A SHARP ANALYSIS ON THE ASYMPTOTIC BEHAVIOR OF THE DURBIN–WATSON STATISTIC FOR THE FIRST-ORDER AUTOREGRESSIVE PROCESS

Bernard Bercu[1] and Frédéric Proïa[1]

Abstract. The purpose of this paper is to provide a sharp analysis on the asymptotic behavior of the Durbin–Watson statistic. We focus our attention on the first-order autoregressive process where the driven noise is also 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 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. It allows us to establish the almost sure convergence and the asymptotic normality for the Durbin–Watson statistic. Finally, we propose a new bilateral statistical test for residual autocorrelation. We show how our statistical test procedure performs better, from a theoretical and a practical point of view, than the commonly used Box–Pierce and Ljung–Box procedures, even on small-sized samples.

1. INTRODUCTION

The Durbin–Watson statistic is very well-known in Econometry and Statistics. It was introduced by the pioneer works of Durbin and Watson [10–12], in order to test the serial independence of the driven noise of a linear regression model. The statistical test based on the Durbin–Watson statistic works pretty well for linear regression models, and its power was investigated by Tillman [26]. However, as it was observed by Malinvaud [19] and Nerlove and Wallis [20], its widespread use in inappropriate situations may lead to inadequate conclusions. More precisely, for linear regression models containing lagged dependent random variables, the Durbin–Watson statistic may be asymptotically biased. In order to prevent this misuse, Durbin [8] proposed alternative tests based on the redesign of the original one. Then, he explained how to use them in the particular case of the first-order autoregressive process previously investigated in [19,20]. Maddala and Rao [18] and Park [21] showed by simulations that alternative tests significantly outperform the inappropriate one even on small-sized samples. Inder [15, 16] and Durbin [9] went even deeper in the approximation of the critical values and distributions of the alternative tests under the null hypothesis. Afterwards, additional improvements were brought by King and

[1] Université Bordeaux 1, Institut de Mathématiques de Bordeaux, UMR 5251, and INRIA Bordeaux, team ALEA, 351 Cours de la Libération, 33405 Talence Cedex, France. `Bernard.Bercu@math.u-bordeaux1.fr`; `Frederic.Proia@inria.fr`

Wu [17] and more recently, Stocker [23] gave substantial contributions to the study of the asymptotic bias in the Durbin–Watson statistic resulting from the presence of lagged dependent random variables. Nevertheless, one can observe that statisticians usually tend to evaluate the presence of a significative serial correlation through the widely applied Ljung–Box [4] and Box–Pierce [5] statistical tests. This approach does not take into account the interaction between the dynamics of the model and the dynamic nature of the autoregressive residuals, and thereby overestimates the null hypothesis of white noise. This is an additional motivation for bringing the Durbin–Watson statistic back into light, insofar as it better fits to the residual set structure.

Our purpose is to investigate several open questions left unanswered during four decades on the Durbin–Watson statistic [8, 9, 20]. We shall focus our attention on the first-order autoregressive process given, for all $n \geq 1$, by

$$\begin{cases} X_n = \theta X_{n-1} + \varepsilon_n \\ \varepsilon_n = \rho \varepsilon_{n-1} + V_n \end{cases} \tag{1.1}$$

where the unknown parameters $|\theta| < 1$, $|\rho| < 1$. *Via* an extensive use of the theory of martingales [7, 14], we shall provide a sharp and rigorous analysis on the asymptotic behavior of the least squares estimators of θ and ρ. The previous results of convergence were only established in probability [19, 20]. We shall prove the almost sure convergence as well as the asymptotic normality of the least squares estimators of θ and ρ. 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 bilateral test for residual autocorrelation under the null hypothesis as well as under the alternative hypothesis. Indeed, we will explain why, on the basis of the empirical power, our test procedure performs better than Ljung–Box [4] and Box–Pierce [5] *portmanteau* tests, and why it is asymptotically equivalent to the most powerful alternative *h-test* of Durbin [8], in the case of a first-order autoregressive process.

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 estimator $\widehat{\theta}_n$ to the limiting value

$$\theta^* = \frac{\theta + \rho}{1 + \theta\rho}. \tag{1.2}$$

One can observe that $\theta^* = \theta$ if and only if $\rho = 0$. The asymptotic normality of $\widehat{\theta}_n$ as well as the quadratic strong law and the law of iterated logarithm are also provided. Section 3 deals with the estimation of the serial correlation parameter. We prove the almost sure convergence of the least squares estimator $\widehat{\rho}_n$ to

$$\rho^* = \theta\rho\theta^* = \frac{\theta\rho(\theta + \rho)}{1 + \theta\rho}. \tag{1.3}$$

As before, the asymptotic normality of $\widehat{\rho}_n$, the quadratic strong law and the law of iterated logarithm are also provided. It enables us to establish in Section 4 the almost sure convergence of the Durbin–Watson statistic $\widehat{D}_n$ to

$$D^* = 2(1 - \rho^*) \tag{1.4}$$

together with its asymptotic normality. Our sharp analysis on the asymptotic behavior of $\widehat{D}_n$ is true whatever the values of the parameters θ and ρ inside the interval $]-1, 1[$. Consequently, we are able in Section 4 to propose a new bilateral statistical test for residual autocorrelation. In Section 5, we compare the empirical power of the test procedure based on the Durbin–Watson statistic with commonly used existing tests, 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. A short conclusion is given in Section 6. All the technical proofs of Sections 2–4 are postponed in Appendices A–C, respectively.

2. On the autoregressive parameter

Consider the first-order autoregressive process given by (1.1) where the initial values X_0 and ε_0 may be arbitrarily chosen. 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 $\sigma^2 > 0$. In order to estimate the unknown parameter θ, it is natural to make use of the least squares estimator $\widehat{\theta}_n$ which minimizes

$$\Delta_n(\theta) = \sum_{k=1}^{n} \left(X_k - \theta X_{k-1} \right)^2.$$

Consequently, we obviously have for all $n \geq 1$,

$$\widehat{\theta}_n = \frac{\sum_{k=1}^{n} X_k X_{k-1}}{\sum_{k=1}^{n} X_{k-1}^2}. \tag{2.1}$$

Our first result concerns the almost sure convergence of $\widehat{\theta}_n$ to the limiting value θ^* given by (1.2). One can observe that the convergence in probability of $\widehat{\theta}_n$ to θ^* was already proven in [19, 20]. We improve this previous result by establishing the almost sure convergence of $\widehat{\theta}_n$ to θ^*.

Theorem 2.1. *We have the almost sure convergence*

$$\lim_{n \to \infty} \widehat{\theta}_n = \theta^* \qquad \text{a.s.} \tag{2.2}$$

Our second result deals with the asymptotic normality of $\widehat{\theta}_n$ where we denote

$$\sigma_\theta^2 = \frac{(1 - \theta^2)(1 - \theta\rho)(1 - \rho^2)}{(1 + \theta\rho)^3}. \tag{2.3}$$

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

$$\sqrt{n} \left(\widehat{\theta}_n - \theta^* \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\theta^2). \tag{2.4}$$

Remark 2.3. In the well-known case where the residuals are not correlated, which means that $\rho = 0$, we clearly have $\theta^* = \theta$, $\sigma_\theta^2 = 1 - \theta^2$ and we find again the asymptotic normality

$$\sqrt{n} \left(\widehat{\theta}_n - \theta \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1 - \theta^2).$$

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

$$\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \left(\widehat{\theta}_k - \theta^* \right)^2 = \sigma_\theta^2 \qquad \text{a.s.} \tag{2.5}$$

where σ_θ^2 is given by (2.3). In addition, we also have the law of iterated logarithm

$$\limsup_{n \to \infty} \left(\frac{n}{2 \log \log n} \right)^{1/2} \left(\widehat{\theta}_n - \theta^* \right) = - \liminf_{n \to \infty} \left(\frac{n}{2 \log \log n} \right)^{1/2} \left(\widehat{\theta}_n - \theta^* \right)$$

$$= \sigma_\theta \qquad \text{a.s.} \tag{2.6}$$

Consequently,

$$\limsup_{n \to \infty} \left(\frac{n}{2 \log \log n} \right) \left(\widehat{\theta}_n - \theta^* \right)^2 = \sigma_\theta^2 \qquad \text{a.s.} \tag{2.7}$$

Proof. The proofs are given in Appendix A. □

Remark 2.5. It clearly follows from (2.7) that

$$\left(\widehat{\theta}_n - \theta^*\right)^2 = O\left(\frac{\log \log n}{n}\right) \qquad \text{a.s.} \tag{2.8}$$

This almost sure rate of convergence will be useful in all the sequel.

3. On the serial correlation parameter

This section is devoted to the estimation of the serial correlation parameter ρ. First of all, it is necessary to evaluate, at step n, the least squares residuals given, for all $1 \leq k \leq n$, by

$$\widehat{\varepsilon}_k = X_k - \widehat{\theta}_n X_{k-1}. \tag{3.1}$$

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

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

The asymptotic behavior of $\widehat{\theta}_n$ and $\widehat{\rho}_n$ are quite similar. However, one can 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 ρ^*.

Theorem 3.1. *We have the almost sure convergence*

$$\lim_{n \to \infty} \widehat{\rho}_n = \rho^* \qquad \text{a.s.} \tag{3.3}$$

Our next result deals with the joint asymptotic normality of $\widehat{\theta}_n$ and $\widehat{\rho}_n$. Denote

$$\sigma_\rho^2 = \frac{(1 - \theta\rho)}{(1 + \theta\rho)^3}\left((\theta + \rho)^2(1 + \theta\rho)^2 + (\theta\rho)^2(1 - \theta^2)(1 - \rho^2)\right). \tag{3.4}$$

In addition, let Γ be the positive semidefinite covariance matrix given by

$$\Gamma = \begin{pmatrix} \sigma_\theta^2 & \theta\rho\sigma_\theta^2 \\ \theta\rho\sigma_\theta^2 & \sigma_\rho^2 \end{pmatrix}. \tag{3.5}$$

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

$$\sqrt{n}\begin{pmatrix} \widehat{\theta}_n - \theta^* \\ \widehat{\rho}_n - \rho^* \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma). \tag{3.6}$$

In particular,

$$\sqrt{n}\left(\widehat{\rho}_n - \rho^*\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\rho^2). \tag{3.7}$$

Remark 3.3. The covariance matrix Γ is invertible if and only if $\theta \neq -\rho$ since one can see by a straightforward calculation that

$$\det(\Gamma) = \frac{\sigma_\theta^2(\theta + \rho)^2(1 - \theta\rho)}{(1 + \theta\rho)}.$$

Moreover, in the particular case where $\theta = -\rho$,

$$\sqrt{n}\,\widehat{\theta}_n \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1 + \theta^2}{1 - \theta^2}\right) \qquad \text{and} \qquad \sqrt{n}\,\widehat{\rho}_n \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\theta^4(1 + \theta^2)}{1 - \theta^2}\right).$$

Finally, if the residuals are not correlated which means that $\rho = 0$,

$$\sqrt{n}\,\widehat{\rho}_n \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \theta^2\right).$$

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

Theorem 3.4. *Assume that (V_n) has a finite moment of order 4. Then, we have the quadratic strong law*

$$\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \left(\widehat{\rho}_k - \rho^*\right)^2 = \sigma_\rho^2 \qquad \text{a.s.} \tag{3.8}$$

where σ_ρ^2 is given by (3.4). *In addition, we also have the law of iterated logarithm*

$$\limsup_{n \to \infty} \left(\frac{n}{2 \log \log n}\right)^{1/2} \left(\widehat{\rho}_n - \rho^*\right) = -\liminf_{n \to \infty} \left(\frac{n}{2 \log \log n}\right)^{1/2} \left(\widehat{\rho}_n - \rho^*\right)$$

$$= \sigma_\rho \qquad \text{a.s.} \tag{3.9}$$

Consequently,

$$\limsup_{n \to \infty} \left(\frac{n}{2 \log \log n}\right) \left(\widehat{\rho}_n - \rho^*\right)^2 = \sigma_\rho^2 \qquad \text{a.s.} \tag{3.10}$$

Proof. The proofs are given in Appendix B. □

Remark 3.5. We obviously deduce from (3.10) that

$$\left(\widehat{\rho}_n - \rho^*\right)^2 = O\left(\frac{\log \log n}{n}\right) \qquad \text{a.s.} \tag{3.11}$$

The estimators $\widehat{\theta}_n$ and $\widehat{\rho}_n$ are self-normalized. Consequently, the asymptotic variances σ_θ^2 and σ_ρ^2 do not depend on the variance σ^2 associated with the driven noise (V_n). Nevertheless, we can estimate the variance σ^2 by

$$\widehat{\sigma}_n^2 = \left(1 - \frac{\widehat{\rho}_n^2}{\widehat{\theta}_n^2}\right) \frac{1}{n} \sum_{k=1}^{n} \widehat{\varepsilon}_k^2. \tag{3.12}$$

We clearly have the almost sure convergence

$$\lim_{n \to \infty} \widehat{\sigma}_n^2 = \sigma^2 \qquad \text{a.s.} \tag{3.13}$$

The proof is left to the reader as it follows essentially the same lines as that of (3.3).

4. On the Durbin–Watson statistic

We shall now investigate the asymptotic behavior of the Durbin–Watson statistic [10–12] given, for all $n \geq 1$, by

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

One can observe that $\widehat{D}_n$ and $\widehat{\rho}_n$ are asymptotically linked together by an affine transformation. Consequently, the results of the previous section allow us to establish the asymptotic behavior of $\widehat{D}_n$. We start with the almost sure convergence of $\widehat{D}_n$ to D^*.

Theorem 4.1. *We have the almost sure convergence*

$$\lim_{n \to \infty} \widehat{D}_n = D^* \qquad \text{a.s.} \tag{4.2}$$

Our next result deals with the asymptotic normality of $\widehat{D}_n$. It will be the keystone of a new bilateral statistical test deciding in particular, for a given significance level, whether residuals are autocorrelated or not. Denote $\sigma_D^2 = 4\sigma_\rho^2$ where the variance σ_ρ^2 is given by (3.4).

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

$$\sqrt{n}\left(\widehat{D}_n - D^*\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_D^2). \tag{4.3}$$

Remark 4.3. We immediately deduce from (4.3) that

$$\frac{n}{\sigma_D^2}\left(\widehat{D}_n - D^*\right)^2 \xrightarrow{\mathcal{L}} \chi^2 \tag{4.4}$$

where χ^2 has a Chi-square distribution with one degree of freedom.

Before providing our statistical test, we focus our attention on the almost sure rates of convergence for $\widehat{D}_n$ which are based on the asymptotic linear relation between $\widehat{D}_n$ and $\widehat{\rho}_n$.

Theorem 4.4. *Assume that (V_n) has a finite moment of order 4. Then, we have the quadratic strong law*

$$\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n}\left(\widehat{D}_k - D^*\right)^2 = \sigma_D^2 \qquad \text{a.s.} \tag{4.5}$$

In addition, we also have the law of iterated logarithm

$$\limsup_{n \to \infty}\left(\frac{n}{2\log\log n}\right)^{1/2}\left(\widehat{D}_n - D^*\right) = -\liminf_{n \to \infty}\left(\frac{n}{2\log\log n}\right)^{1/2}\left(\widehat{D}_n - D^*\right)$$
$$= \sigma_D \qquad \text{a.s.} \tag{4.6}$$

Consequently,

$$\limsup_{n \to \infty}\left(\frac{n}{2\log\log n}\right)\left(\widehat{D}_n - D^*\right)^2 = \sigma_D^2 \qquad \text{a.s.} \tag{4.7}$$

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 trivial case $\theta = 0$ since our statistical test procedure is of interest only for autoregressive processes. In addition, we shall note the existence of a critical case as introduced in Remark 3.3. Indeed, if $\theta = -\rho$, the covariance matrix Γ given by (3.5) is not invertible and the distribution of the statistic associated with the test we plan to establish will be degenerate. For this reason, we suggest a preliminary test for the hypothesis "$\theta = -\rho$", allowing us to switch from one test to another if necessary. More precisely, we first wish to test

$$\mathcal{H}_0 \; : \; \text{``}\theta = -\rho\text{''} \qquad \text{against} \qquad \mathcal{H}_1 \; : \; \text{``}\theta \neq -\rho\text{''}.$$

Under the null hypothesis $\mathcal{H}_0$, it is easy to see that $D^* = 2$. According to Remark 3.3, we have

$$\frac{n\left(1 - \theta^2\right)}{4\theta^4\left(1 + \theta^2\right)}\left(\widehat{D}_n - 2\right)^2 \xrightarrow{\mathcal{L}} \chi^2 \tag{4.8}$$

where χ^2 has a Chi-square distribution with one degree of freedom. Moreover, the model can be rewritten under $\mathcal{H}_0$, for all $n \geq 2$, as

$$X_n = \theta^2 X_{n-2} + V_n. \tag{4.9}$$

Then, we propose to make use of the standard least squares estimator $\widehat{\vartheta}_n^2$ of θ^2,

$$\widehat{\vartheta}_n^2 = \frac{\sum_{k=2}^n X_{k-2} X_k}{\sum_{k=2}^n X_{k-2}^2}. \tag{4.10}$$

Under $\mathcal{H}_0$, we have the almost sure convergence of $\widehat{\vartheta}_n^2$ to θ^2. In addition, we obviously have $D^* \neq 2$ under $\mathcal{H}_1$. These results under the null and the alternative hypothesis lead to Theorem 4.5, whose proof immediately follows from (4.8).

Theorem 4.5. *Assume that (V_n) has a finite moment of order 4, $\theta \neq 0$ and $\rho \neq 0$. Then, under the null hypothesis $\mathcal{H}_0$: "$\theta = -\rho$",*

$$\frac{n\left(1 - \widehat{\vartheta}_n^2\right)}{4\left(\widehat{\vartheta}_n^2\right)^2 \left(1 + \widehat{\vartheta}_n^2\right)} \left(\widehat{D}_n - 2\right)^2 \xrightarrow{\mathcal{L}} \chi^2 \tag{4.11}$$

where χ^2 has a Chi-square distribution with one degree of freedom. In addition, under the alternative hypothesis $\mathcal{H}_1$: "$\theta \neq -\rho$",

$$\lim_{n \to \infty} \frac{n\left(1 - \widehat{\vartheta}_n^2\right)}{4\left(\widehat{\vartheta}_n^2\right)^2 \left(1 + \widehat{\vartheta}_n^2\right)} \left(\widehat{D}_n - 2\right)^2 = +\infty \qquad \text{a.s.} \tag{4.12}$$

For a significance level α where $0 < \alpha < 1$, the acceptance and rejection regions are given by $\mathcal{A} = [0, z_\alpha]$ and $\mathcal{R} =]z_\alpha, +\infty[$ where z_α stands for the $(1-\alpha)$-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

$$\frac{n\left(1 - \widehat{\vartheta}_n^2\right)}{4\left(\widehat{\vartheta}_n^2\right)^2 \left(1 + \widehat{\vartheta}_n^2\right)} \left(\widehat{D}_n - 2\right)^2 \leq z_\alpha,$$

and will be rejected otherwise. Assume now that we cannot reject $\mathcal{H}_0$, which means that we admit *de facto* the hypothesis "$\theta = -\rho$". For a given value ρ_0 such that $|\rho_0| < 1$, we wish to test whether or not the serial correlation parameter is equal to ρ_0, setting

$$\mathcal{H}_0 : \text{"}\rho = \rho_0\text{"} \qquad \text{against} \qquad \mathcal{H}_1 : \text{"}\rho \neq \rho_0\text{"}.$$

One shall proceed once again to the test described by Theorem 4.5, taking ρ_0^2 in lieu of $\widehat{\vartheta}_n^2$, insofar as one can easily agree that our test statistic satisfies the same properties, under $\mathcal{H}_0$ as under $\mathcal{H}_1$, by virtue of Remark 3.3. This alternative solution is necessary to avoid the degenerate situation implied by the critical case $\theta = -\rho$. Let us now focus on the more widespread case where the preliminary test leads to a rejection of $\mathcal{H}_0$, admitting "$\theta \neq -\rho$". For that purpose, denote $\widetilde{\theta}_n = \widehat{\theta}_n + \widehat{\rho}_n - \rho_0$ and $\widetilde{D}_n = 2\left(1 - \widetilde{\rho}_n\right)$ where

$$\widetilde{\rho}_n = \frac{\rho_0 \widetilde{\theta}_n \left(\widetilde{\theta}_n + \rho_0\right)}{1 + \rho_0 \widetilde{\theta}_n}. \tag{4.13}$$

One can observe that our statistical test procedure works whatever the value of the autoregressive parameter θ inside the interval $]-1, 1[$ with $\theta \neq -\rho$. Moreover, it follows from (2.2) and (3.3) that under the null hypothesis $\mathcal{H}_0$,

$$\lim_{n \to \infty} \widetilde{\theta}_n = \theta + \rho_0 - \rho_0 = \theta \qquad \text{a.s.}$$

To construct our statistical test, we need to introduce more notations. Denote

$$\widehat{a}_n = -\rho_0\left(\widehat{\theta}_n + \widetilde{\theta}_n\right) = -\rho_0\left(2\widehat{\theta}_n + \widehat{\rho}_n - \rho_0\right),$$
$$\widehat{b}_n = 1 - \rho_0\widehat{\theta}_n,$$

and let $\widehat{w}_n$ be the vector of $\mathbb{R}^2$ given by $\widehat{w}_n' = (\widehat{a}_n, \widehat{b}_n)$. In addition, let

$$\widehat{\Gamma}_n = \begin{pmatrix} \widehat{\alpha}_n & \rho_0\widetilde{\theta}_n\widehat{\alpha}_n \\ \rho_0\widetilde{\theta}_n\widehat{\alpha}_n & \widehat{\beta}_n \end{pmatrix}. \tag{4.14}$$

where $\widehat{\alpha}_n$ and $\widehat{\beta}_n$ are defined as

$$\widehat{\alpha}_n = \frac{\left(1 - \widetilde{\theta}_n^2\right)\left(1 - \rho_0\widetilde{\theta}_n\right)(1 - \rho_0^2)}{\left(1 + \rho_0\widetilde{\theta}_n\right)^3},$$

$$\widehat{\beta}_n = \frac{\left(1 - \rho_0\widetilde{\theta}_n\right)}{\left(1 + \rho_0\widetilde{\theta}_n\right)^3}\left(\left(\widetilde{\theta}_n + \rho_0\right)^2\left(1 + \rho_0\widetilde{\theta}_n\right)^2 + \left(\rho_0\widetilde{\theta}_n\right)^2\left(1 - \widetilde{\theta}_n^2\right)(1 - \rho_0^2)\right).$$

Furthermore, denote

$$\widehat{\tau}_n^2 = \frac{4}{\left(1 + \rho_0\widetilde{\theta}_n\right)^2}\widehat{w}_n'\widehat{\Gamma}_n\widehat{w}_n. \tag{4.15}$$

Theorem 4.6. *Assume that (V_n) has a finite moment of order 4, $\theta \neq \rho_0$ and $\theta \neq -\rho$. Then, under the null hypothesis $\mathcal{H}_0$: "$\rho = \rho_0$",*

$$\frac{n}{\widehat{\tau}_n^2}\left(\widehat{D}_n - \widetilde{D}_n\right)^2 \xrightarrow{\mathcal{L}} \chi^2 \tag{4.16}$$

where χ^2 has a Chi-square distribution with one degree of freedom. In addition, under the alternative hypothesis $\mathcal{H}_1$: "$\rho \neq \rho_0$",

$$\lim_{n\to\infty} \frac{n}{\widehat{\tau}_n^2}\left(\widehat{D}_n - \widetilde{D}_n\right)^2 = +\infty \qquad \text{a.s.} \tag{4.17}$$

One can observe by a symmetry argument on the role played by θ and ρ, that the assumption $\theta = \rho_0$ is not restrictive since the latter can be seen as another way of expressing Theorem 4.6. From a practical point of view, for a significance level α where $0 < \alpha < 1$, the acceptance and rejection regions are given by $\mathcal{A} = [0, z_\alpha]$ and $\mathcal{R} =]z_\alpha, +\infty[$ where z_α stands for the $(1 - \alpha)$-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

$$\frac{n}{\widehat{\tau}_n^2}\left(\widehat{D}_n - \widetilde{D}_n\right)^2 \leq z_\alpha,$$

and will be rejected otherwise. Moreover, if one wishes to test

$$\mathcal{H}_0 : \text{``}\rho = 0\text{''} \qquad \text{against} \qquad \mathcal{H}_1 : \text{``}\rho \neq 0\text{''},$$

our statistical test procedure is very simple. As a matter of fact, we are in the particular case $\rho_0 = 0$ which means that $\widetilde{D}_n = 2$, $\widehat{a}_n = 0$ and $\widehat{b}_n = 1$. We can also replace $\widetilde{\theta}_n$ by $\widehat{\theta}_n$ so $\widehat{\tau}_n^2$ reduces to $\widehat{\tau}_n^2 = 4\widehat{\theta}_n^2$. In the next section, we will explain why the following test is asymptotically equivalent to the *h-test* proposed by Durbin [8].

Theorem 4.7. *Assume that (V_n) has a finite moment of order 4, $\theta \neq 0$ and $\theta \neq -\rho$. Then, under the null hypothesis $\mathcal{H}_0$: "$\rho = 0$",*

$$\frac{n}{4\widehat{\theta}_n^2} \left(\widehat{D}_n - 2\right)^2 \xrightarrow{\mathcal{L}} \chi^2 \tag{4.18}$$

where χ^2 has a Chi-square distribution with one degree of freedom. In addition, under the alternative hypothesis $\mathcal{H}_1$: "$\rho \neq 0$",

$$\lim_{n \to \infty} \frac{n}{4\widehat{\theta}_n^2} \left(\widehat{D}_n - 2\right)^2 = +\infty \qquad \text{a.s.} \tag{4.19}$$

Proof. The proofs are given in Appendix C.   □

5. COMPARISONS WITH PREVIOUS STATISTICAL TESTS

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. More precisely, we will focus our attention on the well-known Ljung–Box [4] and Box–Pierce [5] tests widely used in Econometry on the one hand, and on the alternative procedures proposed by Durbin [8] in order to prevent the inappropriate use of the original one in the case of lagged-dependent random variables, on the other hand. First, we will briefly recall some theoretical aspects of the procedures to be compared. Then, we shall explain how, on the basis of the empirical power, the test procedure in Theorem 4.7 performs better than Ljung–Box [4] and Box–Pierce [5] *portmanteau* tests, and why it is asymptotically equivalent to the most powerful alternative test of Durbin [8], better known as *h-test*, in the case of a first-order autoregressive process. We shall assume in all the sequel that $\theta \neq 0$ is a statistically significant parameter and we will test

$$\mathcal{H}_0 : \text{``}\rho = 0\text{''} \qquad \text{against} \qquad \mathcal{H}_1 : \text{``}\rho \neq 0\text{''}.$$

5.1. The Ljung–Box and Box–Pierce tests

It is important to note that these *portmanteau* tests are built to detect the presence of a significative serial correlation of any order. Of course, in this paper, we shall only focus our attention on the first-order restriction. The Box–Pierce procedure uses the test statistic given by

$$\widehat{Q}_n^{\text{BP}} = n\widehat{\rho}_n^2 \tag{5.1}$$

where $\widehat{\rho}_n$ is given in (3.2). The Ljung–Box test is an adjustement of the latter, to improve the precision on small-sized samples. It is defined as

$$\widehat{Q}_n^{\text{LB}} = \frac{n(n+2)}{n-1}\widehat{\rho}_n^2. \tag{5.2}$$

If there exists a white noise (W_n) such that the estimated residual set $(\widehat{\varepsilon}_n)$ given by (3.1) could by rewritten, for all $n \geq 1$, as $\widehat{\varepsilon}_n = \rho\widehat{\varepsilon}_{n-1} + W_n$, then under $\mathcal{H}_0$,

$$\widehat{Q}_n^{\text{BP}} \xrightarrow{\mathcal{L}} \chi^2 \qquad \text{and} \qquad \widehat{Q}_n^{\text{LB}} \xrightarrow{\mathcal{L}} \chi^2 \tag{5.3}$$

where χ^2 has a Chi-square distribution with one degree of freedom, and go a.s. to infinity under $\mathcal{H}_1$. Moreover, the asymptotic normality given in Theorem 3.2 ensures that, under $\mathcal{H}_0$,

$$\frac{n\widehat{\rho}_n^2}{\widehat{\theta}_n^2} \xrightarrow{\mathcal{L}} \chi^2. \tag{5.4}$$

This explains why our statistical procedure is more precise for testing the existence of a significative first-order serial correlation in a first-order autoregressive process than the Box–Pierce and Ljung–Box ones as this will be observed on simulated data.

5.2. The alternative Durbin tests

Park [21] has observed by simulations that the alternative *t-test* and *h-test* proposed by Durbin [8] for autocorrelated error terms in a lagged-dependent random variables framework significantly outperform the original Durbin–Watson test. Let us try to compare our statistical procedure with both of them. In fact, we will only focus our attention on the *h-test* since, as it is clearly specified by Park [21], the *t-test* merely consists in fitting by least squares a linear regression of $(\widehat{\varepsilon}_n)$ on the one-period lagged residuals $(\widehat{\varepsilon}_{n-1})$ and in testing the significance of the coefficient of $(\widehat{\varepsilon}_{n-1})$. One has to note that such a procedure is quasi equivalent to the one of Box–Pierce, in our precise framework. The test statistic associated with the *h-test* is given by

$$\widehat{H}_n = \left(1 - \frac{\widehat{D}_n}{2}\right)^2 \left[\frac{n}{1 - n\widehat{\mathrm{Var}}(\widehat{\theta}_n)}\right] \tag{5.5}$$

where $\widehat{D}_n$ is the Durbin–Watson statistic given in (4.1). Under $\mathcal{H}_0$, it is shown by Durbin [8] that

$$\widehat{H}_n \xrightarrow{\mathcal{L}} \chi^2 \tag{5.6}$$

where χ^2 stands for the Chi-square distribution with one degree of freedom. Moreover, it immediately follows from the very definition of $\widehat{\theta}_n$ that, under $\mathcal{H}_0$,

$$\widehat{\mathrm{Var}}(\widehat{\theta}_n) = \widehat{\sigma}_n^2 S_{n-1}^{-1} \qquad \text{where} \qquad S_n = \sum_{k=0}^{n} X_k^2$$

and $\widehat{\sigma}_n^2$ is the least squares estimator of the variance of the sequence (ε_n),

$$\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{k=1}^{n} \left(X_k - \widehat{\theta}_n X_{k-1}\right)^2.$$

As a consequence,

$$n\widehat{\mathrm{Var}}(\widehat{\theta}_n) = 1 - \widehat{\theta}_n^2 + X_n^2 S_{n-1}^{-1}$$

where the remainder term converges a.s. to zero by virtue of Lemma A.2 with $a = 2$. Accordingly, it follows from a straightforward calculation that $\widehat{H}_n$ is almost surely equivalent to the test statistic given by (4.18), that is

$$\widehat{Z}_n = \frac{n}{4\widehat{\theta}_n^2} \left(\widehat{D}_n - 2\right)^2. \tag{5.7}$$

5.3. Empirical power and comparisons

We shall now compare the empirical power of the statistical tests based on $\widehat{Q}_n^{\mathrm{BP}}$, $\widehat{Q}_n^{\mathrm{LB}}$, $\widehat{H}_n$ and $\widehat{Z}_n$ by simulations. These statistics are used to test $\mathcal{H}_0$ against $\mathcal{H}_1$ at the 0.05 level of significance. We repeat the procedure 1000 times and compute the frequency with which $\mathcal{H}_0$ is rejected, for different values of ρ. At the end of each experiment, we obtain the empirical power of the procedures and compare their performance, by estimating

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

To minimize the impact of the initial values, the first 50 observations are discarded and the remaining observations taken as a sample, according to the protocol of Park [21]. We choose a significative $\theta = 0.5$ and standard normal residuals. Note that σ^2 has no influence on the test statistics since $\widehat{\rho}_n$ is self-normalized. Finally, sample sizes $n = 30$, $n = 100$ and $n = 500$ are used in order to evaluate the empirical power even on small-sized samples. All numerical results are given in Tables 1–3.

TABLE 1. Figures give the power of the tests for $n = 30$, the frequencies of $\mathcal{H}_0$ not rejected are in parentheses.

$n = 30$	Values of ρ										
	-0.9	-0.6	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.6	0.9
Box–Pierce	0.20	0.00	0.00	0.00	0.00	–	0.00	0.01	0.02	0.22	0.63
	(0.80)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(0.99)	(0.98)	(0.78)	(0.40)
Ljung–Box	0.24	0.00	0.00	0.00	0.00	–	0.00	0.01	0.03	0.26	0.60
	(0.76)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(0.99)	(0.97)	(0.74)	(0.37)
h-test	0.58	0.24	0.16	0.14	0.17	–	0.18	0.22	0.28	0.49	0.71
	(0.42)	(0.76)	(0.84)	(0.86)	(0.83)	(0.85)	(0.82)	(0.78)	(0.72)	(0.51)	(0.29)
Durbin–Watson	0.63	0.54	0.31	0.22	0.17	–	0.15	0.18	0.23	0.46	0.70
	(0.37)	(0.47)	(0.69)	(0.78)	(0.84)	(0.87)	(0.85)	(0.82)	(0.77)	(0.55)	(0.30)

TABLE 2. Figures give the power of the tests for $n = 100$, the frequencies of $\mathcal{H}_0$ not rejected are in parentheses.

$n = 100$	Values of ρ										
	-0.9	-0.6	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.6	0.9
Box–Pierce	0.89	0.00	0.00	0.00	0.00	–	0.00	0.02	0.10	0.75	0.99
	(0.11)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(0.98)	(0.90)	(0.26)	(0.01)
Ljung–Box	0.90	0.00	0.00	0.00	0.00	–	0.00	0.02	0.11	0.75	0.99
	(0.10)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(0.98)	(0.89)	(0.25)	(0.01)
h-test	0.99	0.62	0.25	0.15	0.10	–	0.14	0.22	0.41	0.87	1.00
	(0.01)	(0.38)	(0.75)	(0.85)	(0.90)	(0.92)	(0.86)	(0.79)	(0.59)	(0.13)	(0.00)
Durbin–Watson	0.99	0.82	0.26	0.13	0.08	–	0.13	0.20	0.39	0.87	1.00
	(0.01)	(0.18)	(0.74)	(0.87)	(0.92)	(0.94)	(0.87)	(0.80)	(0.61)	(0.14)	(0.00)

TABLE 3. Figures give the power of the tests for $n = 500$, the frequencies of $\mathcal{H}_0$ not rejected are in parentheses.

$n = 500$	Values of ρ										
	-0.9	-0.6	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.6	0.9
Box–Pierce	1.00	0.01	0.00	0.00	0.00	–	0.00	0.19	0.70	1.00	1.00
	(0.00)	(0.99)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(0.81)	(0.30)	(0.00)	(0.00)
Ljung–Box	1.00	0.01	0.00	0.00	0.00	–	0.00	0.19	0.70	1.00	1.00
	(0.00)	(0.99)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(0.81)	(0.30)	(0.00)	(0.00)
h-test	1.00	0.94	0.87	0.54	0.17	–	0.24	0.63	0.93	1.00	1.00
	(0.00)	(0.06)	(0.13)	(0.46)	(0.83)	(0.95)	(0.76)	(0.37)	(0.07)	(0.00)	(0.00)
Durbin–Watson	1.00	1.00	0.86	0.54	0.17	–	0.24	0.63	0.93	1.00	1.00
	(0.00)	(0.00)	(0.14)	(0.46)	(0.83)	(0.95)	(0.77)	(0.37)	(0.07)	(0.00)	(0.00)

Figures 1 and 2 below represent the frequencies of $\mathcal{H}_0$ not rejected when ρ varies from -0.9 to 0.9, for $n = 100$ and then for $n = 500$. BP, LB, HT and DW respectively stand for Box–Pierce, Ljung–Box, H-test and Durbin–Watson.

It clearly appears that the *h-test* and our procedure are more powerful than the *portmanteau* tests, in our precise framework of first-order processes. Note that an unexpected peak occurs when $\rho = -0.6$, due to the fact that we are close to the critical situation where $\theta = -\rho$, in which $\widehat{\rho}_n$ converges a.s. to zero even if there is a significative serial correlation. As we have seen before in (5.4), 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. That is the reason why we constantly observe that these procedures lead to inadequate conclusions in the region
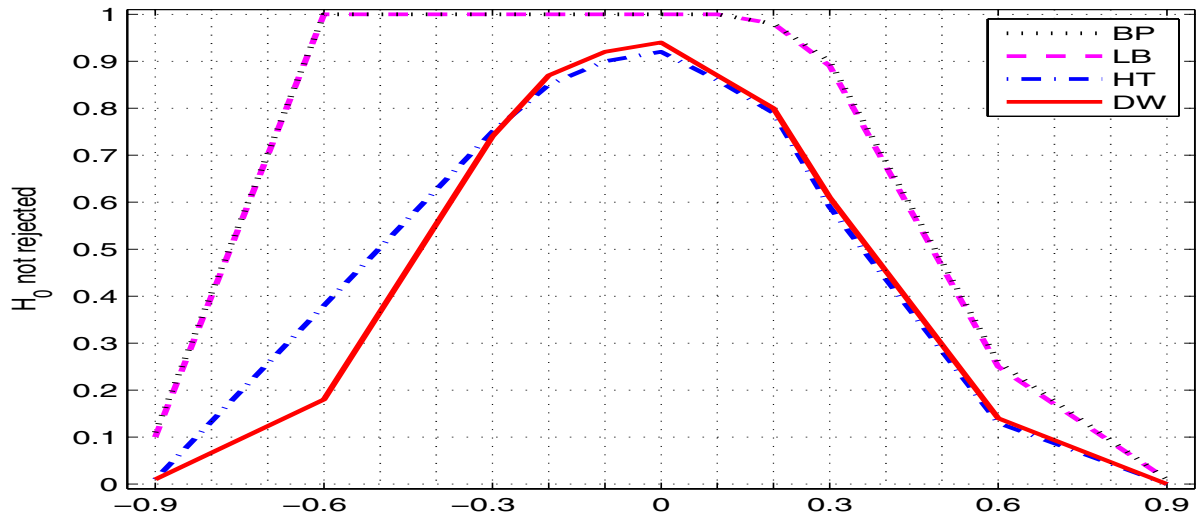
FIGURE 1. Frequencies of $\mathcal{H}_0$ not rejected for sample size $n = 100$ and ρ varying from -0.9 to 0.9 on the abscissa.
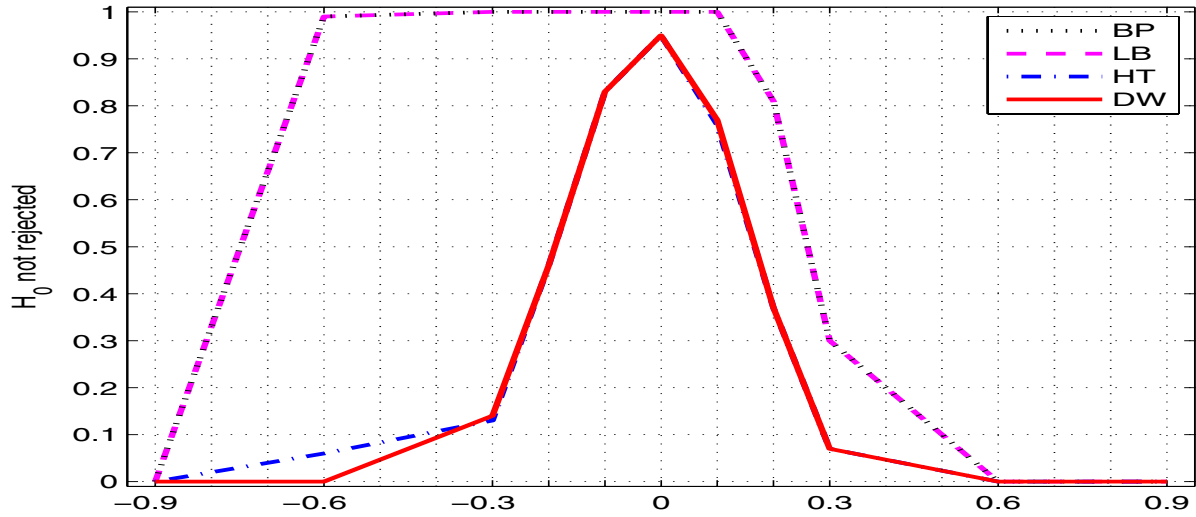

FIGURE 2. Frequencies of $\mathcal{H}_0$ not rejected for sample size $n = 500$ and ρ varying from -0.9 to 0.9 on the abscissa.
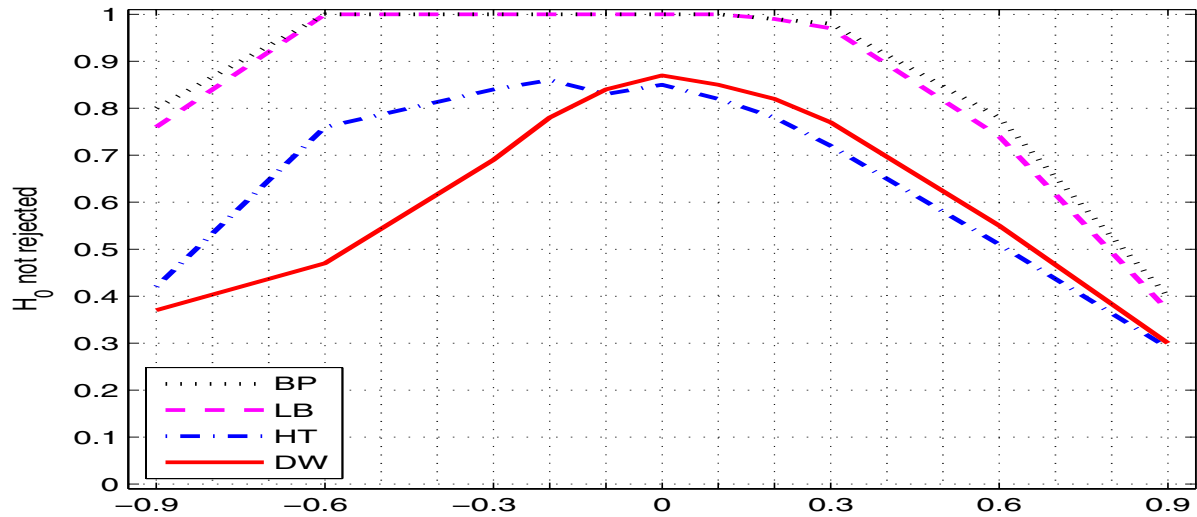
FIGURE 3. Frequencies of $\mathcal{H}_0$ not rejected for sample size $n = 30$ and ρ varying from -0.9 to 0.9 on the abscissa.

$0 < |\rho| \leq 0.3$ even if, under $\mathcal{H}_0$ and for the same reason, they are more reliable than the *h-test* and the test we proposed. Under $\mathcal{H}_1$, we suggest to make use of the latter ones, clearly more powerful. The asymptotic equivalence between them implies that they share the same behavior on large samples, as it appears on the graphs above. However, our procedure is less impacted by the critical situation where $\theta = -\rho$. Figure 3 displays the results for $n = 30$.

On small-sized samples, all procedures are obviously less powerful, but once again we observe that our procedure outperforms both the *portmanteau* tests and the *h-test*, particularly when θ and ρ are of different sign. Nevertheless, results are quite satisfying in the area of uncertainty where $|\rho| \leq 0.2$. In consequence, even if the *h-test* and our procedure are equally powerful on large samples, it seems advisable to make use of the one described in Theorem 4.7 on small-sized ones, to detect a significative first-order serial correlation in the residuals of a first-order autoregressive process. Note that the so-called Breusch-Godfrey test [6, 13] is not as efficient as our procedure on small-sized samples, but we have observed that it would lead to similar results on large samples. In order to keep this section brief, we have chosen a particular $\theta = 0.5$. However, it is clear that all those simulation results lead to the same conclusions when θ varies in a significative area, say $|\theta|$ not too close to 0. On the other side, when $|\theta|$ is close to 1, we have from (5.4) that all the statistical tests are asymptotically equivalent.

6. CONCLUSION

Via an extensive use of the theory of martingales, we have provided a sharp analysis on the asymptotic behavior of the least squares estimators $\widehat{\theta}_n$ and $\widehat{\rho}_n$ which has allowed us to deduce the asymptotic behavior of the Durbin–Watson statistic $\widehat{D}_n$ for the first-order autoregressive process. More precisely, we have established the almost sure convergence and the asymptotic normality for all three estimators $\widehat{\theta}_n$, $\widehat{\rho}_n$ and $\widehat{D}_n$. In addition, we have proposed a new bilateral statistical procedure for testing serial correlation, built on $\widehat{D}_n$. All these results give a new light on the well-known test of Durbin–Watson in a context of lagged dependent random variables. From a practical standpoint and for a matter of completeness, we may wonder about the estimation of the true

values of the parameters θ and ρ. As soon as $\theta \neq -\rho$, we can estimate θ and ρ by solving the nonlinear system of two equations given, for $\widehat{a}_n = \widehat{\theta}_n + \widehat{\rho}_n$ and $\widehat{b}_n = \widehat{\rho}_n/\widehat{\theta}_n$, by

$$\begin{cases} \lim_{n\to\infty} \widehat{a}_n = \theta + \rho \\ \lim_{n\to\infty} \widehat{b}_n = \theta\rho \end{cases} \qquad a.s.$$

One can easily find two couples of solutions, symmetrically linked together. For example, assuming $\theta < \rho$, we propose to make use of

$$\widetilde{\theta}_n = \frac{\widehat{a}_n - \sqrt{\widehat{a}_n^2 - 4\widehat{b}_n}}{2} \qquad \text{and} \qquad \widetilde{\rho}_n = \frac{\widehat{a}_n + \sqrt{\widehat{a}_n^2 - 4\widehat{b}_n}}{2},$$

merely inverting the values of $\widetilde{\theta}_n$ and $\widetilde{\rho}_n$ whether, for some statistical argument, we would rather choose $\theta > \rho$. This work lifts the veil on a set of questions for long left unanswered about the pioneer works of Durbin and Watson [10–12,19,20] and contributes to bring the test procedure back into light. Moderate deviation principles on the Durbin–Watson statistic have recently been established by Bitseki Penda *et al.* [3] and the extension to the more general stable p-order autoregressive processes may be found in [22]. In addition, it would be very interesting to investigate the asymptotic behavior of the Durbin–Watson statistic in the explosive case where $|\theta| > 1$ or $|\rho| > 1$.

APPENDIX A. PROOFS OF THE AUTOREGRESSIVE PARAMETER RESULTS

A.1. Proof of Theorem 2.1

We start with some useful technical lemmas we shall make repeatedly use of. The proof of the first one may be found in [7] page 24.

Lemma A.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,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n |V_k|^a = \mathbb{E}[|V_1|^a] \qquad a.s. \tag{A.1}$$

and

$$\sup_{1\leq k\leq n} |V_k| = o(n^{1/a}) \qquad a.s. \tag{A.2}$$

Lemma A.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 $|\theta| < 1$, $|\rho| < 1$, then*

$$\sum_{k=1}^n |X_k|^a = O(n) \qquad a.s. \tag{A.3}$$

and

$$\sup_{1\leq k\leq n} |X_k| = o(n^{1/a}) \qquad a.s. \tag{A.4}$$

Remark A.3. In the particular case $a = 2$, we obtain that

$$\sum_{k=1}^n X_k^2 = O(n) \qquad \text{and} \qquad \sup_{1\leq k\leq n} X_k^2 = o(n) \qquad a.s.$$

Proof. It follows from (1.1) that for all $n \geq 1$,

$$|X_n| \leq |\theta|^n |X_0| + \sum_{k=1}^n |\theta|^{n-k} |\varepsilon_k|.$$ (A.5)

Consequently, as $|\theta| < 1$, we obtain that

$$\sup_{1 \leq k \leq n} |X_k| \leq \frac{1}{1 - |\theta|} \left(|X_0| + \sup_{1 \leq k \leq n} |\varepsilon_k| \right).$$ (A.6)

By the same token, as $|\rho| < 1$, we also deduce from (1.1) that

$$\sup_{1 \leq k \leq n} |\varepsilon_k| \leq \frac{1}{1 - |\rho|} \left(|\varepsilon_0| + \sup_{1 \leq k \leq n} |V_k| \right).$$ (A.7)

Hence, (A.2) together with (A.6) and (A.7) obviously imply

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

Furthermore, let b be the conjugate exponent of a,

$$\frac{1}{a} + \frac{1}{b} = 1.$$

It follows from (A.5) and Holder's inequality that for all $n \geq 1$,

$$|X_n| \leq \left(|\theta|^n |X_0|^a + \sum_{k=1}^n |\theta|^{n-k} |\varepsilon_k|^a \right)^{1/a} \left(\sum_{k=0}^n |\theta|^{n-k} \right)^{1/b}$$

which implies that

$$
\begin{aligned}
|X_n|^a &\leq \left(\sum_{k=0}^n |\theta|^{n-k} \right)^{a/b} \left(|\theta|^n |X_0|^a + \sum_{k=0}^n |\theta|^{n-k} |\varepsilon_k|^a \right), \\
&\leq \left(\sum_{k=0}^\infty |\theta|^k \right)^{a/b} \left(|\theta|^n |X_0|^a + \sum_{k=1}^n |\theta|^{n-k} |\varepsilon_k|^a \right), \\
&\leq \left(1 - |\theta| \right)^{-a/b} \left(|\theta|^n |X_0|^a + \sum_{k=1}^n |\theta|^{n-k} |\varepsilon_k|^a \right).
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
\sum_{k=1}^n |X_k|^a &\leq \left(1 - |\theta| \right)^{-a/b} \left(\sum_{k=1}^n |\theta|^k |X_0|^a + \sum_{k=1}^n \sum_{\ell=1}^k |\theta|^{k-\ell} |\varepsilon_\ell|^a \right), \\
&\leq \left(1 - |\theta| \right)^{-a/b} \left(|X_0|^a \sum_{k=1}^n |\theta|^k + \sum_{\ell=1}^n |\varepsilon_\ell|^a \sum_{k=\ell}^n |\theta|^{k-\ell} \right), \\
&\leq \left(1 - |\theta| \right)^{-a} \left(|X_0|^a + \sum_{k=1}^n |\varepsilon_k|^a \right).
\end{aligned}
$$ (A.8)

Via the same lines, we also obtain that

$$\sum_{k=1}^{n} |\varepsilon_k|^a \le \left(1 - |\rho|\right)^{-a} \left(|\varepsilon_0|^a + \sum_{k=1}^{n} |V_k|^a\right). \tag{A.9}$$

Finally, (A.1) together with (A.8) and (A.9) lead to (A.3), which completes the proof of Lemma A.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 $n \ge 2$, by

$$X_n = (\theta + \rho)X_{n-1} - \theta\rho X_{n-2} + V_n. \tag{A.10}$$

For all $n \ge 0$, let

$$S_n = \sum_{k=0}^{n} X_k^2, \tag{A.11}$$

$$P_n = \sum_{k=1}^{n} X_k X_{k-1}, \tag{A.12}$$

$$M_n = \sum_{k=1}^{n} X_{k-1} V_k \tag{A.13}$$

where $P_0 = 0$ and $M_0 = 0$. It is not hard to see from (A.10) that for all $n \ge 2$,

$$P_n = (\theta + \rho)S_{n-1} - \theta\rho P_{n-1} + M_n + \rho X_0(\varepsilon_0 - X_0)$$

which implies that

$$(1 + \theta\rho)P_n = (\theta + \rho)S_{n-1} + M_n + \theta\rho X_n X_{n-1} + \rho X_0(\varepsilon_0 - X_0). \tag{A.14}$$

Via (2.1), (A.14) leads to the main decomposition

$$\widehat{\theta}_n = \frac{\theta + \rho}{1 + \theta\rho} + \frac{1}{1 + \theta\rho}\frac{M_n}{S_{n-1}} + \frac{1}{1 + \theta\rho}\frac{R_n}{S_{n-1}} \tag{A.15}$$

where the remainder term

$$R_n = \theta\rho X_n X_{n-1} + \rho X_0(\varepsilon_0 - X_0).$$

For all $n \ge 1$, denote by $\mathcal{F}_n$ the σ-algebra of the events occurring up to time n, $\mathcal{F}_n = \sigma(X_0, \varepsilon_0, V_1, \ldots, V_n)$. We infer from (A.13) that (M_n) is a locally square-integrable real martingale [7,14] with predictable quadratic variation given by $\langle M \rangle_0 = 0$ and for all $n \ge 1$,

$$\langle M \rangle_n = \sum_{k=1}^{n} \mathbb{E}[(M_k - M_{k-1})^2 | \mathcal{F}_{k-1}],$$

$$= \sum_{k=1}^{n} \mathbb{E}[X_{k-1}^2 V_k^2 | \mathcal{F}_{k-1}] = \sigma^2 S_{n-1}.$$

Furthermore, it follows from (A.10) and Corollary 1.3.25 of [7] that $n = O(S_n)$ a.s. Then, we deduce from the strong law of large numbers for martingales given *e.g.* by Theorem 1.3.15 of [7] that

$$\lim_{n \to \infty} \frac{M_n}{\langle M \rangle_n} = 0 \qquad \text{a.s.}$$

which of course ensures that

$$\lim_{n\to\infty} \frac{M_n}{S_{n-1}} = 0 \qquad \text{a.s.} \tag{A.16}$$

It remains to show that the remainder $R_n = o(S_{n-1})$ a.s. We have from (A.4) with $a = 2$ that $|X_n| = o(\sqrt{n})$ a.s. which implies that $R_n = o(n)$ a.s. However, we already saw that $n = O(S_n)$ a.s. Hence,

$$\lim_{n\to\infty} \frac{R_n}{S_{n-1}} = 0 \qquad \text{a.s.} \tag{A.17}$$

Finally, it follows from (A.15) together with (A.16) and (A.17) that

$$\lim_{n\to\infty} \widehat{\theta}_n = \frac{\theta + \rho}{1 + \theta\rho} \qquad \text{a.s.}$$

which achieves the Proof of Theorem 2.1. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

A.2. Proof of Theorem 2.2

In order to establish the asymptotic normality of the least squares estimator $\widehat{\theta}_n$, it is necessary to be more precise in Lemma A.2 with $a = 2$.

Lemma A.4. *Assume that the initial values X_0 and ε_0 are square-integrable and that (V_n) is a sequence of square-integrable, independent and identically distributed random variables with zero mean and variance $\sigma^2 > 0$. Then,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k^2 = \ell \qquad \text{a.s.} \tag{A.18}$$

where the limiting value

$$\ell = \frac{\sigma^2(1 + \theta\rho)}{(1 - \theta^2)(1 - \theta\rho)(1 - \rho^2)}. \tag{A.19}$$

In addition, if $\ell_1 = \theta^*\ell$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k X_{k-1} = \ell_1 \qquad \text{a.s.} \tag{A.20}$$

Proof. We deduce from the fundamental autoregressive equation (A.10) together with straightforward calculations that for all $n \geq 2$,

$$S_n = (\theta + \rho)^2 S_{n-1} + (\theta\rho)^2 S_{n-2} + L_n - 2\theta\rho(\theta + \rho)P_{n-1} + 2(\theta + \rho)M_n - 2\theta\rho N_n + \xi_1$$

where S_n, P_n and M_n are respectively given by (A.11)–(A.13), the last term $\xi_1 = (1 - 2\theta\rho - \rho^2)X_0^2 + \rho^2\varepsilon_0^2 + 2\theta\rho X_0\varepsilon_0 + 2\rho(\varepsilon_0 - X_0)V_1$ and, for all $n \geq 2$,

$$L_n = \sum_{k=1}^{n} V_k^2, \tag{A.21}$$

$$N_n = \sum_{k=2}^{n} X_{k-2}V_k. \tag{A.22}$$

Consequently,

$$(1 - (\theta + \rho)^2 - (\theta\rho)^2)S_n = L_n - 2\theta\rho(\theta + \rho)P_n + 2(\theta + \rho)M_n - 2\theta\rho N_n - T_n \tag{A.23}$$

where the remainder term

$$T_n = ((\theta + \rho)^2 + (\theta\rho)^2)X_n^2 + (\theta\rho)^2 X_{n-1}^2 - 2\theta\rho(\theta + \rho)X_n X_{n-1} - \xi_1.$$

It follows from (A.1) with $a = 2$ that

$$\lim_{n \to \infty} \frac{L_n}{n} = \sigma^2 \qquad \text{a.s.} \tag{A.24}$$

In addition, we already saw from equation (A.14) that $P_n = \theta^* S_{n-1} + o(S_{n-1})$ a.s. which clearly implies

$$P_n = \theta^* S_n + o(S_n) \qquad \text{a.s.} \tag{A.25}$$

Moreover, (N_n) given by (A.22) is a locally square-integrable real martingale sharing the same almost sure properties than (M_n). More precisely, its predictable quadratic variation is given by $\langle N \rangle_n = \sigma^2 S_{n-2}$ which means that

$$\lim_{n \to \infty} \frac{M_n}{S_n} = 0 \qquad \text{and} \qquad \lim_{n \to \infty} \frac{N_n}{S_n} = 0 \qquad \text{a.s.} \tag{A.26}$$

Furthermore, we have from (A.4) with $a = 2$ that $X_n^2 = o(n)$ a.s. It ensures by use of $n = O(S_n)$ a.s. that

$$\lim_{n \to \infty} \frac{T_n}{S_n} = 0 \qquad \text{a.s.} \tag{A.27}$$

Therefore, it follows from the conjunction of (A.23), (A.25), (A.26) and (A.27) that

$$(1 - (\theta + \rho)^2 - (\theta\rho)^2 + 2\theta\rho(\theta + \rho)\theta^*)S_n = L_n + o(S_n) \qquad \text{a.s.} \tag{A.28}$$

Finally, dividing both sides of (A.28) by n and letting n goes to infinity, we deduce from (A.24) that

$$\lim_{n \to \infty} \frac{S_n}{n} = \frac{\sigma^2(1 + \theta\rho)}{(1 - \theta\rho)(1 - \theta^2)(1 - \rho^2)} \qquad \text{a.s.}$$

$$\lim_{n \to \infty} \frac{P_n}{n} = \frac{\sigma^2(\theta + \rho)}{(1 - \theta\rho)(1 - \theta^2)(1 - \rho^2)} \qquad \text{a.s.}$$

These two limits will often be used in all the sequel. $\qquad \square$

Proof of Theorem 2.2. We are now in the position to prove the asymptotic normality of $\widehat{\theta}_n$. We have from the main decomposition (A.15) that for all $n \geq 2$,

$$\sqrt{n}\left(\widehat{\theta}_n - \theta^*\right) = \sqrt{n}\left(\frac{\sigma^2}{1 + \theta\rho}\right)\frac{M_n}{\langle M \rangle_n} + \sqrt{n}\left(\frac{1}{1 + \theta\rho}\right)\frac{R_n}{S_{n-1}}. \tag{A.29}$$

We shall make use of the central limit theorem for martingales given *e.g.* by Corollary 2.1.10 of [7], to establish the asymptotic normality of the first term in the right-hand side of (A.29). On the other hand, we will also show that the second term $\sqrt{n}R_n/S_{n-1}$ goes to zero almost surely. First of all, it follows from (A.18) that

$$\lim_{n \to \infty} \frac{\langle M \rangle_n}{n} = \sigma^2 \ell \qquad \text{a.s.} \tag{A.30}$$

From now on, in order to apply the central limit theorem for martingales, it is necessary to prove that the Lindeberg condition is satisfied. For all $n \geq 1$, denote $\Delta M_n = X_{n-1}V_n$. One only has to show that for all $\varepsilon > 0$,

$$\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\left[|\Delta M_k|^2 \mathrm{I}_{|\Delta M_k| \geq \varepsilon\sqrt{n}}|\mathcal{F}_{k-1}\right] \xrightarrow{\mathcal{P}} 0. \tag{A.31}$$

However, we have assumed that (V_n) has a finite moment of order 4, $\tau^4 = \mathbb{E}[V_1^4]$. Hence, for all $n \geq 1$, $\mathbb{E}[|\Delta M_n|^4|\mathcal{F}_{n-1}] = \mathbb{E}[X_{n-1}^4 V_n^4|\mathcal{F}_{n-1}] = \tau^4 X_{n-1}^4$. In addition, we deduce from (A.3) with $a = 4$ that

$$\sum_{k=1}^{n} X_k^4 = O(n) \qquad \text{a.s.} \tag{A.32}$$

Therefore, for all $\varepsilon > 0$,

$$\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\left[|\Delta M_k|^2 \mathrm{I}_{|\Delta M_k| \geq \varepsilon\sqrt{n}}|\mathcal{F}_{k-1}\right] \leq \frac{1}{\varepsilon^2 n^2}\sum_{k=1}^{n} \mathbb{E}\left[|\Delta M_k|^4|\mathcal{F}_{k-1}\right],$$

$$\leq \frac{\tau^4}{\varepsilon^2 n^2}\sum_{k=1}^{n} X_{k-1}^4.$$

Consequently, (A.32) ensures that

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

and the Lindeberg condition is clearly satisfied. We can conclude from the central limit theorem for martingales that

$$\frac{1}{\sqrt{n}}M_n \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \sigma^2\ell\right) \tag{A.33}$$

in which the asymptotic variance is the deterministic limit given by (A.30). Moreover, as $\ell > 0$, we have from (A.33) and Slutsky's lemma that

$$\sqrt{n}\frac{M_n}{\langle M \rangle_n} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \sigma^{-2}\ell^{-1}\right). \tag{A.34}$$

It only remains to prove that $\sqrt{n}R_n = o(S_{n-1})$ a.s. We have from (A.4) with $a = 4$ that $|X_n| = o(n^{1/4})$ a.s. which implies that $\sqrt{n}R_n = o(n)$ a.s. Hence,

$$\lim_{n\to\infty} \sqrt{n}\frac{R_n}{S_{n-1}} = 0 \qquad \text{a.s.} \tag{A.35}$$

Finally, it follows from (A.29) together with (A.34) and (A.35) that

$$\sqrt{n}\left(\widehat{\theta}_n - \theta^*\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \sigma_\theta^2\right)$$

where the asymptotic variance

$$\sigma_\theta^2 = \frac{\sigma^2}{\ell(1+\theta\rho)^2} = \frac{(1-\theta^2)(1-\theta\rho)(1-\rho^2)}{(1+\theta\rho)^3}$$

which achieves the Proof of Theorem 2.2. $\square$

A.3. Proof of Theorem 2.4

Denote by f_n the explosion coefficient associated with the locally square-integrable real martingale (M_n), given for all $n \geq 0$, by

$$f_n = \frac{X_n^2}{S_n}. \tag{A.36}$$

It clearly follows from (A.18) that f_n tends to zero almost surely. Consequently, by virtue of the quadratic strong law for martingales given by Theorem 3 of [1] or [2],

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} f_k \left(\frac{M_k^2}{S_{k-1}} \right) = \sigma^2 \qquad \text{a.s.} \qquad (A.37)$$

In addition, by summation of equation (A.15), we have for all $n \geq 1$,

$$\sum_{k=1}^{n} f_k S_{k-1} \left(\widehat{\theta}_k - \theta^* \right)^2 = \frac{1}{(1+\theta\rho)^2} \sum_{k=1}^{n} f_k \left(\frac{M_k^2}{S_{k-1}} \right) + \frac{1}{(1+\theta\rho)^2} \sum_{k=1}^{n} f_k \left(\frac{R_k^2}{S_{k-1}} \right)$$
$$+ \frac{2}{(1+\theta\rho)^2} \sum_{k=1}^{n} f_k \left(\frac{M_k R_k}{S_{k-1}} \right).$$

We already saw from (A.35) that $R_n^2 = o(S_{n-1})$ a.s. Moreover, by the elementary inequality $x \leq -\log(1-x)$ where $0 \leq x \leq 1$, we obtain that $f_n \leq -\log(1-f_n)$ which means that $f_n \leq \log S_n - \log S_{n-1}$. Thus,

$$\sum_{k=1}^{n} f_k \left(\frac{R_k^2}{S_{k-1}} \right) = O(1) + o\left(\sum_{k=1}^{n} f_k \right) = O(1) + o\left(\log S_n \right) = o\left(\log n \right) \qquad \text{a.s.}$$

Consequently, the second term of the summation is negligible compared to the first one. Furthermore, the third one is a cross-term and this ensures that it also plays a negligible role compared to the first term. Thereby,

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} f_k S_{k-1} \left(\widehat{\theta}_k - \theta^* \right)^2 = \frac{\sigma^2}{(1+\theta\rho)^2} \qquad \text{a.s.} \qquad (A.38)$$

Finally, as in the proof of Corollary 8 in [1], we deduce from (A.18) and (A.38) that

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \left(\widehat{\theta}_k - \theta^* \right)^2 = \frac{\sigma^2}{\ell(1+\theta\rho)^2} \qquad \text{a.s.}$$
$$= \frac{(1-\theta^2)(1-\theta\rho)(1-\rho^2)}{(1+\theta\rho)^3} \qquad \text{a.s.}$$

which completes the proof of the quadratic strong law (2.5). We shall now proceed to the proof of the law of iterated logarithm given by (2.6). Kolmogorov's law of iterated logarithm was extended to the martingale framework by Stout [24, 25] and a simplified version of this result may be found in Corollary 6.4.25 of [7]. In order to apply the law of iterated logarithm for martingales, it is only necessary to verify that

$$\sum_{k=1}^{+\infty} \frac{X_k^4}{k^2} < +\infty \qquad \text{a.s.} \qquad (A.39)$$

For all $n \geq 0$, denote

$$T_n = \sum_{k=1}^{n} X_k^4$$

with $T_0 = 0$. We clearly have

$$\sum_{k=1}^{+\infty} \frac{X_k^4}{k^2} = \sum_{k=1}^{+\infty} \frac{T_k - T_{k-1}}{k^2} = \sum_{k=1}^{+\infty} \left(\frac{2k+1}{k^2(k+1)^2} \right) T_k.$$

However, we already saw from (A.32) that $T_n = O(n)$ a.s. Consequently,

$$\sum_{k=1}^{+\infty} \frac{X_k^4}{k^2} = O\left(\sum_{k=1}^{+\infty} \frac{T_k}{k^3}\right) = O\left(\sum_{k=1}^{+\infty} \frac{1}{k^2}\right) = O(1) \qquad \text{a.s.}$$

which immediately implies (A.39). Then, we obtain from the law of iterated logarithm for martingales that

$$\limsup_{n\to\infty} \left(\frac{\langle M\rangle_n}{2\log\log\langle M\rangle_n}\right)^{1/2} \frac{M_n}{\langle M\rangle_n} = -\liminf_{n\to\infty} \left(\frac{\langle M\rangle_n}{2\log\log\langle M\rangle_n}\right)^{1/2} \frac{M_n}{\langle M\rangle_n}$$
$$= 1 \qquad \text{a.s.}$$

Whence, as $\langle M\rangle_n = \sigma^2 S_{n-1}$, we deduce from (A.18) that

$$\limsup_{n\to\infty} \left(\frac{n}{2\log\log n}\right)^{1/2} \frac{M_n}{S_{n-1}} = -\liminf_{n\to\infty} \left(\frac{n}{2\log\log n}\right)^{1/2} \frac{M_n}{S_{n-1}}$$
$$= \frac{\sigma}{\sqrt{\ell}} \qquad \text{a.s.} \tag{A.40}$$

Furthermore, we obviously have from (A.35) that

$$\lim_{n\to\infty} \left(\frac{n}{2\log\log n}\right)^{1/2} \frac{R_n}{S_{n-1}} = 0 \qquad \text{a.s.} \tag{A.41}$$

Finally, (2.6) follows from the conjunction of (A.15), (A.40) and (A.41), completing the Proof of Theorem 2.4.

APPENDIX B. PROOFS OF THE SERIAL CORRELATION PARAMETER RESULTS

B.1. Proof of Theorem 3.1

In order to establish the almost sure convergence of the least squares estimator $\widehat{\rho}_n$, it is necessary to start with an useful technical lemma.

Lemma B.1. *Assume that the initial values X_0 and ε_0 are square-integrable and that (V_n) is a sequence of square-integrable, independent and identically distributed random variables with zero mean and variance $\sigma^2 > 0$. Then,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=2}^{n} X_k X_{k-2} = \ell_2 \qquad \text{a.s.} \tag{B.1}$$

where the limiting value

$$\ell_2 = \frac{\sigma^2\left((\theta+\rho)^2 - \theta\rho(1+\theta\rho)\right)}{(1-\theta^2)(1-\theta\rho)(1-\rho^2)}. \tag{B.2}$$

Proof. Proceeding as in the proof of Lemma A.4, we deduce from (A.10) that for all $n \geq 2$,

$$Q_n = \sum_{k=2}^{n} X_k X_{k-2} = (\theta+\rho)P_{n-1} - \theta\rho S_{n-2} + N_n \tag{B.3}$$

where S_n, P_n and N_n are respectively given by (A.11), (A.12) and (A.22). We already saw in Appendix A that $N_n = o(n)$ a.s. Hence, it follows from (A.18) and (A.20) that

$$\lim_{n\to\infty} \frac{Q_n}{n} = (\theta+\rho)\ell_1 - \theta\rho\ell = \ell_2 \qquad \text{a.s.}$$

which achieves the proof of Lemma B.1. $\qquad\qquad\square$

Proof of Theorem 3.1. We are now in the position to prove the almost sure convergence of $\widehat{\rho}_n$ to ρ^* given by (1.3). For all $n \geq 1$, denote

$$I_n = \sum_{k=1}^{n} \widehat{\varepsilon}_k \widehat{\varepsilon}_{k-1} \qquad \text{and} \qquad J_n = \sum_{k=0}^{n} \widehat{\varepsilon}_k^2.$$

It is not hard to see that

$$I_n = P_n - \widehat{\theta}_n (S_{n-1} + Q_n) + \widehat{\theta}_n^2 P_{n-1}, \tag{B.4}$$
$$J_n = S_n - 2\widehat{\theta}_n P_n + \widehat{\theta}_n^2 S_{n-1}. \tag{B.5}$$

Consequently, it follows from convergence (2.2) together with (A.18), (A.20) and (B.1) that

$$\lim_{n \to \infty} \frac{I_n}{n} = \ell_1 - \theta^*(\ell + \ell_2) + (\theta^*)^2 \ell_1 \qquad \text{a.s.}$$
$$= \theta \rho \theta^* \ell (1 - (\theta^*)^2) \qquad \text{a.s.}$$
$$= \rho^* \ell (1 - (\theta^*)^2) \qquad \text{a.s.} \tag{B.6}$$

since $\rho^* = \theta \rho \theta^*$. By the same token,

$$\lim_{n \to \infty} \frac{J_n}{n} = \ell - 2\theta^* \ell_1 + (\theta^*)^2 \ell \qquad \text{a.s.}$$
$$= \ell (1 - (\theta^*)^2) \qquad \text{a.s.} \tag{B.7}$$

One can observe that $\ell > 0$ and $|\theta^*| < 1$, which implies that $\ell(1 - (\theta^*)^2) > 0$. Therefore, we deduce from (3.2), (B.6) and (B.7) that

$$\lim_{n \to \infty} \widehat{\rho}_n = \lim_{n \to \infty} \frac{I_n}{J_{n-1}} = \frac{\rho^* \ell (1 - (\theta^*)^2)}{\ell (1 - (\theta^*)^2)} = \rho^* \qquad \text{a.s.}$$

which completes the Proof of Theorem 3.1. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

B.2. Proof of Theorem 3.2

First of all, we already saw from (A.15) that

$$S_{n-1} \left(\widehat{\theta}_n - \theta^* \right) = \frac{M_n}{1 + \theta \rho} + \frac{R_n(\theta)}{1 + \theta \rho} \tag{B.8}$$

where $R_n(\theta) = \theta \rho X_n X_{n-1} + \rho X_0 (\varepsilon_0 - X_0)$. Our goal is to find a similar decomposition for $\widehat{\rho}_n - \rho^*$. On the one hand, we deduce from (A.14) that

$$P_n = \theta^* S_n + \frac{M_n}{1 + \theta \rho} + \frac{\xi_n^P}{1 + \theta \rho} \tag{B.9}$$

where $\xi_n^P = R_n(\theta) - (\theta + \rho) X_n^2$. On the other hand, we obtain from (B.3) and (B.9) that

$$Q_n = \left((\theta + \rho)\theta^* - \theta \rho) \right) S_n + \theta^* M_n + N_n + \xi_n^Q \tag{B.10}$$

with $\xi_n^Q = \theta^* \xi_n^P - (\theta + \rho) X_n X_{n-1} + \theta \rho (X_n^2 + X_{n-1}^2)$. Then, it follows from (B.4), (B.9) and (B.10) together with tedious but straightforward calculations that

$$I_n = \theta^* (\theta \rho - \theta^* \rho^*) S_n + \left(\frac{1 - \theta^* \rho^*}{1 + \theta \rho} \right) M_n - \theta^* N_n - \left(\widehat{\theta}_n - \theta^* \right) F_n + \xi_n^I \tag{B.11}$$

where $F_n = S_n + Q_n - \left(\widehat{\theta}_n + \theta^*\right)P_n$,

$$\xi_n^I = \widehat{\theta}_n X_n^2 - \widehat{\theta}_n^2 X_n X_{n-1} + \left(\frac{1 + (\theta^*)^2}{1 + \theta\rho}\right)\xi_n^P - \theta^*\xi_n^Q.$$

Via the same lines, we also find from (B.5), (B.9) and (B.10) that

$$J_{n-1} = \left(1 - (\theta^*)^2\right)S_n - \left(\frac{2\theta^*}{1 + \theta\rho}\right)M_n - \left(\widehat{\theta}_n - \theta^*\right)G_n + \xi_n^J \tag{B.12}$$

where $G_n = 2P_n - \left(\widehat{\theta}_n + \theta^*\right)S_n$,

$$\xi_n^J = -X_n^2 + 2\widehat{\theta}_n X_n X_{n-1} - \widehat{\theta}_n^2(X_n^2 + X_{n-1}^2) - \left(\frac{2\theta^*}{1 + \theta\rho}\right)\xi_n^P.$$

Replacing I_n and J_{n-1} by the expansions (B.11) and (B.12), we obtain from the identity $J_{n-1}(\widehat{\rho}_n - \rho^*) = I_n - \rho^* J_{n-1}$ that

$$J_{n-1}\left(\widehat{\rho}_n - \rho^*\right) = \left(\frac{1 + \theta^*\rho^*}{1 + \theta\rho}\right)M_n - \theta^* N_n - \left(\widehat{\theta}_n - \theta^*\right)H_n + \xi_n^I - \rho^*\xi_n^J$$

where $H_n = F_n - \rho^* G_n$. One can observe that the leading term depending on S_n vanishes as it should, since

$$\left(\theta^*\left(\theta\rho - \theta^*\rho^*\right)\right) - \rho^*\left(1 - (\theta^*)^2\right) = 0.$$

Consequently, we deduce from (B.8) that

$$J_{n-1}\left(\widehat{\rho}_n - \rho^*\right) = \frac{T_n M_n}{1 + \theta\rho} - \theta^* N_n + \frac{R_n(\rho)}{1 + \theta\rho} \tag{B.13}$$

where

$$T_n = 1 + \theta^*\rho^* - \frac{H_n}{S_{n-1}},$$

$$R_n(\rho) = (1 + \theta\rho)\left(\xi_n^I - \rho^*\xi_n^J\right) - \frac{R_n(\theta)H_n}{S_{n-1}}.$$

In contrast to (B.8), it was much more tricky to establish relation (B.13). We are now in the position to prove the joint asymptotic normality of $\widehat{\theta}_n$ and $\widehat{\rho}_n$. Using the same approach as in [27], it follows from (B.8) and (B.13) that

$$\sqrt{n}\begin{pmatrix} \widehat{\theta}_n - \theta^* \\ \widehat{\rho}_n - \rho^* \end{pmatrix} = \frac{1}{\sqrt{n}}A_n Z_n + B_n \tag{B.14}$$

where

$$Z_n = \begin{pmatrix} M_n \\ N_n \end{pmatrix},$$

$$A_n = \frac{n}{1 + \theta\rho}\begin{pmatrix} \dfrac{1}{S_{n-1}} & 0 \\ \dfrac{T_n}{J_{n-1}} & -\dfrac{(\theta + \rho)}{J_{n-1}} \end{pmatrix} \quad \text{and} \quad B_n = \frac{\sqrt{n}}{1 + \theta\rho}\begin{pmatrix} \dfrac{R_n(\theta)}{S_{n-1}} \\ \dfrac{R_n(\rho)}{J_{n-1}} \end{pmatrix}.$$

On the one hand, we obtain from (A.18), (A.20), (B.1) and (B.7) that

$$\lim_{n\to\infty} A_n = A \qquad \text{a.s.} \tag{B.15}$$

where A is the limiting matrix given by

$$A = \frac{1}{\ell(1+\theta\rho)(1-(\theta^*)^2)} \begin{pmatrix} 1-(\theta^*)^2 & 0 \\ \theta\rho+(\theta^*)^2 & -(\theta+\rho) \end{pmatrix}. \tag{B.16}$$

On the other hand, as in the proof of (A.35), we clearly have

$$\lim_{n\to\infty} B_n = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{a.s.} \tag{B.17}$$

Furthermore, (Z_n) is a two-dimensional real martingale [7, 14] with increasing process given, for all $n \geq 2$, by

$$\langle Z \rangle_n = \sigma^2 \begin{pmatrix} S_{n-1} & P_{n-1} \\ P_{n-1} & S_{n-2} \end{pmatrix}.$$

We deduce from (A.18) and (A.20) that

$$\lim_{n\to\infty} \frac{1}{n} \langle Z \rangle_n = L \qquad \text{a.s.} \tag{B.18}$$

where L is the positive-definite symmetric matrix given by

$$L = \sigma^2 \ell \begin{pmatrix} 1 & \theta^* \\ \theta^* & 1 \end{pmatrix}. \tag{B.19}$$

We also immediately derive from (A.32) that (Z_n) satisfies the Lindeberg condition. Therefore, we can conclude from the central limit theorem for multidimensional martingales given *e.g.* by Corollary 2.1.10 of [7] that

$$\frac{1}{\sqrt{n}} Z_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, L). \tag{B.20}$$

Finally, we find from the conjunction of (B.14), (B.15), (B.17) and (B.20) together with Slutsky's lemma that

$$\sqrt{n} \begin{pmatrix} \widehat{\theta}_n - \theta^* \\ \widehat{\rho}_n - \rho^* \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, ALA'\right).$$

One can easily check the identity $\Gamma = ALA'$ *via* (B.16) and (B.19), where Γ is given by (3.5), which achieves the Proof of Theorem 3.2.

B.3. Proof of Theorem 3.4

The proof of the quadratic strong law for $\widehat{\theta}_n$ relies on the quadratic strong law for the martingale (M_n) given by (A.37)

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} f_k \left(\frac{M_k^2}{S_{k-1}}\right) = \sigma^2 \qquad \text{a.s.}$$

which implies that

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \left(\frac{M_k}{S_{k-1}}\right)^2 = \frac{\sigma^2}{\ell} \qquad \text{a.s.} \tag{B.21}$$

In order to establish a similar result for $\widehat{\rho}_n$, we shall introduce a suitable martingale (L_n) which is a linear combination of (M_n) and (N_n). The sequence (L_n) is defined by $L_0 = 0$, $L_1 = X_0 V_1$ and, for all $n \geq 2$,

$$L_n = M_n - aN_n = L_1 + \sum_{k=2}^{n} (X_{k-1} - aX_{k-2})V_k \qquad (B.22)$$

where

$$a = \frac{\theta + \rho}{\theta\rho + (\theta^*)^2}.$$

We infer from (A.13) and (A.22) together with (B.22) that (L_n) is a locally square-integrable real martingale with predictable quadratic variation given by $\langle L \rangle_0 = 0$, $\langle L \rangle_1 = \sigma^2 X_0^2$ and for all $n \geq 2$,

$$\langle L \rangle_n = \sigma^2 \big(S_{n-1} - 2aP_{n-1} + a^2 S_{n-2} \big).$$

Moreover, we clearly deduce from (A.18) and (A.20) that

$$\lim_{n \to \infty} \frac{\langle L \rangle_n}{n} = \sigma^2 \ell b \qquad \text{a.s.} \qquad (B.23)$$

where $b = 1 - 2a\theta^* + a^2$. It also comes from a tedious calculation that

$$b = \frac{a^2(1 - \theta^2)(1 - \rho^2)c}{(\theta + \rho)^2(1 + \theta\rho)^4} \qquad (B.24)$$

where $c = (\theta + \rho)^2(1 + \theta\rho)^2 + (\theta\rho)^2(1 - \theta^2)(1 - \rho^2)$. Then, *via* the same arguments as in the proof of (B.21), we obtain from (B.23) that

$$\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \left(\frac{L_k}{S_{k-1}} \right)^2 = \frac{\sigma^2 b}{\ell} \qquad \text{a.s.} \qquad (B.25)$$

Furthermore, it follows from (B.13) that

$$J_{n-1}\big(\widehat{\rho}_n - \rho^*\big) = \left(\frac{\theta\rho + (\theta^*)^2}{1 + \theta\rho} \right) L_n + \zeta_n = \frac{\theta^* L_n}{a} + \zeta_n \qquad (B.26)$$

where

$$\zeta_n = \left(\frac{T_n - \theta\rho - (\theta^*)^2}{1 + \theta\rho} \right) M_n + \frac{R_n(\rho)}{1 + \theta\rho}.$$

We obtain from (B.21) and the almost sure convergence of T_n to $\theta\rho + (\theta^*)^2$ that

$$\sum_{k=1}^{n} \left(\frac{\zeta_k}{S_{k-1}} \right)^2 = o(\log n) \qquad \text{a.s.}$$

Consequently, (B.25) and (B.26) lead to

$$\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \left(\frac{J_{k-1}}{S_{k-1}} \right)^2 \big(\widehat{\rho}_k - \rho^*\big)^2 = \frac{\sigma^2 b (\theta^*)^2}{a^2 \ell} \qquad \text{a.s.}$$

In addition, we get from (B.7) that

$$\lim_{n \to \infty} \frac{J_n}{S_n} = 1 - (\theta^*)^2 \qquad \text{a.s.}$$

which implies that

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \left(\widehat{\rho}_k - \rho^*\right)^2 = \frac{\sigma^2 b(\theta^*)^2}{a^2 \ell (1-(\theta^*)^2)^2} \qquad \text{a.s.} \qquad (B.27)$$

However, we clearly have from (A.19) that

$$\frac{\sigma^2 (\theta^*)^2}{\ell(1-(\theta^*)^2)^2} = \frac{(\theta+\rho)^2(1-\theta\rho)(1+\theta\rho)}{(1-\theta^2)(1-\rho^2)}.$$

Finally, we can deduce from (B.24) and (B.27) that

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \left(\widehat{\rho}_k - \rho^*\right)^2 = \frac{(1-\theta\rho)c}{(1+\theta\rho)^3} = \sigma_\rho^2 \qquad \text{a.s.}$$

which completes the proof of the quadratic strong law (3.8). The law of iterated logarithm given by (3.9) is much more easy to handle. In order to make use of the law of iterated logarithm for the martingale (L_n), it is only necessary to verify that

$$\sum_{k=1}^{+\infty} \frac{(X_k - aX_{k-1})^4}{k^2} < +\infty \qquad \text{a.s.}$$

which of course follows from (A.39). Consequently, we obtain that

$$\limsup_{n\to\infty} \left(\frac{\langle L \rangle_n}{2\log\log\langle L\rangle_n}\right)^{1/2} \frac{L_n}{\langle L\rangle_n} = -\liminf_{n\to\infty} \left(\frac{\langle L \rangle_n}{2\log\log\langle L\rangle_n}\right)^{1/2} \frac{L_n}{\langle L\rangle_n}$$
$$= 1 \qquad \text{a.s.}$$

Therefore, we deduce from (B.23) that

$$\limsup_{n\to\infty} \left(\frac{n}{2\log\log n}\right)^{1/2} \frac{L_n}{\langle L\rangle_n} = -\liminf_{n\to\infty} \left(\frac{n}{2\log\log n}\right)^{1/2} \frac{L_n}{\langle L\rangle_n}$$
$$= \frac{1}{\sigma\sqrt{\ell b}} \qquad \text{a.s.}$$

Whence, the convergence

$$\lim_{n\to\infty} \frac{J_{n-1}}{\langle L\rangle_n} = \frac{1-(\theta^*)^2}{\sigma^2 b} \qquad \text{a.s.}$$

implies that

$$\limsup_{n\to\infty} \left(\frac{n}{2\log\log n}\right)^{1/2} \frac{L_n}{J_{n-1}} = -\liminf_{n\to\infty} \left(\frac{n}{2\log\log n}\right)^{1/2} \frac{L_n}{J_{n-1}}$$
$$= \frac{\sigma\sqrt{b}}{\sqrt{\ell}(1-(\theta^*)^2)} \qquad \text{a.s.} \qquad (B.28)$$

One can be convinced that the remainder term ζ_n at the right-hand side of (B.26) plays a negligible role compared to L_n. Finally, (B.26) and (B.28) ensure that

$$\limsup_{n\to\infty} \left(\frac{n}{2\log\log n}\right)^{1/2} \left(\widehat{\rho}_n - \rho^*\right) = -\liminf_{n\to\infty} \left(\frac{n}{2\log\log n}\right)^{1/2} \left(\widehat{\rho}_n - \rho^*\right)$$
$$= \frac{\sigma\sqrt{b}\theta^*}{a\sqrt{\ell}(1-(\theta^*)^2)} = \sigma_\rho \qquad \text{a.s.}$$

which ends the Proof of Theorem 3.4.

Appendix C. Proofs of the Durbin–Watson statistic results

C.1. Proof of Theorem 4.1

First of all, we establish a very useful linear relation between the Durbin–Watson statistic $\widehat{D}_n$ and the least squares estimator $\widehat{\rho}_n$, which allows us to deduce the asymptotic behavior of $\widehat{D}_n$. For all $n \geq 1$, set

$$I_n = \sum_{k=1}^n \widehat{\varepsilon}_k \widehat{\varepsilon}_{k-1}, \qquad J_n = \sum_{k=0}^n \widehat{\varepsilon}_k^2, \qquad K_n = \sum_{k=1}^n \left(\widehat{\varepsilon}_k - \widehat{\varepsilon}_{k-1}\right)^2,$$

$$f_n = \frac{\widehat{\varepsilon}_n^2}{J_n}.$$

It is not hard to see that

$$K_n = \sum_{k=1}^n \widehat{\varepsilon}_k^2 - 2 \sum_{k=1}^n \widehat{\varepsilon}_k \widehat{\varepsilon}_{k-1} + \sum_{k=1}^n \widehat{\varepsilon}_{k-1}^2 = 2\left(J_{n-1} - I_n\right) + \widehat{\varepsilon}_n^2 - \widehat{\varepsilon}_0^2.$$

Consequently, it follows from (4.1) that

$$\left(J_{n-1} + \widehat{\varepsilon}_n^2\right)\widehat{D}_n = 2\left(J_{n-1} - I_n\right) + \widehat{\varepsilon}_n^2 - \widehat{\varepsilon}_0^2. \tag{C.1}$$

Therefore, dividing both sides of (C.1) by J_{n-1}, we obtain that

$$\widehat{D}_n = 2(1 - f_n)\left(1 - \widehat{\rho}_n\right) + \xi_n \tag{C.2}$$

where

$$\xi_n = \frac{\widehat{\varepsilon}_n^2 - \widehat{\varepsilon}_0^2}{J_n}.$$

We already saw from (B.7) that

$$\lim_{n \to \infty} \frac{J_n}{n} = \ell(1 - (\theta^*)^2) \qquad \text{a.s.} \tag{C.3}$$

with $\ell(1 - (\theta^*)^2) > 0$, which implies that both f_n and ξ_n converge to zero almost surely. Hence, we deduce from (C.2) that

$$\lim_{n \to \infty} \widehat{D}_n = D^* \qquad \text{a.s.}$$

where $D^* = 2(1 - \rho^*)$, which completes the Proof of Theorem 4.1.

C.2. Proof of Theorem 4.2

We shall now prove the asymptotic normality of $\widehat{D}_n$ using (3.7). On the one hand, we clearly have from (C.2) that

$$\sqrt{n}\left(\widehat{D}_n - D^*\right) = -2(1 - f_n)\sqrt{n}\left(\widehat{\rho}_n - \rho^*\right) + 2(\rho^* - 1)\sqrt{n}f_n + \sqrt{n}\xi_n. \tag{C.4}$$

On the other hand, we deduce from (A.4) with $a = 4$ that

$$\sup_{1 \leq k \leq n} \widehat{\varepsilon}_k^2 = o(\sqrt{n}) \qquad \text{a.s.}$$

which, *via* (C.3), implies that

$$\lim_{n \to \infty} \sqrt{n}f_n = 0 \qquad \text{and} \qquad \lim_{n \to \infty} \sqrt{n}\xi_n = 0 \qquad \text{a.s.} \tag{C.5}$$

Then, it follows from (3.7), (C.4) and (C.5) together with Slutsky's lemma that

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

where $\sigma_D^2 = 4\sigma_\rho^2$, which achieves the Proof of Theorem 4.2.

C.3. Proof of Theorem 4.4

We immediately deduce from relation (C.2) that

$$\widehat{D}_n - D^* = -2(1 - f_n)\left(\widehat{\rho}_n - \rho^*\right) + \zeta_n \tag{C.6}$$

where $\zeta_n = 2(\rho^* - 1)f_n + \xi_n$. Consequently, by summation of (C.6), we obtain that for all $n \geq 1$,

$$\sum_{k=1}^{n} \left(\widehat{D}_k - D^*\right)^2 = \sum_{k=1}^{n} \left(4(1 - f_k)^2\left(\widehat{\rho}_k - \rho^*\right)^2 + \zeta_k^2 - 4(1 - f_k)\zeta_k\left(\widehat{\rho}_k - \rho^*\right)\right). \tag{C.7}$$

Since f_n goes to zero almost surely, we have

$$\sum_{k=1}^{n} \zeta_k^2 = O(1) + O\left(\sum_{k=1}^{n} f_k^2\right) = O(1) + o\left(\sum_{k=1}^{n} f_k\right) = o(\log n) \qquad \text{a.s.}$$

Hence, we infer from (3.8) and (C.7) that

$$\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \left(\widehat{D}_k - D^*\right)^2 = 4\sigma_\rho^2 = \sigma_D^2 \qquad \text{a.s.}$$

Furthermore, the law of iterated logarithm (4.6) immediately follows from (3.9) and (C.5), which completes the Proof of Theorem 4.4.

C.4. Proof of Theorem 4.6

We shall now establish the asymptotic behavior associated with our statistical procedure. It follows from the identity $\widetilde{\theta}_n = \widehat{\theta}_n + \widehat{\rho}_n - \rho_0$ and (4.13) that

$$\begin{aligned}
\widehat{\rho}_n - \widetilde{\rho}_n &= \widehat{\rho}_n - \frac{\rho_0 \widetilde{\theta}_n \left(\widetilde{\theta}_n + \rho_0\right)}{1 + \rho_0 \widetilde{\theta}_n}, \\
&= \frac{\widehat{\rho}_n + \rho_0 \widetilde{\theta}_n \widehat{\rho}_n - \rho_0 \widetilde{\theta}_n \left(\widetilde{\theta}_n + \rho_0\right)}{1 + \rho_0 \widetilde{\theta}_n}, \\
&= \frac{\widehat{\rho}_n + \rho_0 \widetilde{\theta}_n \left(\widehat{\rho}_n - \widetilde{\theta}_n - \rho_0\right)}{1 + \rho_0 \widetilde{\theta}_n}, \\
&= \frac{\widehat{\rho}_n - \rho_0 \widetilde{\theta}_n \widehat{\theta}_n}{1 + \rho_0 \widetilde{\theta}_n}.
\end{aligned}$$

Hence, if $\widehat{\gamma}_n = \widehat{\rho}_n - \theta \rho_0 \widehat{\theta}_n$ and $\widehat{\delta}_n = \widehat{\theta}_n + \widehat{\rho}_n - \theta - \rho_0$, we find that

$$\widehat{\rho}_n - \widetilde{\rho}_n = \frac{\widehat{\rho}_n - \rho_0 \widehat{\theta}_n \left(\widehat{\theta}_n + \widehat{\rho}_n - \rho_0\right)}{1 + \rho_0 \widetilde{\theta}_n} = \frac{\widehat{\gamma}_n - \rho_0 \widehat{\theta}_n \widehat{\delta}_n}{1 + \rho_0 \widetilde{\theta}_n}. \tag{C.8}$$

Denote

$$\theta_0^* = \frac{\theta + \rho_0}{1 + \theta \rho_0} \qquad \text{and} \qquad \rho_0^* = \frac{\theta \rho_0 (\theta + \rho_0)}{1 + \theta \rho_0}.$$

Since $\rho_0^* = \theta\rho_0\theta_0^*$ and $\theta_0^* + \rho_0^* = \theta + \rho_0$, we obtain that $\widehat{\gamma}_n = \widehat{\rho}_n - \rho_0^* - \theta\rho_0(\widehat{\theta}_n - \theta_0^*)$ and $\widehat{\delta}_n = \widehat{\theta}_n + \widehat{\rho}_n - \theta_0^* - \rho_0^*$. Consequently, we deduce from (C.8) that

$$\widehat{\rho}_n - \widetilde{\rho}_n = \frac{a\left(\widehat{\theta}_n - \theta_0^*\right) + b\left(\widehat{\rho}_n - \rho_0^*\right)}{1 + \rho_0\widetilde{\theta}_n} - \frac{\rho_0\widehat{\delta}_n\left(\widehat{\theta}_n - \theta_0^*\right)}{1 + \rho_0\widetilde{\theta}_n} \tag{C.9}$$

where $a = -\rho_0(\theta + \theta_0^*)$ and $b = 1 - \rho_0\theta_0^*$. On the other hand, it follows from (C.2) that

$$\begin{aligned} \widehat{D}_n - \widetilde{D}_n &= 2(1 - f_n)(1 - \widehat{\rho}_n) + \xi_n - 2(1 - \widetilde{\rho}_n), \\ &= -2(\widehat{\rho}_n - \widetilde{\rho}_n) - 2f_n(1 - \widehat{\rho}_n) + \xi_n, \\ &= -2(\widehat{\rho}_n - \widetilde{\rho}_n) + \Delta_n \end{aligned} \tag{C.10}$$

where $\Delta_n = \xi_n - 2f_n(1 - \widehat{\rho}_n)$. Therefore, (C.9) together with (C.10) imply that

$$\sqrt{n}\left(\widehat{D}_n - \widetilde{D}_n\right) = -\frac{2w'W_n}{1 + \rho_0\widetilde{\theta}_n} + \frac{2\rho_0\widehat{\delta}_n v'W_n}{1 + \rho_0\widetilde{\theta}_n} + \sqrt{n}\Delta_n \tag{C.11}$$

where v and w are the vectors of $\mathbb{R}^2$ given by $v' = (1, 0)$, $w' = (a, b)$ and

$$W_n = \sqrt{n}\begin{pmatrix} \widehat{\theta}_n - \theta_0^* \\ \widehat{\rho}_n - \rho_0^* \end{pmatrix}.$$

We already saw by (C.5) that

$$\lim_{n\to\infty} \sqrt{n}\Delta_n = 0 \qquad \text{a.s.} \tag{C.12}$$

Moreover, as $|\widehat{\delta}_n| \leq |\widehat{\theta}_n - \theta_0^*| + |\widehat{\rho}_n - \rho_0^*|$, the almost sure rates of convergence given by (2.8) and (3.11) ensure that under the null hypothesis $\mathcal{H}_0$,

$$\left|\widehat{\delta}_n v'W_n\right| = O\left(\frac{\log\log n}{\sqrt{n}}\right) \qquad \text{a.s.}$$

leading to

$$\lim_{n\to\infty} \widehat{\delta}_n v'W_n = 0 \qquad \text{a.s.} \tag{C.13}$$

Consequently, it follows from the joint asymptotic normality (3.6) together with Slutsky's lemma, (C.11), (C.12) and (C.13) that under the null hypothesis $\mathcal{H}_0$,

$$\sqrt{n}\left(\widehat{D}_n - \widetilde{D}_n\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \tau^2) \tag{C.14}$$

with

$$\tau^2 = \frac{4}{(1 + \rho_0\theta)^2}w'\Gamma w$$

where the covariance matrix Γ is given by (3.5). We recall from Remark 3.3 that Γ is invertible as soon as $\theta \neq -\rho$, which implies that $\tau^2 > 0$. In addition, we obtain from (4.14) and (4.15) that

$$\lim_{n\to\infty} \widehat{\tau}_n^2 = \tau^2 \qquad \text{a.s.} \tag{C.15}$$

Finally, we deduce from (C.14), (C.15) and once again from Slutsky's lemma that under the null hypothesis $\mathcal{H}_0$,

$$\frac{\sqrt{n}}{\widehat{\tau}_n}\left(\widehat{D}_n - \widetilde{D}_n\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$$

which obviously implies (4.16). It remains to show that under the alternative hypothesis $\mathcal{H}_1$, our test statistic goes almost surely to infinity. We already saw from (3.3) that

$$\lim_{n\to\infty} \widehat{\rho}_n = \frac{\theta\rho(\theta+\rho)}{1+\theta\rho} \qquad \text{a.s.} \tag{C.16}$$

Moreover, as $\widetilde{\theta}_n$ converges almost surely to $\theta + \rho - \rho_0$, we obtain that

$$\lim_{n\to\infty} \widetilde{\rho}_n = \frac{\rho_0(\theta+\rho)(\theta+\rho-\rho_0)}{1+\rho_0(\theta+\rho-\rho_0)} \qquad \text{a.s.} \tag{C.17}$$

Hence, it follows from (C.16) and (C.17) that

$$\lim_{n\to\infty} \left(\widehat{\rho}_n - \widetilde{\rho}_n\right) = \frac{(\theta+\rho)(\theta-\rho_0)(\rho-\rho_0)}{(1+\theta\rho)(1+\rho_0(\theta+\rho-\rho_0))} \qquad \text{a.s.} \tag{C.18}$$

Under the alternative hypothesis, this limit is equal to zero if and only if $\theta = \rho_0$ or $\theta = -\rho$. However, these particular cases are already excluded from the study of $\mathcal{H}_1$. Consequently, under $\mathcal{H}_1$, we deduce from (C.18) that

$$\lim_{n\to\infty} n\left(\widehat{\rho}_n - \widetilde{\rho}_n\right)^2 = +\infty \qquad \text{a.s.}$$

which, *via* (C.10), clearly leads to (4.17), completing the Proof of Theorem 4.6.

References

[1] B. Bercu, On the convergence of moments in the almost sure central limit theorem for martingales with statistical applications. *Stoch. Process. Appl.* **11** (2004) 157–173.

[2] B. Bercu, P. Cenac and G. Fayolle, On the almost sure central limit theorem for vector martingales: convergence of moments and statistical applications. *J. Appl. Probab.* **46** (2009) 151–169.

[3] V. Bitseki Penda, H. Djellout and F. Proïa, Moderate deviations for the Durbin–Watson statistic related to the first-order autoregressive process. Submitted for publication, `arXiv:1201.3579` (2012).

[4] G. Box and G. Ljung, On a measure of a lack of fit in time series models. *Biometrika* **65** (1978) 297–303.

[5] G. Box and D. Pierce, Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Amer. Statist. Assn. J.* **65** (1970) 1509–1526.

[6] T. Breusch, Testing for autocorrelation in dynamic linear models. *Austral. Econ. Papers.* **17** (1978) 334–355.

[7] M. Duflo, Random iterative models, *Appl. Math.,* vol. 34. Springer-Verlag, Berlin (1997).

[8] J. Durbin, Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables. *Econometrica* **38** (1970) 410–421.

[9] J. Durbin, Approximate distributions of student's t-statistics for autoregressive coefficients calculated from regression residuals. *J. Appl. Probab.* **23A** (1986) 173–185.

[10] J. Durbin and G.S. Watson, Testing for serial correlation in least squares regression I. *Biometrika* **37** (1950) 409–428.

[11] J. Durbin and G.S. Watson, Testing for serial correlation in least squares regression II. *Biometrika* **38** (1951) 159–178.

[12] J. Durbin and G.S. Watson, Testing for serial correlation in least squares regession III. *Biometrika* **58** (1971) 1–19.

[13] L. Godfrey, Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica* **46** (1978) 1293–1302.

[14] P. Hall and C.C. Heyde, *Martingale limit theory and its application*, *Probability and Mathematical Statistics*. Academic Press Inc., New York (1980).

[15] B.A. Inder, Finite-sample power of tests for autocorrelation in models containing lagged dependent variables. *Econom. Lett.* **14** (1984) 179–185.

[16] B.A. Inder, An approximation to the null distribution of the Durbin–Watson statistic in models containing lagged dependent variables. *Econom. Theory* **2** (1986) 413–428.

[17] M.L. King and P.X. Wu, Small-disturbance asymptotics and the Durbin–Watson and related tests in the dynamic regression model. *J. Econometrics* **47** (1991) 145–152.

[18] G.S. Maddala and A.S. Rao, Tests for serial correlation in regression models with lagged dependent variables and serially correlated errors. *Econometrica* **41** (1973) 761–774.

[19] E. Malinvaud, Estimation et prévision dans les modèles économiques autorégressifs. *Review of the International Institute of Statistics* **29** (1961) 1–32.

[20] M. Nerlove and K.F. Wallis, Use of the Durbin–Watson statistic in inappropriate situations. *Econometrica* **34** (1966) 235–238.

[21] S.B. Park, On the small-sample power of Durbin's *h*-test. *J. Amer. Stat. Assoc.* **70** (1975) 60–63.

[22] F. Proïa, A new statistical procedure for testing the presence of a significative correlation in the residuals of stable autoregressive processes. Submitted for publication, `arXiv:1203.1871` (2012).

[23] T. Stocker, On the asymptotic bias of OLS in dynamic regression models with autocorrelated errors. *Statist. Papers* **48** (2007) 81–93.

[24] W.F. Stout, A martingale analogue of Kolmogorov's law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **15** (1970) 279–290.

[25] W.F. Stout, Almost sure convergence, *Probab. Math. Statist.* Academic Press, New York, London **24** (1974).

[26] J.A. Tillman, The power of the Durbin–Watson test. *Econometrica* **43** (1975) 959–974.

[27] C. Wei and J. Winnicki, Estimation on the means in the branching process with immigration. *Ann. Statist.* **18** (1990) 1757–1773.