [9], the spectral density of the process  $(Y_n)$  is given, for all  $x$  in the torus  $\mathbb{T} = [-\pi, \pi]$ , by

$$(A.10) \quad f_Y(x) = \frac{\sigma^2}{2\pi |A(e^{-ix})|^2}$$

where the polynomial  $A$  is defined, for all  $\mu \neq 0$ , as

$$(A.11) \quad A(\mu) = \mu^{p+1} P_A(\mu^{-1}),$$

in which  $P_A$  is the polynomial given in (A.7), and  $A(0) = (-1)^p$ . In light of foregoing,  $A$  has no zeroes on the unit circle. In addition, for all  $k \in \mathbb{Z}$ , denote by

$$\hat{f}_k = \int_{\mathbb{T}} f_Y(x) e^{-ikx} dx$$

the Fourier coefficient of order  $k$  associated with  $f_Y$ . It is well-known that, for all  $p \geq 1$ , the covariance matrix of the vector  $\Psi_n^p$  coincides with the Toeplitz matrix of order  $p$  of the spectral density  $f_Y$  in (A.10). More precisely, for all  $p \geq 1$ , we have

$$(A.12) \quad T_p(f_Y) = \left( \hat{f}_{i-j} \right)_{1 \leq i, j \leq p} = \Delta_p$$

where  $\Delta_p$  is given by (2.9) and  $T$  stands for the Toeplitz operator. As a matter of fact, since  $\rho(C_A) < 1$ , we have

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \Phi_n^p \Phi_n^{p'} \right] = \mathbb{E} \left[ \Psi_p^p \Psi_p^{p'} \right] = \Delta_p.$$

Finally, we deduce from Proposition 4.5.3 of [9] or [19] that

$$(A.13) \quad 2\pi m_f \leq \lambda_{\min}(T_p(f_Y)) \leq \lambda_{\max}(T_p(f_Y)) \leq 2\pi M_f$$

where

$$m_f = \min_{x \in \mathbb{T}} f_Y(x) \quad \text{and} \quad M_f = \max_{x \in \mathbb{T}} f_Y(x).$$

Therefore, as  $m_f > 0$ ,  $T_p(f_Y)$  is positive definite, which clearly ensures that for all  $p \geq 1$ ,  $\Delta_p$  is also positive definite. This achieves the proof of Lemma 2.2.  $\square$

## APPENDIX B

## PROOFS OF THE AUTOREGRESSIVE PARAMETER RESULTS

**B.1. Preliminary Lemmas.**

We start with some useful technical lemmas we shall make repeatedly use of. The proof of Lemma B.1 may be found in the one of Corollary 1.3.21 in [11].

**Lemma B.1.** *Assume that  $(V_n)$  is a sequence of independent and identically distributed random variables such that, for some  $a \geq 1$ ,  $\mathbb{E}[|V_1|^a]$  is finite. Then,*

$$(B.1) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n |V_k|^a = \mathbb{E}[|V_1|^a] \quad \text{a.s.}$$

and

$$(B.2) \quad \sup_{1 \leq k \leq n} |V_k| = o(n^{1/a}) \quad \text{a.s.}$$

**Lemma B.2.** *Assume that  $(V_n)$  is a sequence of independent and identically distributed random variables such that, for some  $a \geq 1$ ,  $\mathbb{E}[|V_1|^a]$  is finite. If  $(X_n)$  satisfies (1.1) with  $\|\theta\|_1 < 1$  and  $|\rho| < 1$ , then*

$$(B.3) \quad \sum_{k=0}^n |X_k|^a = O(n) \quad \text{a.s.}$$

and

$$(B.4) \quad \sup_{0 \leq k \leq n} |X_k| = o(n^{1/a}) \quad \text{a.s.}$$

**Remark B.1.** *In the particular case where  $a = 4$ , we obtain that*

$$\sum_{k=0}^n X_k^4 = O(n) \quad \text{and} \quad \sup_{0 \leq k \leq n} X_k^2 = o(\sqrt{n}) \quad \text{a.s.}$$

*Proof.* The reader may find an approach following essentially the same lines in the proof of Lemma A.2 in [4], merely considering the stability condition  $\|\theta\|_1 < 1$  in lieu of  $|\theta| < 1$ .  $\square$

**Lemma B.3.** *Assume that the initial values  $X_0, X_1, \dots, X_{p-1}$  with  $\varepsilon_0 = X_0$  are square-integrable and that  $(V_n)$  is a sequence of independent and identically distributed random variables with zero mean and variance  $\sigma^2 > 0$ . Then, under the stability conditions  $\|\theta\|_1 < 1$  and  $|\rho| < 1$ , we have the almost sure convergence*

$$(B.5) \quad \lim_{n \rightarrow \infty} \frac{S_n}{n} = \sigma^2 \Delta_p \quad \text{a.s.}$$

where the matrix  $\Delta_p$  is given by (2.9).

*Proof.* By adopting the same approach as the one used to prove Theorem 2.2 in [4], it follows from the fundamental autoregressive equation (B.8), that will be detailed in the next section, that for all  $0 \leq d \leq p+1$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_{k-d} V_k = \sigma^2 \delta_d \quad \text{a.s.}$$

where  $\delta_d$  stands for the Kronecker delta function equal to 1 when  $d = 0$ , and 0 otherwise. Denote by  $\ell_d$  the limiting value which verifies, by virtue of Lemma B.2 together with Corollary 1.3.25 of [11],

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_{k-d} X_k = \ell_d \quad \text{a.s.}$$

Finally, let also  $L \in \mathbb{R}^{p+2}$  and, for  $0 \leq d \leq p+1$ ,  $L_p^d \in \mathbb{R}^p$  be vectors of limiting values such that,

$$L = (\ell_0 \ \ell_1 \ \dots \ \ell_{p+1})' \quad \text{and} \quad L_p^d = (\ell_d \ \ell_{d-1} \ \dots \ \ell_{d-p+1})'.$$

From (B.8), an immediate development leads to

$$\sum_{k=1}^n X_{k-d} X_k = \beta' \sum_{k=1}^n \Phi_{k-1}^p X_{k-d} - \theta_p \rho \sum_{k=1}^n X_{k-p-1} X_{k-d} + \sum_{k=1}^n X_{k-d} V_k,$$

considering that  $X_{-1}, X_{-2}, \dots, X_{-p} = 0$ . Consequently, we obtain a set of relations between almost sure limits, for all  $0 \leq d \leq p+1$ ,

$$(B.6) \quad \ell_d = \beta' L_p^{d-1} - \theta_p \rho \ell_{d-p-1} + \sigma^2 \delta_d$$

where  $\ell_{-d} = \ell_d$ . Hereafter, if  $d$  varies from 0 to  $p+1$ , one can build a  $(p+2) \times (p+2)$  linear system of equations verifying

$$(B.7) \quad BL = \sigma^2 e$$

where  $B$  is precisely given by (2.7). We know from Lemma 2.1 that under the stability conditions, the matrix  $B$  is invertible. Therefore, it follows that

$$L = \sigma^2 B^{-1} e,$$

meaning via (2.8) that  $L = \sigma^2 \Lambda$ , or else, for all  $0 \leq d \leq p+1$ ,  $\ell_d = \sigma^2 \lambda_d$ , which completes the proof of Lemma B.3.  $\square$

## B.2. Proof of Theorem 2.1.

We easily deduce from (1.1) that the process  $(X_n)$  satisfies the fundamental autoregressive equation given, for all  $n \geq p+1$ , by

$$(B.8) \quad X_n = \beta' \Phi_{n-1}^p - \theta_p \rho X_{n-p-1} + V_n$$

where  $\beta$  is given by (2.5). On the basis of (B.8), consider the summation

$$(B.9) \quad \sum_{k=1}^n \Phi_{k-1}^p X_k = \sum_{k=1}^n \Phi_{k-1}^p \beta' \Phi_{k-1}^p - \theta_p \rho \sum_{k=1}^n \Phi_{k-1}^p X_{k-p-1} + \sum_{k=1}^n \Phi_{k-1}^p V_k.$$

First of all, an immediate calculation leads to

$$(B.10) \quad \sum_{k=1}^n \Phi_{k-1}^p \beta' \Phi_{k-1}^p = (S_{n-1} - S) \beta$$

where  $S_{n-1}$  and  $S$  are given in (2.2). Let us focus now on the more intricate term

$$\sum_{k=1}^n \Phi_{k-1}^p X_{k-p-1}$$

in which we shall expand each element of  $\Phi_{k-1}^p$  according to (B.8). A direct calculation infers the equality, for all  $n \geq p+1$ ,

$$(B.11) \quad \sum_{k=1}^n \Phi_{k-1}^p X_{k-p-1} = S_{n-1} J_p \beta - \theta_p \rho \sum_{k=1}^n \Phi_{k-1}^p X_k + J_p \sum_{k=1}^n \Phi_{k-1}^p V_k + \xi_n$$

where Lemma B.2 ensures that the remainder term  $\xi_n$  is made of isolated terms such that  $\|\xi_n\| = o(n)$  a.s. Let also  $M_n$  be the  $p$ -dimensional martingale

$$(B.12) \quad M_n = \sum_{k=1}^n \Phi_{k-1}^p V_k.$$

We deduce from (B.9) together with (B.10) and (B.11) that

$$\sum_{k=1}^n \Phi_{k-1}^p X_k = \alpha S_{n-1} (I_p - \theta_p \rho J_p) \beta + \alpha (I_p - \theta_p \rho J_p) M_n + \alpha \xi_n$$

where  $\alpha$  is given by (2.4). Thus, taking into account the expression of the estimator (2.3), we get the main decomposition, for all  $n \geq p+1$ ,

$$(B.13) \quad \hat{\theta}_n = \alpha (I_p - \theta_p \rho J_p) \beta + \alpha (S_{n-1})^{-1} (I_p - \theta_p \rho J_p) M_n + \alpha (S_{n-1})^{-1} \xi_n.$$

For all  $n \geq 1$ , denote by  $\mathcal{F}_n$  the  $\sigma$ -algebra of the events occurring up to stage  $n$ ,  $\mathcal{F}_n = \sigma(X_0, \dots, X_p, V_1, \dots, V_n)$ . The random sequence  $(M_n)$  given by (B.12) is a locally square-integrable real vector martingale [11], [20], adapted to  $\mathcal{F}_n$ , with predictable quadratic variation given, for all  $n \geq 1$ , by

$$(B.14) \quad \begin{aligned} \langle M \rangle_n &= \sum_{k=1}^n \mathbb{E}[(\Delta M_k)(\Delta M_k)' | \mathcal{F}_{k-1}], \\ &= \sigma^2 \sum_{k=1}^n \Phi_{k-1}^p \Phi_{k-1}^{p'} = \sigma^2 (S_{n-1} - S) \end{aligned}$$

where  $\Delta M_k$  stands for the difference  $M_k - M_{k-1}$ . We know from Lemma B.3 that

$$(B.15) \quad \lim_{n \rightarrow \infty} \frac{S_n}{n} = \sigma^2 \Delta_p \quad \text{a.s.}$$

and  $\Delta_p$  is positive definite as a result of Lemma 2.2. Then, (B.15) implies that

$$(B.16) \quad \lim_{n \rightarrow \infty} \frac{\text{tr}(S_n)}{n} = \sigma^2 p \lambda_0 \quad \text{a.s.}$$

where  $\lambda_0 > 0$ . Moreover, since  $\Delta_p$  is positive definite, we also have that

$$(B.17) \quad \lambda_{\max}(S_n) = O(\lambda_{\min}(S_n)) \quad \text{a.s.}$$

Consequently, we deduce from (B.14), (B.16), (B.17) and the strong law of large numbers for vector martingales given e.g. in Theorem 4.3.15 of [11], or [12] that,

$$(B.18) \quad \lim_{n \rightarrow \infty} \langle M \rangle_n^{-1} M_n = 0 \quad \text{a.s.}$$

and obviously,

$$(B.19) \quad \lim_{n \rightarrow \infty} (S_{n-1})^{-1} (I_p - \theta_p \rho J_p) M_n = 0 \quad \text{a.s.}$$

As mentioned above,  $(V_n)$  having a finite moment of order 2 implies, via Lemma B.2 and (B.15), that

$$(B.20) \quad \lim_{n \rightarrow \infty} (S_{n-1})^{-1} \xi_n = 0 \quad \text{a.s.}$$

Finally, (B.13) together with (B.19) and (B.20) achieve the proof of Theorem 2.1,

$$\lim_{n \rightarrow \infty} \hat{\theta}_n = \alpha(I_p - \theta_p \rho J_p) \beta \quad \text{a.s.}$$

□

### B.3. Proof of Theorem 2.2.

The main decomposition (B.13) enables us to write, for all  $n \geq p + 1$ ,

$$(B.21) \quad \sqrt{n} (\hat{\theta}_n - \theta^*) = \alpha \sqrt{n} (S_{n-1})^{-1} (I_p - \theta_p \rho J_p) M_n + \alpha \sqrt{n} (S_{n-1})^{-1} \xi_n.$$

On the one hand, we have from Lemma B.2 with  $a = 4$  that  $\|\xi_n\| = o(\sqrt{n})$  a.s. assuming the existence of a finite moment of order 4 for  $(V_n)$ . Hence, via (B.15),

$$(B.22) \quad \lim_{n \rightarrow \infty} \sqrt{n} (S_{n-1})^{-1} \xi_n = 0 \quad \text{a.s.}$$

On the other hand, we shall make use of the central limit theorem for vector martingales given e.g. by Corollary 2.1.10 of [11], to establish the asymptotic normality of the first term in the right-hand side of (B.21). Foremost, it is necessary to prove that the Lindeberg's condition is satisfied. We have to prove that, for all  $\varepsilon > 0$ ,

$$(B.23) \quad \frac{1}{n} \sum_{k=1}^n \mathbb{E} \left[ \|\Delta M_k\|^2 \mathbf{I}_{\{\|\Delta M_k\| \geq \varepsilon \sqrt{n}\}} | \mathcal{F}_{k-1} \right] \xrightarrow{\mathcal{P}} 0$$

where  $\Delta M_k = M_k - M_{k-1} = \Phi_{k-1}^p V_k$ . We have from Lemma B.2 with  $a = 4$  that

$$(B.24) \quad \sum_{k=1}^n \|\Phi_{k-1}^p\|^4 = O(n) \quad \text{a.s.}$$

Moreover, for all  $\varepsilon > 0$ ,

$$\begin{aligned} \frac{1}{n} \sum_{k=1}^n \mathbb{E} \left[ \|\Delta M_k\|^2 \mathbf{I}_{\{\|\Delta M_k\| \geq \varepsilon \sqrt{n}\}} | \mathcal{F}_{k-1} \right] &\leq \frac{1}{\varepsilon^2 n^2} \sum_{k=1}^n \mathbb{E} [\|\Delta M_k\|^4 | \mathcal{F}_{k-1}] , \\ &\leq \frac{\tau^4}{\varepsilon^2 n^2} \sum_{k=1}^n \|\Phi_{k-1}^p\|^4 \end{aligned}$$

where  $\tau^4$  stands for the moment of order 4 associated with  $(V_n)$ . Consequently, (B.24) ensures that

$$\frac{1}{n} \sum_{k=1}^n \mathbb{E} \left[ \|\Delta M_k\|^2 \mathbf{I}_{\{\|\Delta M_k\| \geq \varepsilon \sqrt{n}\}} | \mathcal{F}_{k-1} \right] = O(n^{-1}) \quad \text{a.s.}$$

and the Lindeberg's condition (B.23) is satisfied. We conclude from the central limit theorem for vector martingales together with Lemma 2.2 and Lemma B.3 that

$$(B.25) \quad \sqrt{n} \langle M \rangle_n^{-1} M_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^{-4} \Delta_p^{-1})$$

where  $\Delta_p$  is given by (2.9), which leads to

$$(B.26) \quad \alpha \sqrt{n} (S_{n-1})^{-1} (I_p - \theta_p \rho J_p) M_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_\theta).$$

Finally, (B.21), (B.22) and (B.26) complete the proof of Theorem 2.2.  $\square$

#### B.4. Proof of Theorem 2.3.

Let  $(W_n)$  be the sequence of standardization matrices defined as  $W_n = \sqrt{n} I_p$ . Consider the locally square-integrable real vector martingale  $(M_n)$  with predictable quadratic variation  $\langle M \rangle_n$  given by (B.14). Via Lemma B.3, we have the almost sure convergence

$$(B.27) \quad \lim_{n \rightarrow \infty} W_n^{-1} \langle M \rangle_n W_n^{-1} = \sigma^4 \Delta_p \quad \text{a.s.}$$

where  $\Delta_p$  is given by (2.9). For all  $n \geq 0$ , denote

$$(B.28) \quad T_n = \sum_{k=1}^n X_k^4$$

with  $T_0 = 0$ . From Lemma B.2 with  $a = 4$ , we have that  $T_n = O(n)$  a.s. Thus,

$$\begin{aligned} \sum_{n=1}^{+\infty} \frac{X_n^4}{n^2} &= \sum_{n=1}^{+\infty} \frac{T_n - T_{n-1}}{n^2} = \sum_{n=1}^{+\infty} \left( \frac{2n+1}{n^2(n+1)^2} \right) T_n, \\ &= O \left( \sum_{n=1}^{+\infty} \frac{T_n}{n^3} \right) = O \left( \sum_{n=1}^{+\infty} \frac{1}{n^2} \right) < +\infty \quad \text{a.s.} \end{aligned}$$



which immediately implies that

$$(B.29) \quad \sum_{n=1}^{+\infty} \frac{\|\Phi_{n-1}^p\|^4}{n^2} < +\infty \quad \text{a.s.}$$

From (B.27) and (B.29), we can deduce that  $(M_n)$  satisfies the quadratic strong law for vector martingales given e.g. by Theorem 2.1 of [10],

$$(B.30) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n^p} \sum_{k=1}^n \left[ 1 - \frac{k^p}{(k+1)^p} \right] W_k^{-1} M_k M_k' W_k^{-1} = \sigma^4 \Delta_p \quad \text{a.s.}$$

Hereafter, it follows from (B.13) that, for all  $n \geq p+1$ ,

$$(B.31) \quad \begin{aligned} \left( \widehat{\theta}_n - \theta^* \right) \left( \widehat{\theta}_n - \theta^* \right)' &= \alpha^2 (S_{n-1})^{-1} \left[ K_p M_n + \xi_n \right] \left[ M_n' K_p + \xi_n' \right] (S_{n-1})^{-1}, \\ &= \alpha^2 (S_{n-1})^{-1} K_p M_n M_n' K_p (S_{n-1})^{-1} + \zeta_n \end{aligned}$$

where  $K_p = (I_p - \theta_p \rho J_p)$  and the remainder term

$$\zeta_n = \alpha^2 (S_{n-1})^{-1} (\xi_n M_n' K_p + K_p M_n \xi_n' + \xi_n \xi_n') (S_{n-1})^{-1}.$$

However, we have from Lemma 2.2 and Lemma B.3 that

$$(B.32) \quad \lim_{n \rightarrow \infty} n (S_{n-1})^{-1} = \sigma^{-2} \Delta_p^{-1} \quad \text{a.s.}$$

As a result, (B.30), (B.32) and a set of additional steps of calculation lead to the almost sure convergence

$$(B.33) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n (S_{k-1})^{-1} K_p M_k M_k' K_p (S_{k-1})^{-1} = K_p \Delta_p^{-1} K_p \quad \text{a.s.}$$

since  $K_p \Delta_p^{-1} = \Delta_p^{-1} K_p$  due to the bisymmetry of  $\Delta_p^{-1}$ . Assuming a finite moment of order 4 for  $(V_n)$ , one can easily be convinced that  $\zeta_n$  is going to play a negligible role compared to the first one in the right-hand side of (B.31). Indeed, we clearly have that  $\|M_n\| \|\xi_n\| = o(n^{3/4} \sqrt{\log n})$  a.s. It follows that

$$(B.34) \quad \sum_{k=1}^n \zeta_k = O(1) \quad \text{a.s.}$$

Finally, (B.33) and (B.34) complete the proof of the first part of Theorem 2.3,

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \left( \widehat{\theta}_k - \theta^* \right) \left( \widehat{\theta}_k - \theta^* \right)' = \Sigma_\theta \quad \text{a.s.}$$

since  $\Sigma_\theta = \alpha^2 K_p \Delta_p^{-1} K_p$ .

The law of iterated logarithm (2.13) is much more easy to handle. It is based on the law of iterated logarithm for vector martingales given e.g. by Lemma C.2 in [1].

Under the assumption (B.29) already verified, for any vector  $v \in \mathbb{R}^p$ , we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{1/2} v'(S_{n-1})^{-1} M_n &= - \liminf_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{1/2} v'(S_{n-1})^{-1} M_n, \\ (B.35) \qquad \qquad \qquad &= (v' \Delta_p^{-1} v)^{1/2} \quad \text{a.s.} \end{aligned}$$

Via (B.35) and the negligibility of  $\zeta_n$ , we immediately get

$$\begin{aligned} \limsup_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{1/2} v' (\hat{\theta}_n - \theta^*) &= - \liminf_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{1/2} v' (\hat{\theta}_n - \theta^*), \\ (B.36) \qquad \qquad \qquad &= \alpha (v' K_p \Delta_p^{-1} K_p v)^{1/2} \quad \text{a.s.} \end{aligned}$$

Since (B.36) is true whatever the value of  $v \in \mathbb{R}^p$ , we obtain a matrix formulation of the law of iterated logarithm,

$$(B.37) \qquad \limsup_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right) (\hat{\theta}_n - \theta^*) (\hat{\theta}_n - \theta^*)' = \Sigma_\theta \quad \text{a.s.}$$

Passing through the trace in (B.37), we find that

$$(B.38) \qquad \limsup_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right) \|\hat{\theta}_n - \theta^*\|^2 = \text{tr}(\Sigma_\theta) \quad \text{a.s.}$$

which completes the proof of Theorem 2.3.  $\square$

## APPENDIX C

### PROOFS OF THE SERIAL CORRELATION PARAMETER RESULTS

#### C.1. Proof of Theorem 3.1.

Let us introduce some additional notations to make this technical proof more understandable. Recall that, for all  $d \in \{0, \dots, p+1\}$ , we have the almost sure convergence

$$(C.1) \qquad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_{k-d} X_k = \sigma^2 \lambda_d \quad \text{a.s.}$$

Let  $\Lambda_p^0$ ,  $\Lambda_p^1$  and  $\Lambda_p^2$  be a set of  $p$ -dimensional vectors of limiting values such that, for  $d = \{0, 1, 2\}$ ,

$$(C.2) \qquad \Lambda_p^d = (\lambda_d \quad \lambda_{d+1} \quad \dots \quad \lambda_{d+p-1})',$$

and note that the almost sure convergence follows,

$$(C.3) \qquad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \Phi_{k-d}^p X_k = \sigma^2 \Lambda_p^d \quad \text{a.s.}$$

For all  $n \geq 1$ , denote by  $A_n$  the square matrix of order  $p$  defined as

$$(C.4) \quad A_n = \sum_{k=1}^n \Phi_k^p \Phi_{k-1}^{p'}$$

Following a reasoning very similar to the proof of Theorem 2.1, it is possible to obtain the decomposition, for all  $n \geq p+1$ ,

$$(C.5) \quad \sum_{k=1}^n \Phi_k^p X_k = A_n \theta^* + \alpha \sum_{k=1}^n \Phi_k^p V_k - \theta_p \rho \alpha J_p \sum_{k=1}^n \Phi_{k-2}^p V_k + \eta_n$$

where the residual  $\eta_n$  is made of isolated terms such that  $\|\eta_n\| = o(n)$  a.s. As an immediate consequence, we have the relation between the limiting values

$$(C.6) \quad \Lambda_p^0 = A_p \theta^* + \alpha e$$

where the almost sure limiting matrix of  $\sigma^{-2} A_n/n$  is given by

$$(C.7) \quad A_p = \begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 & \dots & \dots & \lambda_p \\ \lambda_0 & \lambda_1 & \lambda_2 & \dots & \dots & \lambda_{p-1} \\ \vdots & \vdots & \vdots & & & \vdots \\ \vdots & \vdots & \vdots & & & \vdots \\ \lambda_{p-2} & \lambda_{p-3} & \lambda_{p-4} & \dots & \dots & \lambda_1 \end{pmatrix}.$$

The reader may find more details about the way to establish these almost sure convergences e.g. in the proof of Lemma B.3. Likewise, one proves that

$$(C.8) \quad \Lambda_p^2 = A_p' \theta^* - \alpha \theta_p \rho J_p e.$$

Finally, the very definition of the estimator  $\hat{\theta}_n$  directly implies another relation, involving the matrix  $\Delta_p$  given by (2.9),

$$(C.9) \quad \Lambda_p^1 = \Delta_p \theta^*.$$

Relations (C.6), (C.8) and (C.9) will be useful thereafter. Let us now consider the expression of  $\hat{\rho}_n$  given by (3.2). On the one hand, in light of foregoing,

$$(C.10) \quad \begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \hat{\varepsilon}_k \hat{\varepsilon}_{k-1} &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \left( X_k - \hat{\theta}_n' \Phi_{k-1}^p \right) \left( X_{k-1} - \hat{\theta}_n' \Phi_{k-2}^p \right), \\ &= \sigma^2 \left( \lambda_1 - \left( \Lambda_p^{0'} + \Lambda_p^{2'} \right) \theta^* + \theta^{*'} A_p \theta^* \right), \\ &= \sigma^2 \left( \lambda_1 - \Lambda_p^{2'} \theta^* - \alpha \theta_1^* \right) \quad \text{a.s.} \end{aligned}$$

On the other hand, similarly,

$$(C.11) \quad \begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \hat{\varepsilon}_{k-1}^2 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \left( X_{k-1} - \hat{\theta}_n' \Phi_{k-2}^p \right)^2, \\ &= \sigma^2 \left( \lambda_0 - 2 \Lambda_p^{1'} \theta^* + \theta^{*'} \Delta_p \theta^* \right), \\ &= \sigma^2 \left( \lambda_0 - \Lambda_p^{1'} \theta^* \right) \quad \text{a.s.} \end{aligned}$$

Via the set of relations (B.6), we find that  $\lambda_0 = \beta' \Lambda_p^1 - \theta_p \rho \lambda_{p+1} + 1$  for  $d = 0$ , and  $\lambda_{p+1} = \beta' J_p \Lambda_p^1 - \theta_p \rho \lambda_0$  for  $d = p + 1$ , in particular. Hence, with  $\theta^* = \alpha(I_p - \theta_p \rho J_p) \beta$ ,

$$\begin{aligned}
 \lambda_1 - \Lambda_p^{2'} \theta^* - \alpha \theta_1^* &= \lambda_1 - \Lambda_p^{2'} \theta^* - \alpha \theta_1^* (\lambda_0 - \beta' \Lambda_p^1 + \theta_p \rho \lambda_{p+1}), \\
 &= \lambda_1 - \Lambda_p^{2'} \theta^* - \alpha \theta_1^* (\lambda_0 - \beta' \Lambda_p^1 + \theta_p \rho (\beta' J_p \Lambda_p^1 - \theta_p \rho \lambda_0)), \\
 &= \lambda_1 - \Lambda_p^{2'} \theta^* - \theta_1^* (\lambda_0 - \Lambda_p^{1'} \theta^*), \\
 (C.12) \quad &= \lambda_1 - \Lambda_p^{2'} \theta^* - (\theta_1 + \rho) (\lambda_0 - \Lambda_p^{1'} \theta^*) + \theta_p \rho \theta_p^* (\lambda_0 - \Lambda_p^{1'} \theta^*)
 \end{aligned}$$

since one has to note that  $\theta_1^* = \theta_1 + \rho - \theta_p \rho \theta_p^*$ . Via (C.9),  $\lambda_1 = \Lambda_p^{0'} \theta^*$ . Thus,

$$\begin{aligned}
 \lambda_1 - \Lambda_p^{2'} \theta^* &= \theta^{*'} (\Lambda_p^0 - \Lambda_p^2), \\
 &= \theta^{*'} A_p' \theta^* - \theta^{*'} A_p \theta^* + \alpha (\theta_1 + \rho), \\
 &= \alpha (\theta_1 + \rho) (\lambda_0 - \beta' \Lambda_p^1 + \theta_p \rho \lambda_{p+1}), \\
 (C.13) \quad &= (\theta_1 + \rho) (\lambda_0 - \Lambda_p^{1'} \theta^*).
 \end{aligned}$$

To conclude, (C.12) together with (C.13) lead to

$$\lambda_1 - \Lambda_p^{2'} \theta^* - \alpha \theta_1^* = \theta_p \rho \theta_p^* (\lambda_0 - \Lambda_p^{1'} \theta^*)$$

which, via (C.10) and (C.11), achieves the proof of Theorem 3.1,

$$\lim_{n \rightarrow \infty} \hat{\rho}_n = \theta_p \rho \theta_p^* \quad \text{a.s.}$$

□

## C.2. Proof of Theorem 3.2.

First of all, we have already seen from (B.13) that, for all  $n \geq p + 1$ ,

$$(C.14) \quad S_{n-1} \left( \hat{\theta}_n - \theta^* \right) = \alpha (I_p - \theta_p \rho J_p) M_n + \alpha \xi_n$$

where Lemma B.2 involves  $\|\xi_n\| = o(\sqrt{n})$  a.s., assuming a finite moment of order 4 for  $(V_n)$ . Our goal is to find a similar decomposition for  $\hat{\rho}_n - \rho^*$ . For a better readability, let us introduce two specific backward and forward difference operators  $\mathcal{B}$  and  $\mathcal{F}$ , such that  $\mathcal{B}(X_n)$  and  $\mathcal{F}(X_n)$  operate on  $X_n$  to produce the following linear combinations,

$$\mathcal{B}(X_n) = X_n - \rho^* X_{n-1} \quad \text{and} \quad \mathcal{F}(X_n) = X_{n-1} - \rho^* X_n.$$

The extension of  $\mathcal{B}$  and  $\mathcal{F}$  to the vectorial framework merely consists in operating component by component. Denote by  $F_n$  the recurrent  $p$ -dimensional expression that appears repeatedly in the decomposition, given, for all  $n \geq 1$ , by

$$(C.15) \quad F_n = \Phi_n^p \theta^{*'} \mathcal{F}(\Phi_n^p) - (\mathcal{F}(\Phi_{n-1}^p) + \mathcal{B}(\Phi_n^p)) X_n.$$

From the residual estimation (3.1), the development of  $\hat{\rho}_n - \rho^*$  reduces to

$$(C.16) \quad J_{n-1} \left( \hat{\rho}_n - \rho^* \right) = W_n + \left( \hat{\theta}_n - \theta^* \right)' H_n$$

where  $H_n$  is a  $p$ -dimensional vector and, for all  $n \geq p + 1$ ,

$$(C.17) \quad J_n = \sum_{k=0}^n \widehat{\varepsilon}_k^2,$$

$$(C.18) \quad W_n = \sum_{k=1}^n \mathcal{F}(X_k) X_k + \theta^{*'} \sum_{k=1}^n F_k + \nu_n,$$

$$(C.19) \quad H_n = \sum_{k=1}^n (\mathcal{F}(\Phi_k^p) \theta^{*'} \Phi_k^p + F_k) + \sum_{k=1}^n \Phi_k^p (\widehat{\theta}_n - \theta^*)' \mathcal{F}(\Phi_k^p) + \mu_n,$$

with  $\|\mu_n\| = o(\sqrt{n})$  a.s. and  $\nu_n = o(\sqrt{n})$  a.s. The reasoning develops in two stages. At first, we shall prove that  $W_n$  reduces to a martingale, except for a residual term. Then, using Theorem 2.2 and the central limit theorem for vector martingales, we will be in the position to prove the joint asymptotic normality of our estimates.

Let  $C$  be the square submatrix of order  $p + 1$  obtained by removing from  $B$  given by (2.7) its first row and first column,

$$(C.20) \quad C = \begin{pmatrix} 1 - \beta_2 & -\beta_3 & \dots & \dots & -\beta_p & \theta_p \rho & 0 \\ -\beta_1 - \beta_3 & 1 - \beta_4 & \dots & \dots & \theta_p \rho & 0 & 0 \\ \vdots & \vdots & & & \vdots & \vdots & \vdots \\ \vdots & \vdots & & & \vdots & \vdots & \vdots \\ -\beta_{p-1} + \theta_p \rho & -\beta_{p-2} & \dots & \dots & -\beta_1 & 1 & 0 \\ -\beta_p & -\beta_{p-1} & \dots & \dots & -\beta_2 & -\beta_1 & 1 \end{pmatrix}.$$

By Corollary 2.1, we have already seen that the matrix  $C$  is invertible under the stability conditions. Denote by  $N_n$  be the  $(p + 1)$ -dimensional martingale

$$(C.21) \quad N_n = \sum_{k=1}^n \Phi_{k-1}^{p+1} V_k$$

where  $\Phi_n^{p+1}$  stands for the extension of  $\Phi_n^p$  to the next dimension. A straightforward calculation based on (B.8) shows that the following linear system is satisfied,

$$C \sum_{k=1}^n \Phi_{k-1}^{p+1} X_k = T \sum_{k=1}^n X_k^2 + N_n$$

in which  $T$  is defined as

$$(C.22) \quad T = (\beta_1 \quad \beta_2 \quad \dots \quad \beta_p \quad -\theta_p \rho)'.$$

As a result of the invertibility of  $C$ , we get the substantial equality, for all  $n \geq p + 1$ ,

$$(C.23) \quad \sum_{k=1}^n \Phi_{k-1}^{p+1} X_k = C^{-1} T \sum_{k=1}^n X_k^2 + C^{-1} N_n.$$

A large manipulation of  $W_n$  given in (C.18) still based on the fundamental autoregressive form (B.8) shows, after further calculations, that there exists an isolated

term  $\nu_n$  such that  $\nu_n = o(\sqrt{n})$  a.s., and, for all  $n \geq p + 1$ ,

$$\begin{aligned} W_n = & \sum_{k=1}^n \mathcal{F}(X_k)X_k - \theta^{*'} \sum_{k=1}^n \mathcal{F}(\Phi_{k-1}^p)X_k - \alpha\theta^{*'} \sum_{k=1}^n (\Phi_k^p - \theta_p \rho J_p \Phi_{k-2}^p) V_k \\ & + \alpha\rho^* \theta^{*'} (I_p - \theta_p \rho J_p) \sum_{k=1}^n \Phi_{k-1}^p V_k + \nu_n, \end{aligned}$$

leading, together with (C.23), to

$$(C.24) \quad W_n = (G'C^{-1}T - \rho^* - \alpha\theta_1^*) \sum_{k=1}^n X_k^2 + G'C^{-1}N_n + L_n + \nu_n$$

where, for all  $n \geq p + 1$ ,

$$(C.25) \quad L_n = \alpha\theta^{*'} \left( \rho^* (I_p - \theta_p \rho J_p) M_n - \sum_{k=1}^n (\Phi_k^p - \theta_p \rho J_p \Phi_{k-2}^p) V_k \right) + \alpha\theta_1^* \sum_{k=1}^n X_k V_k,$$

and where the  $(p + 1)$ -dimensional vector  $G$  is given by

$$(C.26) \quad G = \rho^* \vartheta^* + \alpha\theta_1^* T - \delta^*$$

with  $\vartheta^* = (\theta_1^* \ \theta_2^* \ \dots \ \theta_p^* \ 0)'$  and  $\delta^* = (-1 \ \theta_1^* \ \dots \ \theta_{p-1}^* \ \theta_p^*)'$ . In terms of almost sure limits, by using the same methodology as e.g. in the proof of Lemma B.3, (C.23) directly implies

$$(C.27) \quad \lambda_0 C^{-1}T = \Lambda_{p+1}^1$$

where  $\Lambda_{p+1}^1 = (\lambda_1 \ \lambda_2 \ \dots \ \lambda_{p+1})'$  is the extension of  $\Lambda_p^1$  in (C.2) to the next dimension. Hence, following the same lines as in the proof of Theorem 3.1,

$$\begin{aligned} \lambda_0 (G'C^{-1}T - \rho^* - \alpha\theta_1^*) &= G'\Lambda_{p+1}^1 - \lambda_0 (\rho^* + \alpha\theta_1^*), \\ &= \rho^* (\Lambda_p^{1'} \theta^* - \lambda_0) + \alpha\theta_1^* (T' \Lambda_{p+1}^1 - \lambda_0) + (\lambda_1 - \Lambda_p^{2'} \theta^*), \\ &= \theta_1^* (\alpha\Lambda_p^{1'} (I_p - \theta_p \rho J_p) \beta - \alpha(1 - \theta_p \rho)(1 + \theta_p \rho) \lambda_0) \\ &\quad + \rho^* (\Lambda_p^{1'} \theta^* - \lambda_0) + (\lambda_1 - \Lambda_p^{2'} \theta^*), \\ &= \theta_1^* (\Lambda_p^{1'} \theta^* - \lambda_0) + \rho^* (\Lambda_p^{1'} \theta^* - \lambda_0) + (\lambda_1 - \Lambda_p^{2'} \theta^*), \\ &= -\alpha(\rho^* + \theta_1^*) + \alpha(\rho^* + \theta_1^*) = 0. \end{aligned}$$

One can see from Lemma 2.2 that  $\lambda_0 > 0$ . The latter development ensures that the pathological term of (C.24) vanishes, as it should. Finally,  $W_n$  reduces to

$$(C.28) \quad W_n = G'C^{-1}N_n + L_n + \nu_n,$$

and one shall observe that  $G'C^{-1}N_n + L_n$  is a locally square-integrable real martingale [11], [20]. One is now able to combine (C.14) and (C.16), via (C.28), to establish the decomposition, for all  $n \geq p + 1$ ,

$$(C.29) \quad J_{n-1}(\hat{\rho}_n - \rho^*) = G'C^{-1}N_n + L_n + \alpha M'_n (I_p - \theta_p \rho J_p) (S_{n-1})^{-1} H_n + r_n$$

where the remainder term  $r_n = \alpha \xi'_n (S_{n-1})^{-1} H_n + \nu_n$  is such that  $r_n = o(\sqrt{n})$  a.s. Taking tediously advantage of the  $(p+2) \times (p+2)$  linear system of equations (B.6), one shall observe that  $G' C^{-1} = \alpha \begin{pmatrix} U'_p & u_{p+1} \end{pmatrix}$  with

$$U_p = \begin{pmatrix} 1 + \beta_2 & \beta_3 - \beta_1 & \dots & \beta_p - \beta_{p-2} & -\beta_{p-1} - \theta_p \rho \end{pmatrix}',$$

and  $u_{p+1} = -\alpha^{-1} \theta_p^* - \theta_p \rho \theta_1^*$ . The combination of (C.25) and (C.28) results in

$$(C.30) \quad W_n = \alpha (U_p + (I_p - \theta_p \rho J_p)(\rho^* \theta^* - \tau^*))' M_n - \theta_p^* \sum_{k=1}^n X_{k-p-1} V_k + \nu_n$$

where  $\tau^* = (\theta_2^* \ \theta_3^* \ \dots \ \theta_p^* \ 0)'$ . Consequently, it follows from (C.14) together with (C.29) and (C.30) that

$$(C.31) \quad \sqrt{n} \begin{pmatrix} \hat{\theta}_n - \theta^* \\ \hat{\rho}_n - \rho^* \end{pmatrix} = \frac{1}{\sqrt{n}} P_n N_n + R_n$$

where the square matrix  $P_n$  of order  $p+1$  is given by

$$(C.32) \quad P_n = \begin{pmatrix} P_n^{(1,1)} & 0 \\ P_n^{(2,1)} & P_n^{(2,2)} \end{pmatrix}$$

with

$$\begin{aligned} P_n^{(1,1)} &= n(S_{n-1})^{-1} \alpha (I_p - \theta_p \rho J_p), \\ P_n^{(2,1)} &= n(J_{n-1})^{-1} \left( \alpha (U_p + (I_p - \theta_p \rho J_p)(\rho^* \theta^* - \tau^*))' + \alpha H'_n (S_{n-1})^{-1} (I_p - \theta_p \rho J_p) \right), \\ P_n^{(2,2)} &= -n(J_{n-1})^{-1} \theta_p^*, \end{aligned}$$

and where the  $(p+1)$ -dimensional remainder term

$$(C.33) \quad R_n = \sqrt{n} \begin{pmatrix} \alpha (S_{n-1})^{-1} \xi_n \\ (J_{n-1})^{-1} r_n \end{pmatrix}$$

is such that  $\|R_n\| = o(1)$  a.s. Via some simplifications on  $H_n$ , (C.6), (C.8) and (C.9), we obtain that

$$(C.34) \quad \lim_{n \rightarrow \infty} \frac{H_n}{n} = -\alpha (I_p - \theta_p \rho J_p) e \quad \text{a.s.}$$

Furthermore, it is not hard to see, via Lemma B.3, (C.11), (C.34) and some simplifications on  $P_n^{(2,1)}$ , that

$$(C.35) \quad \lim_{n \rightarrow \infty} P_n = \sigma^{-2} P \quad \text{a.s.}$$

where  $P$  is the limiting matrix precisely given by (3.4). The locally square-integrable real vector martingale  $(N_n)$  introduced in (C.21) and adapted to  $\mathcal{F}_n$  has a predictable quadratic variation  $\langle N \rangle_n$  such that

$$(C.36) \quad \lim_{n \rightarrow \infty} \frac{\langle N \rangle_n}{n} = \sigma^4 \Delta_{p+1} \quad \text{a.s.}$$

where  $\Delta_{p+1}$  is given by (3.5). This convergence can be achieved following e.g. the same lines as in the proof of Lemma B.3. On top of that, we also immediately

deduce from (B.24) that  $(N_n)$  satisfies the Lindeberg's condition. We conclude from the central limit theorem for martingales, given e.g. in Corollary 2.1.10 of [11], that

$$(C.37) \quad \frac{1}{\sqrt{n}} N_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^4 \Delta_{p+1}).$$

Whence, from (C.31), (C.33), (C.35), (C.37) and Slutsky's lemma,

$$(C.38) \quad \sqrt{n} \begin{pmatrix} \hat{\theta}_n - \theta^* \\ \hat{\rho}_n - \rho^* \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, P \Delta_{p+1} P').$$

This concludes the proof of Theorem 3.2 where, for readability purposes, we omitted most of the calculations which the attentive reader might easily deduce.  $\square$

### C.3. Proof of Theorem 3.3.

In the proof of Theorem 3.2, we have established a particular relation that we shall develop from now on, to achieve the proof of Theorem 3.3. Indeed, from (C.31), for all  $n \geq p+1$ ,

$$(C.39) \quad \hat{\rho}_n - \rho^* = n^{-1} \pi'_n N_n + (J_{n-1})^{-1} r_n$$

where  $N_n$  and  $J_{n-1}$  are given by (C.21) and (C.17), respectively, where  $r_n$  is such that  $r_n = o(\sqrt{n})$  a.s. and where  $\pi_n$  of order  $p+1$  is given from (C.32) by

$$(C.40) \quad \pi_n = \begin{pmatrix} P_n^{(2,1)} & P_n^{(2,2)} \end{pmatrix}'.$$

Denote by  $\pi$  the almost sure limit of  $\pi_n$ , accordingly given by

$$(C.41) \quad \pi = \sigma^{-2} \begin{pmatrix} P'_L & \varphi \end{pmatrix}'$$

where  $P_L$  and  $\varphi$  are defined in (3.4). Hence, (C.39) can be rewritten as

$$(C.42) \quad \hat{\rho}_n - \rho^* = n^{-1} \pi'_n N_n + n^{-1} (\pi_n - \pi)' N_n + (J_{n-1})^{-1} r_n.$$

One can note that  $(\pi'_n N_n)$  is a locally square-integrable real martingale with predictable quadratic variation given, for all  $n \geq 1$ , by

$$(C.43) \quad \langle \pi'_n N \rangle_n = \sigma^2 \pi' (T_{n-1} - T) \pi$$

where the square matrix  $T_n$  of order  $p+1$  is the extension of  $S_n$  given by (2.2) to the next dimension defined, for all  $n \geq 1$ , as

$$(C.44) \quad T_n = \sum_{k=1}^n \Phi_k^{p+1} \Phi_k^{p+1'} + T,$$

and  $T$  is a symmetric positive definite matrix. In addition,  $(\pi'_n N_n)$  satisfies a non-explosion condition summarized by

$$\lim_{n \rightarrow \infty} \frac{\pi'_n \Phi_n^{p+1} \Phi_n^{p+1'} \pi}{\pi'_n T_n \pi} = 0 \quad \text{a.s.}$$



by application of Lemma B.2 with  $a = 4$ . By virtue of the quadratic strong law for martingales given e.g. by Theorem 3 of [2] or [3],

$$(C.45) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \left( \frac{\pi' N_k}{\pi' T_{k-1} \pi} \right)^2 = \frac{1}{\pi' \Delta_{p+1} \pi} \quad \text{a.s.}$$

where  $\Delta_{p+1}$  given by (3.5) is the almost sure limit of  $\sigma^{-2} T_n / n$ . We refer the reader to Lemma B.3 to have more details on the latter remark. Note that  $\pi' \Delta_{p+1} \pi > 0$  since  $\Delta_{p+1}$  is a positive definite matrix, as a result of Lemma 2.2. The same goes for  $\pi' T_n \pi$ , for all  $n \geq 1$ , assuming a suitable choice of  $T$ . Besides, the almost sure convergence of  $\pi_n$  to  $\pi$ , the finite moment of order 4 for  $(V_n)$  together with (C.45) ensure that

$$(C.46) \quad \begin{aligned} \sum_{k=1}^n \left( \frac{(\pi_k - \pi)' N_k}{k} + \frac{r_k}{J_{k-1}} \right)^2 &= O \left( \sum_{k=1}^n \frac{((\pi_k - \pi)' N_k)^2}{k^2} + \sum_{k=1}^n \frac{r_k^2}{J_{k-1}^2} \right), \\ &= O(1) + o \left( \sum_{k=1}^n \frac{(\pi' N_k)^2}{k^2} \right), \\ &= o(\log n) \quad \text{a.s.} \end{aligned}$$

since  $r_n$  is made of isolated terms of order 2 and  $J_n = O(n)$  a.s. It follows that

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \left( \hat{\rho}_k - \rho^* \right)^2 &= \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \left( \frac{\pi' N_k}{k} \right)^2, \\ &= \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \left( \frac{\pi' N_k}{\pi' T_{k-1} \pi} \right)^2 \left( \frac{\pi' T_{k-1} \pi}{k} \right)^2, \\ &= \frac{\sigma^4 (\pi' \Delta_{p+1} \pi)^2}{\pi' \Delta_{p+1} \pi} = \sigma^4 \pi' \Delta_{p+1} \pi \quad \text{a.s.} \end{aligned}$$

via (C.45) and (C.46), since the cross-term also plays a negligible role compared to the leading one. The definition of  $\pi$  in (C.41) combined with the one of  $\Gamma$  in (3.6) achieves the proof of the first part of Theorem 3.3.

Furthermore, it follows from the law of iterated logarithm for martingales [31], [32], see also Corollary 6.4.25 of [11], that

$$\begin{aligned} \limsup_{n \rightarrow \infty} \left( \frac{\langle \pi' N \rangle_n}{2 \log \log \langle \pi' N \rangle_n} \right)^{1/2} \frac{\pi' N_n}{\langle \pi' N \rangle_n} &= - \liminf_{n \rightarrow \infty} \left( \frac{\langle \pi' N \rangle_n}{2 \log \log \langle \pi' N \rangle_n} \right)^{1/2} \frac{\pi' N_n}{\langle \pi' N \rangle_n}, \\ &= 1 \quad \text{a.s.} \end{aligned}$$

since we have via (B.29) that

$$(C.47) \quad \sum_{k=1}^{+\infty} \frac{(\pi' \Phi_{k-1}^p)^4}{k^2} < +\infty \quad \text{a.s.}$$

Recall that we have the almost sure convergence,

$$(C.48) \quad \lim_{n \rightarrow \infty} \frac{\langle \pi' N \rangle_n}{n} = \sigma^4 \pi' \Delta_{p+1} \pi \quad \text{a.s.}$$

Therefore, we immediately obtain that

$$\begin{aligned}
 \limsup_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{1/2} \frac{\pi' N_n}{\langle \pi' N \rangle_n} &= - \liminf_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{1/2} \frac{\pi' N_n}{\langle \pi' N \rangle_n}, \\
 \text{(C.49)} \qquad \qquad \qquad &= \sigma^{-2} (\pi' \Delta_{p+1} \pi)^{-1/2} \quad \text{a.s.}
 \end{aligned}$$

As in the previous proof and by virtue of the same arguments, one can easily be convinced that the remainder term in the right-hand side of (C.42) is negligible. It follows from (C.42) together with (C.48) and (C.49) that,

$$\begin{aligned}
 \limsup_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{1/2} (\hat{\rho}_n - \rho^*) &= - \liminf_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{1/2} (\hat{\rho}_n - \rho^*), \\
 &= \sigma^2 (\pi' \Delta_{p+1} \pi)^{1/2} \quad \text{a.s.}
 \end{aligned}$$

which achieves the proof of Theorem 3.3.  $\square$

**Acknowledgments.** *The author thanks Bernard Bercu for all his advices and suggestions during the preparation of this work.*

## REFERENCES

- [1] BERCU, B. Central limit theorem and law of iterated logarithm for least squares algorithms in adaptive tracking. *SIAM J. Control. Optim.* **36** (1998), 910–928.
- [2] BERCU, B. On the convergence of moments in the almost sure central limit theorem for martingales with statistical applications. *Stochastic Process. Appl.* **11** (2004), 157–173.
- [3] BERCU, B., CÉNAC, P., AND FAYOLLE, G. On the almost sure central limit theorem for vector martingales: convergence of moments and statistical applications. *J. Appl. Probab.* **46** (2009), 151–169.
- [4] BERCU, B., AND PROÏA, F. A sharp analysis on the asymptotic behavior of the Durbin-Watson statistic for the first-order autoregressive process. *ESAIM Probab. Stat.* **16** (2012).
- [5] BITSEKI PENDA, V., DJELLOUT, H., AND PROÏA, F. Moderate deviations for the Durbin-Watson statistic related to the first-order autoregressive process. *arXiv 1201.3579. Submitted for publication.* (2012).
- [6] BOX, G., AND LJUNG, G. On a measure of a lack of fit in time series models. *Biometrika*. **65-2** (1978), 297–303.
- [7] BOX, G., AND PIERCE, D. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Am. Stat. Assn. Jour.* **65** (1970), 1509–1526.
- [8] BREUSCH, T. Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*. **17-31** (1978), 334–355.
- [9] BROCKWELL, P., AND DAVIS, R. *Time Series: Theory and Methods*. Springer-Verlag, New-York, 1991.
- [10] CHAABANE, F., AND MAAOUIA, F. Théorèmes limites avec poids pour les martingales vectorielles. *ESAIM Probab. Stat.* **4** (2000), 137–189.
- [11] DUFLO, M. *Random iterative models*, vol. 34 of *Applications of Mathematics*, New York. Springer-Verlag, Berlin, 1997.
- [12] DUFLO, M., SENOUSI, R., AND TOUATI, A. Sur la loi des grands nombres pour les martingales vectorielles et l'estimateur des moindres carrés d'un modèle de régression. *Ann. Inst. Henri Poincaré*. **26** (1990), 549–566.
- [13] DURBIN, J. Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables. *Econometrica* **38** (1970), 410–421.

- [14] DURBIN, J. Approximate distributions of Student's  $t$ -statistics for autoregressive coefficients calculated from regression residuals. *J. Appl. Probab. Special Vol. 23A* (1986), 173–185.
- [15] DURBIN, J., AND WATSON, G. S. Testing for serial correlation in least squares regression. I. *Biometrika* 37 (1950), 409–428.
- [16] DURBIN, J., AND WATSON, G. S. Testing for serial correlation in least squares regression. II. *Biometrika* 38 (1951), 159–178.
- [17] DURBIN, J., AND WATSON, G. S. Testing for serial correlation in least squares regression. III. *Biometrika* 58 (1971), 1–19.
- [18] GODFREY, L. Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica*. 46 (1978), 1293–1302.
- [19] GRENDER, U., AND SZEGÖ, G. *Toeplitz forms and their applications*. California Monographs in Mathematical Sciences. University of California Press, Berkeley, 1958.
- [20] HALL, P., AND HEYDE, C. C. *Martingale limit theory and its application*. Probability and Mathematical Statistics. Academic Press Inc., New York, 1980.
- [21] HORN, R., AND JOHNSON, C. *Matrix Analysis*. Cambridge University Press, Cambridge, New-York, 1985.
- [22] INDER, B. A. Finite-sample power of tests for autocorrelation in models containing lagged dependent variables. *Economics Letters* 14 (1984), 179–185.
- [23] INDER, B. A. An approximation to the null distribution of the durbin-watson statistic in models containing lagged dependent variables. *Econometric Theory* 2 (1986), 413–428.
- [24] KING, M. L., AND WU, P. X. Small-disturbance asymptotics and the Durbin-Watson and related tests in the dynamic regression model. *J. Econometrics* 47 (1991), 145–152.
- [25] LAI, T. L., AND WEI, C. Asymptotic properties of general autoregressive models and strong consistency of least-squares estimates of their parameters. *Jour. Multi. Analysis*. 13 (1983), 1–23.
- [26] MADDALA, G. S., AND RAO, A. S. Tests for serial correlation in regression models with lagged dependent variables and serially correlated errors. *Econometrica* 41 (1973), 761–774.
- [27] MALINVAUD, E. Estimation et prévision dans les modèles économiques autorégressifs. *Review of the International Institute of Statistics* 29 (1961), 1–32.
- [28] NERLOVE, M., AND WALLIS, K. F. Use of the Durbin-Watson statistic in inappropriate situations. *Econometrica* 34 (1966), 235–238.
- [29] PARK, S. B. On the small-sample power of Durbin's  $h$  test. *Journal of the American Statistical Association* 70 (1975), 60–63.
- [30] STOCKER, T. On the asymptotic bias of OLS in dynamic regression models with autocorrelated errors. *Statist. Papers* 48 (2007), 81–93.
- [31] STOUT, W. F. A martingale analogue of Kolmogorov's law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 15 (1970), 279–290.
- [32] STOUT, W. F. *Almost sure convergence*, vol. 24 of *Probability and Mathematical Statistics*. Academic Press, New York-London, 1974.
- [33] TILLMAN, J. A. The power of the Durbin-Watson test. *Econometrica* 43 (1975), 959–974.

*E-mail address:* Frederic.Proia@inria.fr

UNIVERSITÉ BORDEAUX 1, INSTITUT DE MATHÉMATIQUES DE BORDEAUX, UMR 5251, AND INRIA BORDEAUX, TEAM ALEA, 351 COURS DE LA LIBÉRATION, 33405 TALENCE CEDEX, FRANCE.