Electronic Journal of Statistics
Vol. 17 (2023) 1160–1239
ISSN: 1935-7524
https://doi.org/10.1214/23-EJS2125

Diagnostic checking in FARIMA models with uncorrelated but non-independent error terms

Yacouba Boubacar Maïnassara, Youssef Esstafa and Bruno Saussereau

Université de Franche-Comté, Laboratoire de mathématiques de Besançon, UMR CNRS 6623, 16 route de Gray, 25030 Besançon, France
e-mail: yacouba.boubacar_mainassara@univ-fcomte.fr
youssef.esstafa@univ-fcomte.fr; bruno.saussereau@univ-fcomte.fr

Abstract: This work considers the problem of modified Portmanteau tests for testing the adequacy of FARIMA models under the assumption that the errors are uncorrelated but not necessarily independent (*i.e.* weak FARIMA). We first study the joint distribution of the least squares estimator and the noise empirical autocovariances. We then derive the asymptotic distribution of residual empirical autocovariances and autocorrelations. We deduce the asymptotic distribution of the Ljung-Box (or Box-Pierce) modified Portmanteau statistics for weak FARIMA models. We also propose another method based on a self-normalization approach to test the adequacy of FARIMA models. Finally some simulation studies are presented to corroborate our theoretical work. An application to the Standard & Poor's 500 and Nikkei returns also illustrates the practical relevance of our theoretical results.

MSC2020 subject classifications: Primary 62M10, 62F03, 62F05; secondary 91B84, 62P05.
Keywords and phrases: Nonlinear processes, long-memory processes, weak FARIMA models, least squares estimator, Box-Pierce and Ljung-Box Portmanteau tests, residual autocorrelations, self-normalization.

Received September 2021.

Contents

1. Introduction

To model the long memory phenomenon, a widely used model is the fractional autoregressive integrated moving average (FARIMA, for short) model (see for instance [25], [18], [11], [29], [3] and [42] among others). This model plays an important role in many scientific disciplines and applied fields such as hydrology, climatology, economics, finance, to name a few.

We consider a centered stationary process $X := (X_t)_{t\in\mathbb{Z}}$ which satisfies a FARIMA(p, d_0, q) representation of the form

$$a(L)(1-L)^{d_0}X_t = b(L)\epsilon_t, \tag{1}$$

where d_0 is the long memory parameter, L stands for the back-shift operator and $a(L) = 1-\sum_{i=1}^{p} a_iL^i$, respectively $b(L) = 1-\sum_{i=1}^{q} b_iL^i$, is the autoregressive, respectively the moving average, operator. These operators represent the short memory part of the model (by convention $a_0 = b_0 = 1$). In the standard situation $\epsilon := (\epsilon_t)_{t\in\mathbb{Z}}$ is assumed to be a sequence of independent and identically distributed (iid for short) random variables with zero mean and with a common variance. In this standard framework, ϵ is said to be a *strong white noise* and the representation (1) is called a strong FARIMA(p, d_0, q) process. In contrast with this previous definition, the representation (1) is said to be a weak FARIMA(p, d_0, q) if the noise process ϵ is a *weak white noise*, that is, if it satisfies

(A0): $\mathbb{E}(\epsilon_t) = 0$, $\mathrm{Var}\,(\epsilon_t) = \sigma_\epsilon^2$ and $\mathrm{Cov}\,(\epsilon_t, \epsilon_{t-h}) = 0$ for all $t \in \mathbb{Z}$ and all $h \neq 0$.

A strong white noise is obviously a weak white noise because independence entails uncorrelatedness. Of course the converse is not true. The strong FARIMA model was introduced by [29]. The particular strong FARIMA$(0, d_0, 0)$ process was discussed by [25]. To ensure the stationarity and the invertibility of the model defined by (1), we assume that $0 < d_0 < 1/2$ and all roots of $a(z)b(z) = 0$ are outside the unit disk (see [25] and [29] for details). It is also assumed that $a(z)$ and $b(z)$ have no common factors in order to insure unique identifiability of the parameters.

The validity of the different steps of the traditional methodology of Box and Jenkins (identification, estimation and validation) depends on the noise properties. After estimating the FARIMA process, the next important step in the modeling consists in checking if the estimated model fits satisfactorily the data. Thus, under the null hypothesis that the model has been correctly identified, the residuals $(\hat{\epsilon}_t)$ are approximately a white noise. This adequacy checking step allows to validate or invalidate the choice of the orders p and q. The choice of p and q is particularly important because the number of parameters $(p+q+1)$ quickly increases with p and q, which entails statistical difficulties. In particular, the selection of too large orders p and q may introduce terms that are not necessarily relevant in the model. Conversely, the selection of too small orders p and q causes loss of some information, that can be detected by the correlation of the residuals.

Thus it is important to check the validity of a FARIMA(p, d_0, q) model, for given orders p and q. Based on the residual empirical autocorrelation, [8] have proposed a goodness-of-fit test, the so-called Portmanteau test, for strong ARMA models. The intuition behind these Portmanteau tests is that if a given time series model with iid innovation is appropriate for the data at hand, the autocorrelations of the residuals $\hat{\epsilon}_t$ should be close to zero, which is the theoretical value of the autocorrelations of ϵ_t (see Assumption **(A0)** below). A modification of the test of [8] has been proposed by [36] which is nowadays one of the most popular diagnostic checking tools in strong ARMA modeling of time series. A modified Portmanteau test statistic was proposed by [33] for checking the overall significance of the residual autocorrelations of a strong FARIMA(p, d_0, q) model. All these above test statistics have been obtained under the iid assumption on the noise and they may be invalid when the series is uncorrelated but dependent (see [43], [38], [39], [19], [53], [7], [52], to name a few).

As mentioned above, the works on the Portmanteau statistic are generally performed under the assumption that the errors ϵ_t are independent (see for instance [33]). This independence assumption is often considered too restrictive by practitioners. It precludes conditional heteroscedasticity and/or other forms of nonlinearity (see [21] for a review on weak univariate ARMA models) which can not be generated by FARIMA models with iid noises.[1] Relaxing this inde-

[1] To cite few examples of nonlinear processes, let us mention: the generalized autoregressive conditional heteroscedastic (GARCH) model (see [24]), the self-exciting threshold autoregressive (SETAR), the smooth transition autoregressive (STAR), the exponential autoregressive (EXPAR), the bilinear, the random coefficient autoregressive (RCA), the functional autoregressive (FAR) (see [49] and [17], for references on these nonlinear time series models).

pendence assumption allows to cover linear representations of general nonlinear processes and to extend the range of application of the FARIMA models.

This paper is devoted to the problem of the validation step of weak FARIMA processes. For the asymptotic theory of weak FARIMA model validation, recently [46] studied the diagnostic checking for long memory time series models with nonparametric conditionally heteroscedastic martingale difference errors. This author also generalized the test statistic based on the kernel-based spectral proposed by [28] under weak assumptions on the innovation process. Note also that [35] have studied the [8] type test for FARIMA-GARCH models by assuming a parametric form for the GARCH model.

To our knowledge, it does not exist any diagnostic checking methodology for FARIMA models when the (possibly dependent) error is subject to unknown conditional heteroscedasticity. We think that this is due to the difficulty that arises when one has to estimate the asymptotic covariance matrix of the parameter estimates. In our paper, thanks to the asymptotic results obtained by [6], we are able to extend for weak FARIMA models the diagnostic checking methodology proposed by [19] as well as the self-normalized approach proposed by [7].

The paper is organized as follows. In Section 2, we recall the results on the least squares estimator asymptotic distribution of weak FARIMA models obtained by [6]. In Section 3, a modified version of the Portmanteau test is proposed thanks to the investigation of the asymptotic distribution of the residual autocorrelations. Our first main result is stated in Theorem 2. The second main result of this section is obtained in Theorem 7 by means of a self-normalized approach. Some numerical illustrations are gathered in Section 4. They corroborate our theoretical work. An application to the Standard & Poor's 500 and Nikkei returns also illustrate the practical relevance of our theoretical results. All our proofs are given in Section 5.

2. Assumptions and estimation procedure

In this section, we recall the results on the least squares estimator asymptotic distribution of weak FARIMA models obtained by [6] in order to have a self-contained paper.

Let Θ^* be the parameter space

$$\Theta^* = \Big\{(\theta_1, \theta_2, \dots, \theta_{p+q}) \in \mathbb{R}^{p+q}, \text{ where } a_\theta(z) = 1 - \sum_{i=1}^{p} \theta_i z^i, \text{ and } b_\theta(z) = 1 - \sum_{j=1}^{q} \theta_{p+j} z^j \text{ have all their zeros outside the unit disk}\Big\}.$$

Denote by Θ the Cartesian product $\Theta^* \times (0, 1/2)$. The unknown parameter of interest $\theta_0 = (a_1, a_2, \dots, a_p, b_1, b_2, \dots, b_q, d_0)'$ is supposed to belong to the parameter space Θ.

The fractional difference operator $(1-L)^{d_0}$ is defined, using the generalized binomial series, by

$$(1-L)^{d_0} = \sum_{j\geq 0} \alpha_j(d_0)L^j,$$

where for all $j \geq 0$, $\alpha_j(d_0) = \Gamma(j-d_0)/\left\{\Gamma(j+1)\Gamma(-d_0)\right\}$ and $\Gamma(\cdot)$ is the Gamma function. Using the Stirling formula we obtain that for large j, $\alpha_j(d_0) \sim j^{-d_0-1}/\Gamma(-d_0)$ (one refers to [3] for further details).

For all $\theta \in \Theta$ we define $(\epsilon_t(\theta))_{t\in\mathbb{Z}}$ as the second order stationary process which is the solution of

$$\epsilon_t(\theta) = \sum_{j\geq 0} \alpha_j(d)X_{t-j} - \sum_{i=1}^{p} \theta_i \sum_{j\geq 0} \alpha_j(d)X_{t-i-j} + \sum_{j=1}^{q} \theta_{p+j}\epsilon_{t-j}(\theta). \tag{2}$$

Observe that, for all $t \in \mathbb{Z}$, $\epsilon_t(\theta_0) = \epsilon_t$ a.s. Given a realization $X_1, \dots, X_n$ of length n, $\epsilon_t(\theta)$ can be approximated, for $0 < t \leq n$, by $\tilde{\epsilon}_t(\theta)$ defined recursively by

$$\tilde{\epsilon}_t(\theta) = \sum_{j=0}^{t-1} \alpha_j(d)X_{t-j} - \sum_{i=1}^{p} \theta_i \sum_{j=0}^{t-i-1} \alpha_j(d)X_{t-i-j} + \sum_{j=1}^{q} \theta_{p+j}\tilde{\epsilon}_{t-j}(\theta), \tag{3}$$

with $\tilde{\epsilon}_t(\theta) = X_t = 0$ if $t \leq 0$.

As shown in Proposition 8 (see Subsection 5.1), these initial values are asymptotically negligible and in particular it holds that $\epsilon_t(\theta) - \tilde{\epsilon}_t(\theta) \to 0$ almost-surely as $t \to \infty$ uniformly in θ. Let Θ_δ^* denotes the compact set

$$\Theta_\delta^* = \Big\{\theta \in \mathbb{R}^{p+q};\ \text{the roots of the polynomials } a_\theta(z) \text{ and } b_\theta(z) \text{ have modulus } \geq 1+\delta\Big\}.$$

We define the set Θ_δ as the Cartesian product of Θ_δ^* by $[d_1, d_2]$, *i.e.* $\Theta_\delta = \Theta_\delta^* \times [d_1, d_2]$, where $[d_1, d_2] \subset (0, 1/2)$ and where δ is a positive constant chosen such that θ_0 belongs to Θ_δ.

The least squares estimator is defined, almost-surely, by

$$\hat{\theta}_n = \underset{\theta\in\Theta_\delta}{\operatorname{argmin}}\ Q_n(\theta), \text{ where } Q_n(\theta) = \frac{1}{n}\sum_{t=1}^{n} \tilde{\epsilon}_t^2(\theta). \tag{4}$$

The asymptotic properties of this estimator are well known when the innovation process $(\epsilon_t)_{t\in\mathbb{Z}}$ is a strong or a semi-strong white noise (see for instance [30], [41] and [10] who have considered the problem of conditional sum-of squares estimation with d_0 allowed to lie in an arbitrary large compact set). To ensure the consistency of the least squares estimator in our context, we assume as in [6] that the parametrization satisfies the following condition.

(A1): The process $(\epsilon_t)_{t\in\mathbb{Z}}$ is strictly stationary and ergodic.

The consistency of the estimator is obtained under the assumptions **(A0)** and **(A1)**. Additional assumptions are required in order to establish the asymptotic normality of the least squares estimator. We assume that θ_0 is not on the boundary of the parameter space Θ_δ.

(A2): We have $\theta_0 \in \overset{\circ}{\Theta}_\delta$, where $\overset{\circ}{\Theta}_\delta$ denotes the interior of Θ_δ.

The stationary process ϵ is not supposed to be an independent sequence. So one needs to control its dependency by means of its strong mixing coefficients $\{\alpha_\epsilon(h)\}_{h\geq 0}$ defined by

$$\alpha_\epsilon\left(h\right) = \sup_{A\in\mathcal{F}_{-\infty}^t, B\in\mathcal{F}_{t+h}^{\infty}} \left|\mathbb{P}\left(A\cap B\right) - \mathbb{P}(A)\mathbb{P}(B)\right|,$$

where $\mathcal{F}_{-\infty}^t = \sigma(\epsilon_u, u \leq t)$ and $\mathcal{F}_{t+h}^{\infty} = \sigma(\epsilon_u, u \geq t+h)$.

We shall need an integrability assumption on the moment of the noise ϵ and a summability condition on the strong mixing coefficients $(\alpha_\epsilon(h))_{h\geq 0}$.

(A3): There exists an integer τ such that for some $\nu \in]0,1]$, $\mathbb{E}|\epsilon_t|^{\tau+\nu} < \infty$ and $\sum_{h=0}^{\infty}(h+1)^{k-2}\left\{\alpha_\epsilon(h)\right\}^{\frac{\nu}{k+\nu}} < \infty$ for $k = 1,\ldots,\tau$.

Note that **(A3)** implies the following weak assumption on the joint cumulants of the innovation process ϵ (see [16], for more details).

(A3'): There exists an integer $\tau \geq 2$ such that

$$C_\tau := \sum_{i_1,\ldots,i_{\tau-1}\in\mathbb{Z}} \left|\mathrm{cum}(\epsilon_0, \epsilon_{i_1}, \ldots, \epsilon_{i_{\tau-1}})\right| < \infty\ .$$

In the above expression, $\mathrm{cum}(\epsilon_0, \epsilon_{i_1}, \ldots, \epsilon_{i_{\tau-1}})$ denotes the τ-th order joint cumulant of the stationary process ϵ. Due to the fact that the ϵ_t's are centered, we notice that for fixed (i,j,k)

$$\mathrm{cum}(\epsilon_0, \epsilon_i, \epsilon_j, \epsilon_k) = \mathbb{E}\left[\epsilon_0\epsilon_i\epsilon_j\epsilon_k\right] - \mathbb{E}\left[\epsilon_0\epsilon_i\right]\mathbb{E}\left[\epsilon_j\epsilon_k\right] - \mathbb{E}\left[\epsilon_0\epsilon_j\right]\mathbb{E}\left[\epsilon_i\epsilon_k\right] - \mathbb{E}\left[\epsilon_0\epsilon_k\right]\mathbb{E}\left[\epsilon_i\epsilon_j\right].$$

Assumption **(A3)** is a usual technical hypothesis which is useful when one proves the asymptotic normality (see [20] for example). Let us notice however that we impose a stronger convergence speed for the mixing coefficients than in the works on weak ARMA processes. This is due to the fact that the coefficients in the infinite AR or MA representation of $\epsilon_t(\theta)$ have no more exponential decay because of the fractional operator (see Subsection 6.1 in [6] for details and comments).

As mentioned before, Hypothesis **(A3)** implies **(A3')** which is also a technical assumption usually used in the fractional ARIMA processes framework (see for instance [44, 46]) or even in an ARMA context (see [22, 53]).

For all $t \in \mathbb{Z}$, let

$$H_t(\theta) = 2\epsilon_t(\theta)\frac{\partial}{\partial\theta}\epsilon_t(\theta) = \left(2\epsilon_t(\theta)\frac{\partial}{\partial\theta_1}\epsilon_t(\theta), \ldots, 2\epsilon_t(\theta)\frac{\partial}{\partial\theta_{p+q+1}}\epsilon_t(\theta)\right)'.$$

Remind that the sequence $(\epsilon_t(\theta))_{t\in\mathbb{Z}}$ is given by (2). Under the assumptions **(A0)**, **(A1)**, **(A2)** and **(A3)** with $\tau = 4$, [6] showed that $\hat{\theta}_n \to \theta_0$ in probability as $n \to \infty$ and $\sqrt{n}(\hat{\theta}_n - \theta_0)$ is asymptotically normal with mean 0 and covariance matrix $\Sigma_{\hat{\theta}} := J^{-1}IJ^{-1}$, where $J = J(\theta_0)$ and $I = I(\theta_0)$, with

$$I(\theta) = \sum_{h=-\infty}^{+\infty} \operatorname{Cov}\left(H_t(\theta), H_{t-h}(\theta)\right) \text{ and } J(\theta) = 2\mathbb{E}\left(\frac{\partial}{\partial\theta}\epsilon_t(\theta)\frac{\partial}{\partial\theta'}\epsilon_t(\theta)\right) \text{ a.s.}$$

3. Diagnostic checking in weak FARIMA models

After the estimation phase, the next important step consists in checking if the estimated model fits satisfactorily the data. In this section we derive the limiting distribution of the residual autocorrelations and that of the Portmanteau statistics (based on the standard and the self-normalized approaches) in the framework of weak FARIMA models.

For $t \geq 1$, let $\hat{e}_t = \tilde{\epsilon}_t(\hat{\theta}_n)$ be the least squares residuals. By (3) we notice that $\hat{e}_t = 0$ for $t \leq 0$ and $t > n$. By (1) it holds that

$$\hat{e}_t = \sum_{j=0}^{t-1} \alpha_j(\hat{d})\hat{X}_{t-j} - \sum_{i=1}^{p} \hat{\theta}_i \sum_{j=0}^{t-i-1} \alpha_j(\hat{d})\hat{X}_{t-i-j} + \sum_{j=1}^{q} \hat{\theta}_{p+j}\hat{e}_{t-j},$$

for $t = 1, \ldots, n$, with $\hat{X}_t = 0$ for $t \leq 0$ and $\hat{X}_t = X_t$ for $t \geq 1$.

For a fixed integer $m \geq 1$ consider the vector of residual autocovariances

$$\hat{\gamma}_m = (\hat{\gamma}(1), \ldots, \hat{\gamma}(m))' \text{ where } \hat{\gamma}(h) = \frac{1}{n}\sum_{t=h+1}^{n} \hat{e}_t\,\hat{e}_{t-h} \text{ for } 0 \leq h < n.$$

In the sequel we will also need the vector of the first m sample autocorrelations

$$\hat{\rho}_m = (\hat{\rho}(1), \ldots, \hat{\rho}(m))' \text{ where } \hat{\rho}(h) = \hat{\gamma}(h)/\hat{\gamma}(0).$$

Since the papers by [8] and [36], Portmanteau tests have been popular diagnostic checking tools in the ARMA modeling of time series. Based on the residual empirical autocorrelations, their test statistics are defined respectively by

$$Q_m^{\text{BP}} = n\sum_{h=1}^{m} \hat{\rho}^2(h) \text{ and } Q_m^{\text{LB}} = n(n+2)\sum_{h=1}^{m} \frac{\hat{\rho}^2(h)}{n-h}. \tag{5}$$

These statistics are usually used to answer the question whether $(X_t)_{t\in\mathbb{Z}}$ satisfies a FARIMA(p, d_0, q) representation or $(X_t)_{t\in\mathbb{Z}}$ admits a FARIMA(p', d_0, q') representation with $p' > p$ or $q' > q$. We will actually test the null hypothesis

(H0): for all $k \in \{1, \ldots, m\}$, $\mathbb{E}(\epsilon_t\epsilon_{t-k}) = 0$

against the alternative

(H1): there exists $k \in \{1, \ldots, m\}$ such that $\mathbb{E}(\epsilon_t\epsilon_{t-k}) \neq 0$.

These tests are very useful tools to check the global significance of the residual autocorrelations.

3.1. Asymptotic distribution of the residual autocorrelations

First of all, the mixing assumption **(A3)** will entail the asymptotic normality of the "empirical" autocovariances

$$\gamma_m = (\gamma(1), \dots, \gamma(m))' \text{ where } \gamma(h) = \frac{1}{n} \sum_{t=h+1}^{n} \epsilon_t \, \epsilon_{t-h} \text{ for } 0 \le h < n. \tag{6}$$

It should be noted that $\gamma(h)$ is not a computable statistic because it depends on the unobserved innovations $\epsilon_t = \epsilon_t(\theta_0)$. They are introduced as a device to facilitate future derivations. Let Ψ_m be the $m \times (p+q+1)$ matrix defined by

$$\Psi_m = \mathbb{E}\left\{ \begin{pmatrix} \epsilon_{t-1} \\ \vdots \\ \epsilon_{t-m} \end{pmatrix} \frac{\partial \epsilon_t(\theta)}{\partial \theta'}\Big|_{\theta=\theta_0} \right\}. \tag{7}$$

By a Taylor expansion of $\sqrt{n}\hat{\gamma}_m$, one should prove that (see Section 5.3)

$$\sqrt{n}\hat{\gamma}_m = \sqrt{n}\gamma_m + \Psi_m \sqrt{n}\left(\hat{\theta}_n - \theta_0\right) + \mathrm{o}_{\mathbb{P}}(1), \tag{8}$$

where Ψ_m is given in (7). We shall also prove (see Section 5.3 again) that

$$\sqrt{n}\hat{\rho}_m = \sqrt{n}\frac{\hat{\gamma}_m}{\sigma_\epsilon^2} + \mathrm{o}_{\mathbb{P}}(1). \tag{9}$$

Thus from (9) the asymptotic distribution of the residual autocorrelations $\sqrt{n}\hat{\rho}_m$ depends on the distribution of $\hat{\gamma}_m$. In view of (8) the asymptotic distribution of the residual autocovariances $\sqrt{n}\hat{\gamma}_m$ will be obtained from the joint asymptotic behavior of $\sqrt{n}(\hat{\theta}_n' - \theta_0', \gamma_m')'$.

In view of Theorem 1 in [6] and **(A2)**, we have $\hat{\theta}_n \to \theta_0 \in \overset{\circ}{\Theta}$ in probability. Thus $\partial Q_n(\hat{\theta}_n)/\partial\theta = 0$ for sufficiently large n and a Taylor expansion gives

$$\sqrt{n}\frac{\partial}{\partial\theta}O_n(\theta_0) + J(\theta_0)\sqrt{n}(\hat{\theta}_n - \theta_0) = \mathrm{o}_{\mathbb{P}}(1), \tag{10}$$

where $O_n(\theta) = n^{-1}\sum_{t=1}^{n} \epsilon_t^2(\theta)$ and the sequence $(\epsilon_t(\theta))_{t\in\mathbb{Z}}$ is given by (2). The equation (10) is proved in [6] (see the proof of Theorem 2). Consequently from (10) we have

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = -\frac{2}{\sqrt{n}}\sum_{t=1}^{n} J^{-1}(\theta_0)\epsilon_t(\theta_0)\frac{\partial\epsilon_t(\theta_0)}{\partial\theta} + \mathrm{o}_{\mathbb{P}}(1). \tag{11}$$

For integers $m, m' \ge 1$, one needs the matrix $\Gamma_{m,m'} = [\Gamma(\ell,\ell')]_{1\le\ell\le m, 1\le\ell'\le m'}$ where

$$\Gamma(\ell,\ell') = \sum_{h=-\infty}^{\infty} \mathbb{E}\left[\epsilon_t\epsilon_{t-\ell}\epsilon_{t-h}\epsilon_{t-h-\ell'}\right].$$

The existence of $\Gamma(\ell,\ell')$ will be justified in Lemma 3 of the appendix.

Proposition 1. *Under the assumptions **(A0)**, **(A1)**, **(A2)** and **(A3)** with $\tau = 4$, the random vector*

$$\sqrt{n}\left(\left(\hat{\theta}_n - \theta_0\right)', \gamma_m'\right)'$$

has a limiting centered normal distribution with covariance matrix

$$\Xi = \begin{pmatrix} \Sigma_{\hat{\theta}} & \Sigma_{\hat{\theta},\gamma_m} \\ \Sigma'_{\hat{\theta},\gamma_m} & \Gamma_{m,m} \end{pmatrix} = \sum_{h=-\infty}^{\infty} \mathbb{E}\left[U_t U'_{t-h}\right], \tag{12}$$

where from (6) *and* (11) *we have*

$$U_t = \begin{pmatrix} U_{1t} \\ U_{2t} \end{pmatrix} = \begin{pmatrix} -2J^{-1}(\theta_0)\epsilon_t(\theta_0)\frac{\partial}{\partial\theta}\epsilon_t(\theta_0) \\ (\epsilon_{t-1}, \dots, \epsilon_{t-m})'\epsilon_t \end{pmatrix}. \tag{13}$$

The proof of the proposition is given in Subsection 5.2 of the appendix.

The following theorem which is an extension of the result given in [19] provides the limit distribution of the residual autocovariances and autocorrelations of weak FARIMA models.

Theorem 2. *Under the assumptions of Proposition 1, we have*

$$\sqrt{n}\hat{\gamma}_m \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0, \Sigma_{\hat{\gamma}_m}\right) \quad \textit{where} \quad \Sigma_{\hat{\gamma}_m} = \Gamma_{m,m} + \Psi_m \Sigma_{\hat{\theta}} \Psi'_m + \Psi_m \Sigma_{\hat{\theta},\gamma_m} + \Sigma'_{\hat{\theta},\gamma_m} \Psi'_m \tag{14}$$

and

$$\sqrt{n}\hat{\rho}_m \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0, \Sigma_{\hat{\rho}_m}\right) \quad \textit{where} \quad \Sigma_{\hat{\rho}_m} = \frac{1}{\sigma_\epsilon^4}\Sigma_{\hat{\gamma}_m}. \tag{15}$$

The detailed proof of this result is postponed to Subsection 5.3 of Appendix.

Remark 1. It is clear from Theorem 2 that for a given FARIMA(p, d_0, q) model, the asymptotic distribution of the residual autocorrelations depends only on the noise distribution through the quantities $\Gamma(\ell, \ell')$ (which depends on the fourth-order structure of the noise). It is also worth noting that this asymptotic distribution depends on the asymptotic normality of the least squares estimator of the FARIMA(p, d_0, q) only through the matrix $\Sigma_{\hat{\theta}}$.

Remark 2. In the standard strong FARIMA case, *i.e.* when **(A1)** is replaced by the assumption that $(\epsilon_t)_{t\in\mathbb{Z}}$ is iid, [6] have showed in Remark 2 that $I(\theta_0) = 2\sigma_\epsilon^2 J(\theta_0)$. Thus the asymptotic covariance matrix is then reduced as $\Sigma_{\hat{\theta}} = 2\sigma_\epsilon^2 J^{-1}(\theta_0)$. In the strong case, we also have: $\Gamma(\ell, \ell') = 0$ when $\ell \neq \ell'$ and $\Gamma(\ell, \ell) = \sigma_\epsilon^4$. Thus $\Gamma_{m,m}$ is reduced as $\Gamma_{m,m} = \sigma_\epsilon^4 I_m$, where I_m denotes the $m \times m$ identity matrix. Because $\Sigma_{\hat{\theta}} = 2\sigma_\epsilon^2 J^{-1}(\theta_0)$ we obtain that

$$\Sigma_{\hat{\theta},\gamma_m} = -2\sum_{h=-\infty}^{\infty} \mathbb{E}\left\{\epsilon_t J^{-1}(\theta_0)\frac{\partial\epsilon_t(\theta_0)}{\partial\theta}\right\}\left\{\begin{pmatrix} \epsilon_{t-1-h} \\ \vdots \\ \epsilon_{t-m-h} \end{pmatrix}\epsilon_{t-h}\right\}'$$

$$= \quad -\left(2\sigma_\epsilon^2 J^{-1}(\theta_0)\right)\left\{\mathbb{E}\left[\begin{pmatrix}\epsilon_{t-1}\\ \vdots \\ \epsilon_{t-m}\end{pmatrix}\frac{\partial \epsilon_t(\theta_0)}{\partial \theta'}\right]\right\}' = -\Sigma_{\hat{\theta}}\Psi_m'.$$

We denote by $\Sigma^{\mathrm{S}}_{\hat{\gamma}_m}$ and $\Sigma^{\mathrm{S}}_{\hat{\rho}_m}$ the asymptotic variances obtained respectively in (14) and (15) for the strong FARIMA case. Thus we obtain, in the strong case, the following simpler expressions

$$\Sigma^{\mathrm{S}}_{\hat{\gamma}_m} = \sigma_\epsilon^4 I_m - 2\sigma_\epsilon^2 \Psi_m J^{-1}(\theta_0)\Psi_m' \quad \text{and} \quad \Sigma^{\mathrm{S}}_{\hat{\rho}_m} = I_m - \frac{2}{\sigma_\epsilon^2}\Psi_m J^{-1}(\theta_0)\Psi_m',$$

which are the matrices obtained by [33].

To validate a FARIMA(p, d_0, q) model, the most basic technique is to examine the autocorrelation function of the residuals. Theorem 2 can be used to obtain asymptotic significance limits for the residual autocorrelations. However, the asymptotic variance matrices $\Sigma_{\hat{\gamma}_m}$ and $\Sigma_{\hat{\rho}_m}$ depend on the unknown matrices Ξ, Ψ_m and the positive scalar σ_ϵ^2 which need to be estimated. This is the purpose of the following discussion.

3.2. Modified version of the Portmanteau test

From Theorem 2 we can deduce the following result, which gives the limiting distribution of the standard Portmanteau statistics (5) under general assumptions on the innovation process of the fitted FARIMA(p, d_0, q) model.

Theorem 3. *Under the assumptions of Theorem 2 and* ***(H0)****, the statistics* Q_m^{BP} *and* Q_m^{LB} *defined by* (5) *converge in distribution, as* $n \to \infty$*, to*

$$Z_m(\xi_m) = \sum_{k=1}^{m} \xi_{k,m} Z_k^2,$$

where $\xi_m = (\xi_{1,m}, \ldots, \xi_{m,m})'$ *is the vector of the eigenvalues of the matrix* $\Sigma_{\hat{\rho}_m} = \sigma_\epsilon^{-4}\Sigma_{\hat{\gamma}_m}$ *and* $Z_1, \ldots, Z_m$ *are independent* $\mathcal{N}(0,1)$ *variables.*

It is possible to evaluate the distribution of a quadratic form of a Gaussian vector by means of the Imhof algorithm (see [31]).

Remark 3. In view of remark 2 when m is large, $\Sigma^{\mathrm{S}}_{\hat{\rho}_m} \simeq I_m - 2\sigma_\epsilon^{-2}\Psi_m J^{-1}(\theta_0)\Psi_m'$ is close to a projection matrix. Its eigenvalues are therefore equal to 0 and 1. The number of eigenvalues equal to 1 is $\mathrm{Tr}(I_m - 2\sigma_\epsilon^{-2}\Psi_m J^{-1}(\theta_0)\Psi_m') = \mathrm{Tr}(I_{m-(p+q+1)}) = m - (p+q+1)$ and $p+q+1$ eigenvalues equal to 0, $\mathrm{Tr}(\cdot)$ denotes the trace of a matrix. Therefore we retrieve the well-known result obtained by [33]. More precisely, under **(H0)** and in the strong FARIMA case, the asymptotic distributions of the statistics Q_m^{BP} and Q_m^{LB} are approximated by a $\mathcal{X}^2_{m-(p+q+1)}$, where $m > p+q+1$ and $\mathcal{X}^2_k$ denotes the chi-squared distribution with k degrees of freedom. Theorem 3 shows that this approximation is no longer valid in the framework of weak FARIMA(p, d, q) models and that the asymptotic null distributions of the statistics Q_m^{BP} and Q_m^{LB} are more complicated.

Remark 4. When one focuses on the following alternative hypothesis

(H1): there exists $h \in \{1, \dots, m\}$ such that $\mathbb{E}(\epsilon_t \epsilon_{t-h}) \neq 0$,

this means that under **(H1)** at least one $\rho(h) = \gamma^0(h)/\sigma_\epsilon^2 \neq 0$ where $\gamma^0(h) = \mathbb{E}[\epsilon_t \epsilon_{t-h}]$. One may prove that under **(H1)**

$$\hat{\rho}_m' \hat{\rho}_m = \frac{\hat{\gamma}_m' \hat{\gamma}_m}{\sigma_\epsilon^4} + o_{\mathbb{P}}\left(\frac{1}{n}\right) \xrightarrow[n\to\infty]{\mathbb{P}} \rho_m' \rho_m = \frac{\gamma_m^{0\,\prime} \gamma_m^0}{\sigma_\epsilon^4}$$

where the vector $\rho_m = (\rho(1), \dots, \rho(m))' = \gamma_m^{0\,\prime}/\sigma_\epsilon^2$ with $\gamma_m^0 = \left(\gamma^0(1), \dots, \gamma^0(m)\right)'$. Therefore the test statistic $n\hat{\rho}_m'\hat{\rho}_m$ is consistent in detecting **(H1)**.

The proof of this remark is also postponed to Section 5.

The limit distribution $Z_m(\xi_m)$ depends on the nuisance parameter σ_ϵ^2, the matrix Ψ_m and the elements of Ξ. Therefore, the asymptotic distribution of the Portmanteau statistics (5), under weak assumptions on the noise, requires a computation of a consistent estimator of the asymptotic covariance matrix $\Sigma_{\hat{\rho}_m}$. The $m \times (p+q+1)$ matrix Ψ_m and the noise variance σ_ϵ^2 can be estimated by its empirical counterpart. Thus we may use

$$\hat{\Psi}_m = \frac{1}{n} \sum_{t=1}^{n} \left\{ (\hat{e}_{t-1}, \dots, \hat{e}_{t-m})' \frac{\partial \hat{e}_t}{\partial \theta'} \right\} \quad \text{and} \quad \hat{\sigma}_\epsilon^2 = \hat{\gamma}(0) = \frac{1}{n} \sum_{t=1}^{n} \hat{e}_t^2.$$

A consistent estimator of Ξ is obtained by means of an autoregressive spectral estimator, as in [6] (see also [4], [5] and [14], to name a few for a more comprehensive exposition of this method). The process $(U_t)_{t\in\mathbb{Z}}$ admits an AR(∞) representation (see [1]) of the form

$$\Delta(L) U_t := U_t - \sum_{i=1}^{\infty} \Delta_i U_{t-i} = v_t, \tag{16}$$

such that $\sum_{i=1}^{\infty} \|\Delta_i\| < \infty$, where $\|\cdot\|$ denotes any norm on the space of the real $(p+q+1+m) \times (p+q+1+m)$ matrices, and $\det\{\Delta(z)\} \neq 0$ if $|z| \leq 1$. In view of (12), the matrix Ξ can be interpreted as 2π times the spectral density of the stationary process $(U_t)_{t\in\mathbb{Z}} = ((U_{1t}', U_{2t}')')_{t\in\mathbb{Z}}$ evaluated at frequency 0 (see p. 459 of [9]). We then obtain that

$$\Xi = \Delta^{-1}(1) \Sigma_v \Delta'^{\,-1}(1)$$

Since U_t is unobservable, we introduce $\hat{U}_t \in \mathbb{R}^{p+q+1+m}$ obtained by replacing $\epsilon_t(\theta_0)$ by $\tilde{\epsilon}_t(\hat{\theta}_n)$ and $J(\theta_0)$ by its empirical or observable counterpart $\hat{J}_n$ in (13). Let $\hat{\Delta}_r(z) = I_{p+q+1+m} - \sum_{k=1}^{r} \hat{\Delta}_{r,k} z^k$, where $\hat{\Delta}_{r,1}, \dots, \hat{\Delta}_{r,r}$ denote the coefficients of the least squares regression of $\hat{U}_t$ on $\hat{U}_{t-1}, \dots, \hat{U}_{t-r}$. Let $\hat{v}_{r,t}$ be the residuals of this regression, and let $\hat{\Sigma}_{\hat{v}_r}$ be the empirical variance of $\hat{v}_{r,1}, \dots, \hat{v}_{r,n}$. We are now able to state Theorem 4 which is an extension of a result given in [5].

Theorem 4. *We assume* ***(A0)****,* ***(A1)****,* ***(A2)*** *and Assumption* ***(A3')*** *with* $\tau = 8$*. In addition, we assume that process* $(\epsilon_t)_{t\in\mathbb{Z}}$ *of the FARIMA*(p, d_0, q) *model* (1) *is such that the process* $(U_t)_{t\in\mathbb{Z}}$ *defined in* (13) *admits a multivariate AR*(∞) *representation* (16)*, where* $\|\Delta_i\| = \mathrm{o}(i^{-2})$ *as* $i \to \infty$*, the roots of* $\det(\Delta(z)) = 0$ *are outside the unit disk, and* $\Sigma_v = \mathrm{Var}(v_t)$ *is non-singular. Then the spectral estimator of* Ξ *satisfies*

$$\hat{\Xi}_n^{\mathrm{SP}} := \hat{\Delta}_r^{-1}(1)\hat{\Sigma}_{\hat{v}_r}\hat{\Delta}_r'^{-1}(1) \to \Xi = \Delta^{-1}(1)\Sigma_v\Delta^{-1}(1)$$

in probability when $r = r(n) \to \infty$ *and* $r^5(n)/n^{1-2(d_2-d_1)} \to 0$ *as* $n \to \infty$ *(remind that* $d_0 \in [d_1, d_2] \subset (0,1/2)$*).*

The proof of this theorem is similar to the proof of Theorem 3 in [6] and it is omitted.

We are now in a position to define the modified versions of the Box-Pierce (BP) and Ljung-Box (LB) goodness-of-fit Portmanteau tests. The standard versions of the Portmanteau tests are useful tools to detect if the orders p and q of a FARIMA(p, d_0, q) model are well chosen, provided the error terms $(\epsilon_t)_{t\in\mathbb{Z}}$ of the FARIMA(p, d_0, q) equation be a strong white noise and provided the number m of residual autocorrelations is not too small (see Remark 3). Now we define the modified versions which are aimed to detect if the orders p and q of a weak FARIMA(p, d_0, q) model are well chosen. These tests are also asymptotically valid for strong FARIMA(p, d_0, q) even for small m. The modified versions of the Portmanteau tests will be denoted by $\mathrm{BP_W}$ and $\mathrm{LB_W}$, the subscript W referring to the term weak.

Let $\hat{\Sigma}_{\hat{\rho}_m}$ be the matrix obtained by replacing Ξ by $\hat{\Xi}$ and σ_ϵ^2 by $\hat{\sigma}_\epsilon^2$ in $\Sigma_{\hat{\rho}_m}$. Denote by $\hat{\xi}_m = (\hat{\xi}_{1,m}, \dots, \hat{\xi}_{m,m})'$ the vector of the eigenvalues of $\hat{\Sigma}_{\hat{\rho}_m}$. At the asymptotic level α, it holds under the assumptions of Theorem 2 and **(H0)** that

$$\lim_{n\to\infty} \mathbb{P}\left(Q_m^{\mathrm{BP}} > S_m(1-\alpha)\right) = \lim_{n\to\infty} \mathbb{P}\left(Q_m^{\mathrm{LB}} > S_m(1-\alpha)\right) = \alpha,$$

where $S_m(1-\alpha)$ is such that $\mathbb{P}(Z_m(\hat{\xi}_m) > S_m(1-\alpha)) = \alpha$. We emphasize the fact that the proposed modified versions of the Box-Pierce and Ljung-Box statistics are more difficult to implement because their critical values have to be computed from the data while the critical values of the standard method are simply deduced from a χ^2-table. We shall evaluate the p-values

$$\mathbb{P}\left\{Z_m(\hat{\xi}_m) > Q_m^{\mathrm{BP}}\right\} \quad \text{and} \quad \mathbb{P}\left\{Z_m(\hat{\xi}_m) > Q_m^{\mathrm{LB}}\right\}, \text{with } Z_m(\hat{\xi}_m) = \sum_{i=1}^m \hat{\xi}_{i,m} Z_i^2,$$

by means of the Imhof algorithm (see [31]).

A second method avoiding the estimation of the asymptotic matrix is proposed in the next Subsection.

3.3. Self-normalized asymptotic distribution of the residual autocorrelations

In view of Theorem 3, the asymptotic distributions of the statistics defined in (5) are a mixture of chi-squared distributions, weighted by eigenvalues of the asymptotic covariance matrix $\Sigma_{\hat{\rho}_m}$ of the vector of autocorrelations obtained in Theorem 2. However, this asymptotic variance matrix depends on the unknown matrices Ξ, Ψ_m and the noise variance σ_ϵ^2. Consequently, in order to obtain a consistent estimator of the asymptotic covariance matrix $\Sigma_{\hat{\rho}_m}$ of the residual autocorrelations vector we have used an autoregressive spectral estimator of the spectral density of the stationary process $(U_t)_{t\in\mathbb{Z}}$ to get a consistent estimator of the matrix Ξ (see Theorem 4). However, this approach presents the problem of choosing the truncation parameter. Indeed this method is based on an infinite autoregressive representation of the stationary process $(U_t)_{t\in\mathbb{Z}}$ (see (16)). So the choice of the order of truncation is crucial and difficult.

In this section, we propose an alternative method where we do not estimate an asymptotic covariance matrix which is an extension to the results obtained by [7]. It is based on a self-normalization approach to construct a test-statistic which is asymptotically distribution-free under the null hypothesis. This approach has been studied by [7] in the weak ARMA case, by proposing new Portmanteau statistics. In this case the critical values are not computed from the data since they are tabulated by [37]. In some sense this method is finally closer to the standard method in which the critical values are simply deduced from a $\mathcal{X}^2$-table. The idea comes from [37] and has been already extended by [7], [32], [44], [45] and [47] to name a few in more general frameworks. See also [48] for a review on some recent developments on the inference of time series data using the self-normalized approach.

Other alternative methods that avoid the estimation of the covariance of the parameter estimates by directly eliminating the estimation effect of the test statistics can be found in [13] or [51]. [13] developed an asymptotically distribution-free transform of the sample autocorrelations of residuals in general parametric linear time-series models and showed that the proposed Box-Pierce-type test statistic based on the transformed autocorrelation is not affected by the estimation effect. [51] proposed an asymptotic simultaneous distribution-free transform of the sample autocorrelations of standardized residuals and their squares, which extended the approach developed by [13] to the conditional mean and variance models diagnosis.

We denote by Λ the block matrix of $\mathbb{R}^{m\times(p+q+1+m)}$ defined by $\Lambda = (\Psi_m|I_m)$. In view of (8) and (11) we deduce that

$$\sqrt{n}\hat{\gamma}_m = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\Lambda U_t + \mathrm{o}_{\mathbb{P}}(1).$$

At this stage, we do not rely on the classical method that would consist in estimating the asymptotic covariance matrix Ξ. We rather try to apply Lemma 1 in [37]. So we need to check that a functional central limit theorem holds for the process $U := (U_t)_{t\geq 1}$. For that sake, we define the normalization matrix C_m of $\mathbb{R}^{m\times m}$ by

$$C_m = \frac{1}{n^2}\sum_{t=1}^{n} S_t S_t' \text{ where } S_t = \sum_{j=1}^{t}\left(\Lambda U_j - \gamma_m\right).$$

To ensure the invertibility of the normalization matrix C_m (it is the result stated in the next proposition), we need the following technical assumption on the distribution of ϵ_t.

(A4): The process $(\epsilon_t)_{t\in\mathbb{Z}}$ has a positive density on some neighbourhood of zero.

Proposition 5. *Under the assumptions of Theorem 2 and* ***(A4)****, the matrix* C_m *is almost surely non singular.*

The proof of this proposition is given in Subsection 5.5 of the appendix.

Let $(B_K(r))_{r\geq 0}$ be a K-dimensional Brownian motion starting from 0. For $K \geq 1$, we denote by $\mathcal{U}_K$ the random variable defined by:

$$\mathcal{U}_K = B_K'(1)V_K^{-1}B_K(1), \tag{17}$$

where

$$V_K = \int_0^1 \left(B_K(r) - rB_K(1)\right)\left(B_K(r) - rB_K(1)\right)' \mathrm{dr}. \tag{18}$$

The critical values of $\mathcal{U}_K$ have been tabulated by [37].

The following theorem states the asymptotic distributions of the sample autocovariances and autocorrelations.

Theorem 6. *Under the assumptions of Theorem 2,* ***(A4)*** *and under the null hypothesis* ***(H0)*** *we have*

$$n\hat{\gamma}_m' C_m^{-1}\hat{\gamma}_m \xrightarrow[n\to\infty]{in\ law} \mathcal{U}_m \ \text{and}\ n\sigma_\epsilon^4\hat{\rho}_m' C_m^{-1}\hat{\rho}_m \xrightarrow[n\to\infty]{in\ law} \mathcal{U}_m.$$

The proof of this theorem is given in Subsection 5.6 of Appendix.

Of course, the above theorem is useless for practical purpose because the normalization matrix C_m and the nuisance parameter σ_ϵ^2 are not observable. This gap will be fixed below (see Theorem 7) when one replaces the matrix C_m and the scalar σ_ϵ^2 by their empirical or observable counterparts. Then we denote

$$\hat{C}_m = \frac{1}{n^2}\sum_{t=1}^{n}\hat{S}_t\hat{S}_t' \text{ where } \hat{S}_t = \sum_{j=1}^{t}\left(\hat{\Lambda}\hat{U}_j - \hat{\gamma}_m\right),$$

with $\hat{\Lambda} = (\hat{\Psi}_m|I_m)$ and where $\hat{U}_t$ and $\hat{\sigma}_\epsilon^2$ are defined in Subsection 3.2.

The above quantities are all observable and the following result is the applicable counterpart of Theorem 6.

Theorem 7. *Under the assumptions of Theorem 6, we have*

$$n\hat{\gamma}_m'\hat{C}_m^{-1}\hat{\gamma}_m \xrightarrow[n\to\infty]{in\ law} \mathcal{U}_m \ \text{and}\ Q_m^{\text{SN}} = n\hat{\sigma}_\epsilon^4\hat{\rho}_m'\hat{C}_m^{-1}\hat{\rho}_m \xrightarrow[n\to\infty]{in\ law} \mathcal{U}_m.$$

The proof of this result is postponed in Subsection 5.7 of Appendix.

Based on the above result, we propose a modified version of the Ljung-Box statistic when one uses the statistic

$$\tilde{Q}_m^{\text{SN}} = n\hat{\sigma}_\epsilon^4\hat{\rho}_m' D_{n,m}^{1/2}\hat{C}_m^{-1}D_{n,m}^{1/2}\hat{\rho}_m,$$

where $D_{n,m} \in \mathbb{R}^{m\times m}$ is diagonal with $(n+2)/(n-1), \dots, (n+2)/(n-m)$ as diagonal terms. These modified versions of the Portmanteau tests will be denoted by $\mathrm{BP}_{\textsc{sn}}$ and $\mathrm{LB}_{\textsc{sn}}$, the subscript SN referring to the term self-normalized.

4. Numerical illustrations

In this section, by means of Monte Carlo experiments, we investigate the finite sample properties of the asymptotic results that we introduced in this work. The numerical illustrations of this section are made with the open source statistical software R (see http://cran.r-project.org/).

4.1. Simulation studies and empirical sizes

We study numerically the behavior of the least squares estimator for FARIMA models of the form

$$(1-L)^{d_0}(X_t - aX_{t-1}) = \epsilon_t - b\epsilon_{t-1}, \tag{19}$$

where the unknown parameter is $\theta_0 = (a, b, d_0)$. First we assume that in (19) the innovation process $(\epsilon_t)_{t\in\mathbb{Z}}$ is an iid centered Gaussian process with common variance 1 which corresponds to the strong FARIMA case. For the weak FARIMA case, we consider that in (19) the innovation process $(\epsilon_t)_{t\in\mathbb{Z}}$ follows firstly a GARCH$(1,1)$ process given by the model

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \end{cases} \tag{20}$$

with $\omega > 0$, $\alpha_1 \geq 0$ and where $(\eta_t)_{t\in\mathbb{Z}}$ is a sequence of iid centered Gaussian random variables with variance 1. Secondly we consider that in (19) a noise defined by

$$\epsilon_t = \eta_t^2 \eta_{t-1}. \tag{21}$$

The example (21) is an extension of a noise process in [43]. Contrary to the GARCH$(1,1)$ process, the noise defined in Equation (21) is not a martingale difference sequence for which the limit theory is more classical.

We simulate $N = 1,000$ independent trajectories of size $n = 10,000$ of models (19). The same series is partitioned as three series of sizes $n = 1,000$, $n = 5,000$ and $n = 10,000$. For each of these N replications, we use the least squares estimation method to estimate the coefficient θ_0 and we apply Portmanteau tests to the residuals for different values of $m \in \{1, 2, 3, 6, 12, 15\}$, where m is the number of autocorrelations used in the Portmanteau test statistic. For the nominal level $\alpha = 5\%$, the empirical size over the N independent replications should vary between the significant limits 3.6% and 6.4% with probability 95%. When the relative rejection frequencies are outside the 95% significant limits, they are displayed in bold type in Tables 1, 2 and 3.

TABLE 1

Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of a strong FARIMA$(0, d_0, 0)$ defined by (19) with $\theta_0 = (0, 0, d_0)$. The nominal asymptotic level of the tests is $\alpha = 5\%$. The number of replications is $N = 1,000$.

d_0	Length n	Lag m	$\mathrm{LB_{SN}}$	$\mathrm{BP_{SN}}$	$\mathrm{LB_W}$	$\mathrm{BP_W}$	$\mathrm{LB_S}$	$\mathrm{BP_S}$
0.05	$n = 1,000$	1	**3.3**	**3.3**	4.6	4.5	n.a.	n.a.
		2	4.5	4.5	4.9	4.9	5.8	5.8
		3	5.2	5.1	4.7	4.4	4.9	4.8
		6	5.8	5.8	4.6	4.5	5.1	5.0
		12	6.0	5.6	5.3	4.6	5.2	5.0
		15	5.6	5.2	4.7	4.3	5.3	4.7
0.05	$n = 5,000$	1	**6.8**	**6.8**	**6.6**	**6.6**	n.a.	n.a.
		2	**6.8**	**6.8**	6.4	6.4	**7.9**	**7.9**
		3	**6.6**	**6.6**	5.7	5.7	5.8	5.8
		6	**6.5**	6.4	5.6	5.6	5.7	5.6
		12	6.4	6.4	5.3	5.3	6.0	5.9
		15	6.1	6.0	4.7	4.6	5.3	5.2
0.05	$n = 10,000$	1	4.9	4.9	5.3	5.3	n.a.	n.a.
		2	5.4	5.4	6.6	6.6	**7.8**	**7.8**
		3	5.7	5.7	5.9	5.9	6.2	6.2
		6	5.9	5.8	4.5	4.5	4.6	4.6
		12	5.3	5.3	5.4	5.4	5.6	5.6
		15	4.4	4.3	4.8	4.8	4.9	4.9
0.20	$n = 1,000$	1	3.6	**3.5**	4.3	4.3	n.a.	n.a.
		2	4.7	4.7	4.7	4.7	5.8	5.7
		3	5.2	5.0	4.3	4.3	4.9	4.7
		6	6.0	5.9	4.7	4.5	5.0	4.9
		12	5.7	5.4	5.3	4.7	5.2	4.9
		15	5.9	5.6	4.8	4.2	5.2	4.8
0.20	$n = 5,000$	1	**6.6**	**6.6**	**6.5**	**6.5**	n.a.	n.a.
		2	**6.6**	**6.6**	6.4	6.4	**7.9**	**7.9**
		3	**6.7**	**6.7**	5.7	5.7	5.8	5.8
		6	6.3	6.3	5.6	5.6	5.7	5.5
		12	6.3	6.2	5.5	5.3	6.0	5.9
		15	6.1	5.9	4.7	4.6	5.3	5.2
0.20	$n = 10,000$	1	4.8	4.8	5.3	5.3	n.a.	n.a.
		2	5.4	5.4	**6.6**	**6.6**	**7.8**	**7.8**
		3	5.5	5.5	5.9	5.9	6.3	6.3
		6	5.8	5.8	4.5	4.5	4.6	4.6
		12	5.4	5.3	5.5	5.5	5.6	5.6
		15	4.4	4.3	4.7	4.7	4.9	4.9
0.45	$n = 1,000$	1	3.9	3.8	4.9	4.9	n.a.	n.a.
		2	5.1	5.0	4.8	4.6	5.9	5.9
		3	5.2	5.2	4.3	4.3	4.8	4.8
		6	6.2	6.0	4.7	4.3	4.9	4.9
		12	5.8	5.4	4.8	4.7	4.9	4.8
		15	5.6	5.5	4.5	4.2	5.0	4.8
0.45	$n = 5,000$	1	**6.6**	**6.6**	**6.6**	**6.6**	n.a.	n.a.
		2	**6.7**	**6.7**	**6.5**	**6.5**	**8.0**	**8.0**
		3	**6.6**	**6.6**	5.7	5.7	5.8	5.8
		6	6.3	6.3	5.4	5.4	5.6	5.5
		12	6.2	6.2	5.5	5.5	6.0	5.9
		15	6.2	5.9	4.6	4.6	5.5	5.3
0.45	$n = 10,000$	1	5.0	5.0	5.3	5.3	n.a.	n.a.
		2	5.4	5.4	**6.6**	**6.6**	**7.9**	**7.9**
		3	5.3	5.3	5.9	5.9	6.3	6.3
		6	5.8	5.8	4.7	4.6	4.7	4.7
		12	5.4	5.4	5.5	5.5	5.7	5.7
		15	4.6	4.5	4.9	4.8	4.9	4.9

Table 2

Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of a weak FARIMA$(0, d_0, 0)$ defined by (19) with $\theta_0 = (0, 0, d_0)$ and where $\omega = 0.4$, $\alpha_1 = 0.3$ and $\beta_1 = 0.3$ in (21). The nominal asymptotic level of the tests is $\alpha = 5\%$. The number of replications is $N = 1,000$.

d_0	Length n	Lag m	LB_{SN}	BP_{SN}	LB_W	BP_W	LB_S	BP_S
0.05	$n = 1,000$	1	4.4	4.4	5.4	5.4	n.a.	n.a.
		2	4.3	4.2	5.7	5.7	**15.6**	**15.5**
		3	5.9	5.9	5.3	5.0	**14.2**	**14.0**
		6	5.2	5.1	6.0	6.0	**14.6**	**14.4**
		12	4.5	4.1	4.2	4.0	**11.0**	**10.7**
		15	4.0	3.9	4.2	3.9	**11.1**	**10.6**
0.05	$n = 5,000$	1	4.3	4.3	5.1	5.1	n.a.	n.a.
		2	4.4	4.4	5.8	5.8	**16.9**	**16.8**
		3	5.0	5.0	5.5	5.5	**16.5**	**16.5**
		6	5.6	5.6	4.5	4.5	**14.8**	**14.6**
		12	5.1	5.1	5.0	4.9	**12.6**	**12.5**
		15	5.2	5.1	4.9	4.7	**11.8**	**11.6**
0.05	$n = 10,000$	1	5.7	5.7	5.3	5.1	n.a.	n.a.
		2	5.0	5.0	4.5	4.5	**17.4**	**17.4**
		3	5.5	5.5	4.7	4.6	**17.2**	**17.2**
		6	5.3	5.3	5.0	5.0	**14.2**	**14.1**
		12	4.9	4.9	4.7	4.7	**11.0**	**11.0**
		15	4.9	4.8	4.7	4.6	**10.2**	**10.2**
0.20	$n = 1,000$	1	4.9	4.9	4.3	4.3	n.a.	n.a.
		2	4.0	4.0	5.7	5.6	**15.5**	**15.4**
		3	6.0	6.0	5.0	4.8	**14.0**	**13.8**
		6	5.2	5.1	5.7	5.6	**14.3**	**14.2**
		12	4.4	4.0	4.3	4.0	**10.8**	**10.5**
		15	3.9	3.8	4.2	3.9	**10.8**	**10.1**
0.20	$n = 5,000$	1	4.3	4.3	5.0	5.0	n.a.	n.a.
		2	4.3	4.3	5.9	5.8	**16.9**	**16.9**
		3	5.2	5.2	5.4	5.4	**16.7**	**16.7**
		6	5.6	5.5	4.6	4.5	**14.8**	**14.7**
		12	5.2	5.2	5.0	4.9	**12.5**	**12.4**
		15	5.2	5.2	4.8	4.6	11.7	11.7
0.20	$n = 10,000$	1	5.7	5.7	5.2	5.2	n.a.	n.a.
		2	5.1	5.1	4.5	4.5	**17.3**	**17.3**
		3	5.7	5.6	4.7	4.7	**17.2**	**17.2**
		6	5.1	5.1	4.9	4.9	**14.2**	**14.2**
		12	4.8	4.8	4.7	4.7	**11.0**	**11.0**
		15	4.9	4.7	4.6	4.6	**10.2**	**10.2**
0.45	$n = 1,000$	1	4.5	4.5	5.4	5.4	n.a.	n.a.
		2	4.1	4.1	6.0	6.0	**16.2**	**16.1**
		3	5.9	5.7	5.3	5.3	**14.6**	**14.5**
		6	5.2	4.8	5.5	5.4	**14.4**	**14.1**
		12	4.0	3.7	4.2	4.2	**11.2**	**10.8**
		15	3.8	3.7	4.3	3.9	**10.6**	**10.4**
0.45	$n = 5,000$	1	4.6	4.6	5.0	5.0	n.a.	n.a.
		2	4.3	4.3	5.9	5.9	**16.7**	**16.7**
		3	4.9	4.9	5.4	5.4	**16.8**	**16.7**
		6	5.7	5.6	4.6	4.6	**15.1**	**14.9**
		12	5.3	5.3	5.1	5.1	**12.7**	**12.4**
		15	5.1	5.0	4.8	4.8	**11.7**	**11.7**
0.45	$n = 10,000$	1	5.7	5.7	5.2	5.2	n.a.	n.a.
		2	5.0	5.0	4.7	4.7	**17.2**	**17.2**
		3	5.8	5.7	4.7	4.7	**17.5**	**17.4**
		6	5.1	5.1	5.0	4.9	**14.3**	**14.3**
		12	4.8	4.8	4.7	4.7	**10.9**	**10.9**
		15	4.9	4.7	4.6	4.6	**10.2**	**10.2**

Table 3
Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of weak FARIMA$(0, d_0, 0)$ defined by (19)–(21) with $\theta_0 = (0, 0, d_0)$. The nominal asymptotic level of the tests is $\alpha = 5\%$. The number of replications is $N = 1,000$.

d_0	Length n	Lag m	LB_{SN}	BP_{SN}	LB_W	BP_W	LB_S	BP_S
0.05	$n = 1,000$	1	**3.3**	**3.3**	**8.7**	**8.6**	n.a.	n.a.
		2	3.8	3.7	6.1	6.1	**16.9**	**16.9**
		3	**3.5**	**3.5**	4.8	4.7	**14.8**	**14.8**
		6	**3.3**	**3.2**	4.0	4.0	**14.1**	**14.0**
		12	**1.0**	**0.9**	**2.5**	**2.4**	**13.0**	**12.8**
		15	**1.0**	**0.9**	**2.3**	**2.1**	**12.8**	**12.2**
0.05	$n = 5,000$	1	3.9	3.9	5.3	5.3	n.a.	n.a.
		2	4.8	4.8	5.2	5.2	**18.7**	**18.7**
		3	5.6	5.6	5.3	5.3	**15.1**	**15.0**
		6	4.8	4.8	4.3	4.3	**12.4**	**12.4**
		12	3.9	3.9	**3.3**	**3.3**	**11.2**	**11.1**
		15	**3.5**	**3.5**	**2.7**	**2.7**	**10.2**	**10.1**
0.05	$n = 10,000$	1	5.4	5.4	5.2	5.2	n.a.	n.a.
		2	5.6	5.6	5.3	5.3	**18.6**	**18.6**
		3	4.9	4.9	5.3	5.2	**16.6**	**16.5**
		6	4.8	4.8	5.5	5.4	**13.3**	**13.3**
		12	4.1	4.0	4.0	4.0	**12.2**	**12.2**
		15	5.0	5.0	**3.5**	**3.5**	**11.2**	**11.2**
0.20	$n = 1,000$	1	**3.3**	**3.3**	4.9	4.9	n.a.	n.a.
		2	4.2	4.1	4.4	4.3	**14.7**	**14.7**
		3	3.7	3.7	**3.4**	**3.2**	**12.8**	**12.8**
		6	3.6	**3.4**	**2.7**	**2.7**	**12.9**	**12.8**
		12	**1.1**	**1.0**	**1.9**	**1.7**	**11.8**	**11.3**
		15	**0.9**	**0.6**	**1.8**	**1.7**	**12.0**	**11.5**
0.20	$n = 5,000$	1	3.8	3.8	5.5	5.5	n.a.	n.a.
		2	4.7	4.7	5.1	5.1	**18.8**	**18.8**
		3	5.8	5.8	5.2	5.2	**15.0**	**15.0**
		6	4.9	4.9	4.3	4.3	**12.5**	**12.4**
		12	3.9	3.9	**3.4**	**3.4**	**11.1**	**11.1**
		15	**3.5**	**3.3**	**2.7**	**2.7**	**10.2**	**10.1**
0.20	$n = 10,000$	1	5.4	5.4	5.1	5.1	n.a.	n.a.
		2	5.6	5.6	5.3	5.3	**18.8**	**18.8**
		3	5.0	5.0	5.2	5.2	**16.6**	**16.6**
		6	4.8	4.8	5.4	5.4	**13.3**	**13.3**
		12	4.0	4.0	4.0	4.0	**12.1**	**12.1**
		15	5.3	5.3	**3.4**	**3.4**	**11.2**	**11.2**
0.45	$n = 1,000$	1	**3.5**	**3.5**	**9.0**	**9.0**	n.a.	n.a.
		2	4.1	4.1	5.9	5.9	17.5	17.5
		3	3.9	3.7	5.0	4.8	**15.0**	**14.6**
		6	**3.4**	**3.4**	3.7	3.7	**14.1**	**13.9**
		12	**0.9**	**0.9**	**2.0**	**2.0**	**12.9**	**12.2**
		15	**1.0**	**0.5**	**1.9**	**1.7**	**13.1**	**12.8**
0.45	$n = 5,000$	1	4.1	4.1	5.4	5.4	n.a.	n.a.
		2	4.6	4.6	5.2	5.2	**18.8**	**18.7**
		3	5.6	5.6	5.2	5.2	**15.2**	**15.2**
		6	5.1	5.0	4.4	4.4	**12.5**	**12.4**
		12	4.0	3.8	**3.5**	**3.5**	**11.1**	**11.1**
		15	**3.5**	**3.5**	**2.6**	**2.6**	**10.0**	**9.9**
0.45	$n = 10,000$	1	5.5	5.5	5.1	5.1	n.a.	n.a.
		2	5.6	5.6	5.3	5.3	**18.7**	**18.6**
		3	4.7	4.7	5.2	5.2	**16.6**	**16.6**
		6	4.8	4.8	5.3	5.3	**13.3**	**13.3**
		12	4.0	4.0	4.0	4.0	**12.1**	**12.1**
		15	5.2	5.2	**3.5**	**3.5**	**11.1**	**11.1**

For the standard Box-Pierce test, the model is therefore rejected when the statistic Q_m^{BP} or Q_m^{LB} is larger than $\chi^2_{(m-p-q-1)}(0.95)$ in a FARIMA(p, d_0, q) case (see [33]). Consequently the empirical size is not available (n.a.) for the statistic Q_m^{BP} or Q_m^{LB} because they are not applicable for $m \leq p+q+1$. For the proposed self-normalized test $\mathrm{BP}_{\mathrm{SN}}$ or $\mathrm{LB}_{\mathrm{SN}}$, the model is rejected when the statistic Q_m^{SN} or $\tilde{Q}_m^{\mathrm{SN}}$ is larger than $\mathcal{U}_m(0.95)$, where the critical values $\mathcal{U}_K(0.95)$ (for $K = 1, \dots, 20$) are tabulated in Lobato (see Table 1 in [37]).

Table 1 displays the relative rejection frequencies of the null hypothesis **(H0)** that the data generating process (DGP for short) is a strong FARIMA$(0, d_0, 0)$ model (19), over the N independent replications. For all tests, the percentages of rejection belong globally to the confident interval with probabilities 95%, except for LB_{S} and BP_{S} (see Table 8).

Now, we repeat the same experiments on two weak FARIMA models. As expected Tables 2 and 3 show that the standard LB_{S} or BP_{S} test poorly performs in assessing the adequacy of these particular weak FARIMA models. Indeed, we observe that the observed relative rejection frequencies of LB_{S} and BP_{S} are definitely outside the significant limits. Thus we draw the conclusion that the error of the first kind is globally well controlled by all the tests in the strong case, but only by the proposed tests in the weak cases.

4.2. Empirical power

In this section, we repeat the same experiments as in Section 4.1 to examine the power of the tests for the null hypothesis of Model (19) with $a = b = 0$ (*i.e.* a FARIMA$(0, d_0, 0)$) against the FARIMA$(0, d_0, 1)$ alternative defined by Model (19) with $\theta_0 = (0, b, d_0)'$ and where the innovation process $(\epsilon_t)_{t\in\mathbb{Z}}$ follows the two weak white noises introduced in Section 4.1.

For each of these N replications we fit a FARIMA$(0, d_0, 0)$ model (19) and perform standard and modified tests based on $m = 1, 2, 3$, 6, 12 and 15 residual autocorrelations.

Tables 4 and 5 compare the empirical powers of Model (19) with $\theta_0 = (0, 0.2, d_0)'$ over the N independent replications. For these particular weak models, we notice that the standard BP_{S} and LB_{S} and our proposed tests have very similar powers except for $\mathrm{BP}_{\mathrm{SN}}$ and $\mathrm{LB}_{\mathrm{SN}}$ when $n = 5,000$.

In these Monte Carlo experiments, we illustrate that the proposed test statistics have reasonable finite sample performance. Under nonindependent errors, it appears that the standard test statistics are generally non reliable, overrejecting severely, while the proposed tests statistics offer satisfactory levels. Even for independent errors, they seem preferable to the standard ones when the number m of autocorrelations is small. Moreover, the error of first kind is well controlled. Contrarily to the standard tests based on BP_{S} or LB_{S}, the proposed tests can be used safely for m small. For all these above reasons, we think that the modified versions that we propose in this paper are preferable to the standard ones for diagnosing FARIMA models under nonindependent errors.

Table 4

Empirical power (in %) of the modified and standard versions of the LB and BP tests in the case of a weak FARIMA$(0, d_0, 1)$ defined by (19) with $\theta_0 = (0, 0.2, d_0)$ and where $\omega = 0.4$, $\alpha_1 = 0.3$ and $\beta_1 = 0.3$ in (20). The nominal asymptotic level of the tests is $\alpha = 5\%$. The number of replications is $N = 1,000$.

d_0	Length n	Lag m	LB_{SN}	BP_{SN}	LB_W	BP_W	LB_S	BP_S
		1	30.1	30.1	100.0	100.0	n.a.	n.a.
		2	55.7	55.7	100.0	100.0	100.0	100.0
0.05	$n = 5,000$	3	75.7	75.7	100.0	100.0	100.0	100.0
		6	87.1	87.1	100.0	100.0	100.0	100.0
		12	87.0	86.8	100.0	100.0	100.0	100.0
		15	87.3	87.2	100.0	100.0	100.0	100.0
		1	50.0	50.0	100.0	100.0	n.a.	n.a.
		2	79.5	79.4	100.0	100.0	100.0	100.0
0.05	$n = 10,000$	3	95.2	95.2	100.0	100.0	100.0	100.0
		6	98.0	98.0	100.0	100.0	100.0	100.0
		12	98.6	98.6	100.0	100.0	100.0	100.0
		15	99.0	99.0	100.0	100.0	100.0	100.0
		1	98.2	98.2	99.9	99.9	n.a.	n.a.
		2	94.6	94.6	99.5	99.5	100.0	100.0
0.20	$n = 5,000$	3	92.3	92.3	99.6	99.6	100.0	100.0
		6	91.0	91.0	99.6	99.6	100.0	100.0
		12	88.8	88.7	99.8	99.8	100.0	100.0
		15	88.6	88.6	99.8	99.8	100.0	100.0
		1	99.7	99.7	100.0	100.0	n.a.	n.a.
		2	99.2	99.2	100.0	100.0	100.0	100.0
0.20	$n = 10,000$	3	99.3	99.2	100.0	100.0	100.0	100.0
		6	98.8	98.8	100.0	100.0	100.0	100.0
		12	99.3	99.3	100.0	100.0	100.0	100.0
		15	99.3	99.3	100.0	100.0	100.0	100.0
		1	98.2	98.2	99.8	99.8	n.a.	n.a.
		2	94.4	94.3	99.5	99.5	100.0	100.0
0.45	$n = 5,000$	3	92.4	92.4	99.6	99.6	100.0	100.0
		6	90.9	90.8	99.6	99.6	100.0	100.0
		12	88.9	88.9	99.8	99.8	100.0	100.0
		15	88.8	88.5	99.8	99.8	100.0	100.0
		1	99.7	99.7	100.0	100.0	n.a.	n.a.
		2	99.0	99.0	100.0	100.0	100.0	100.0
0.45	$n = 10,000$	3	99.2	99.2	100.0	100.0	100.0	100.0
		6	98.9	98.9	100.0	100.0	100.0	100.0
		12	99.3	99.3	100.0	100.0	100.0	100.0
		15	99.3	99.3	100.0	100.0	100.0	100.0

Table 5

Empirical power (in %) of the modified and standard versions of the LB and BP tests in the case of a weak FARIMA$(0,d_0,1)$ defined by (19)–(21) with $\theta_0=(0,0.2,d_0)$. The nominal asymptotic level of the tests is $\alpha=5\%$. The number of replications is $N=1,000$.

d_0	Length n	Lag m	$\mathrm{LB_{SN}}$	$\mathrm{BP_{SN}}$	$\mathrm{LB_W}$	$\mathrm{BP_W}$	$\mathrm{LB_S}$	$\mathrm{BP_S}$
		1	20.3	20.3	99.9	99.9	n.a.	n.a.
		2	56.7	56.6	99.9	99.9	99.9	99.9
0.05	$n=5,000$	3	69.1	69.1	99.9	99.9	99.9	99.9
		6	75.9	75.9	99.9	99.9	99.9	99.9
		12	71.9	71.4	99.9	99.9	99.9	99.9
		15	68.5	68.0	99.9	99.9	99.9	99.9
		1	60.0	60.0	100.0	100.0	n.a.	n.a.
		2	81.8	81.8	100.0	100.0	100.0	100.0
0.05	$n=10,000$	3	90.3	90.3	100.0	100.0	100.0	100.0
		6	93.9	93.9	100.0	100.0	100.0	100.0
		12	93.8	93.8	100.0	100.0	100.0	100.0
		15	93.7	93.7	100.0	100.0	100.0	100.0
		1	92.3	92.3	99.9	99.9	n.a.	n.a.
		2	86.1	86.0	98.6	98.6	99.8	99.8
0.20	$n=5,000$	3	82.3	82.3	99.2	99.1	99.8	99.8
		6	80.0	80.0	98.9	98.9	99.9	99.9
		12	73.1	72.8	98.7	98.7	99.6	99.6
		15	68.3	68.0	98.4	98.4	99.5	99.5
		1	99.2	99.2	100.0	100.0	n.a.	n.a.
		2	96.4	96.4	100.0	100.0	100.0	100.0
0.20	$n=10,000$	3	94.6	94.6	100.0	100.0	100.0	100.0
		6	95.1	95.1	100.0	100.0	100.0	100.0
		12	95.2	95.2	100.0	100.0	100.0	100.0
		15	94.0	94.0	100.0	100.0	100.0	100.0
		1	92.4	92.4	99.9	99.9	n.a.	n.a.
		2	85.6	85.6	98.6	98.6	99.8	99.8
0.45	$n=5,000$	3	82.1	82.0	99.3	99.3	99.8	99.8
		6	80.3	80.3	98.9	98.9	99.9	99.9
		12	73.0	72.7	98.7	98.7	99.6	99.6
		15	68.2	68.1	98.4	98.4	99.5	99.5
		1	99.2	99.2	100.0	100.0	n.a.	n.a.
		2	96.4	96.4	100.0	100.0	100.0	100.0
0.45	$n=10,000$	3	94.8	94.8	100.0	100.0	100.0	100.0
		6	95.2	95.2	100.0	100.0	100.0	100.0
		12	95.0	95.0	100.0	100.0	100.0	100.0
		15	94.0	94.0	100.0	100.0	100.0	100.0

4.3. Illustrative example

We now consider an application to the daily log returns (also simply called the returns) of the Nikkei and Standard & Poor's 500 indices (S&P 500, for short). The returns are defined by $r_t = 100\log(p_t/p_{t-1})$ where p_t denotes the price index of the S&P 500 index at time t. The observations of the S&P 500 (resp. the Nikkei) index cover the period from January 3, 1950 to February 14, 2019 (resp. from January 5, 1965 to February 14, 2019). The length of the series is $n = 17,391$ (resp. $n = 13,319$) for the S&P 500 (resp. the Nikkei) index. The data can be downloaded from the website Yahoo Finance: `http://fr.finance.yahoo.com/`.

In Financial Econometrics the returns are often assumed to be a white noise. In view of the so-called volatility clustering, it is well known that the strong white noise model is not adequate for these series (see for instance [24, 38, 5, 7]).

A long-range memory property of the stock market returns series was largely investigated by [15] which shows that there are more correlation beetwen power transformation of the absolute return $|r_t|^v$ ($v > 0$) than returns themselves (see also [3], [42], [2] and [35]). We choose here the case where $v = 2$ which corresponds to the squared returns $(r_t^2)_{t\geq 1}$ process. The mean and the standard deviation of $(r_t^2)_{t\geq 1}$ are $0,9347$ and $5,0036$ (resp. $1,6167$ and $5,4759$) for the S&P 500 (resp. the Nikkei) index. Following a similar way as in [34] we denote by $(X_t)_{t\geq 1}$ the centered series of the squared returns, that is, $X_t = r_t^2 - 0,9347$ (resp. $X_t = r_t^2 - 1,6167$) for the S&P 500 (resp. the Nikkei) index. Figure 1 (resp. Figure 3) plots the returns and the sample autocorrelations of $(X_t)_{t\geq 1}$ of the S&P 500 (resp. of the Nikkei). The centered squared returns $(X)_{t\geq 1}$ have significant positive autocorrelations at least up to lag 80 (see Figure 1 and Figure 3) which confirm the claim that stock market returns have long-term memory (see for instance [15], for more details).

We first fit a FARIMA$(1, d_0, 1)$ model defined in (19) to the process $(X)_{t\geq 1}$ of the S&P 500 and the Nikkei returns. Let $\hat{\theta}_n^{\mathrm{SP500}}$ and $\hat{\theta}_n^{\mathrm{Nikkei}}$ be respectively the least squares estimators of the parameter $\theta_0 = (a, b, d_0)'$ for the model (19) in the case of the S&P 500 and the Nikkei. The least squares estimators were obtained as

$$\hat{\theta}_n^{\mathrm{SP500}} = \begin{pmatrix} -0.3371 & [0.1105] & (0.0023) \\ -0.1795 & [0.0788] & (0.0227) \\ 0.2338 & [0.0367] & (0.0000) \end{pmatrix} \text{ and } \hat{\sigma}_\epsilon^2 = 22.9076 \times 10^{-8}$$

and

$$\hat{\theta}_n^{\mathrm{Nikkei}} = \begin{pmatrix} -0.0217 & [0.3670] & (0.9528) \\ 0.1579 & [0.3053] & (0.6050) \\ 0.3217 & [0.0589] & (0.0000) \end{pmatrix} \text{ and } \hat{\sigma}_\epsilon^2 = 25.6844 \times 10^{-8}, \quad (22)$$

where the estimated asymptotic standard errors obtained from $\Sigma_{\hat{\theta}} := J^{-1}IJ^{-1}$ (respectively the p-values), of the estimated parameters (first column), are given

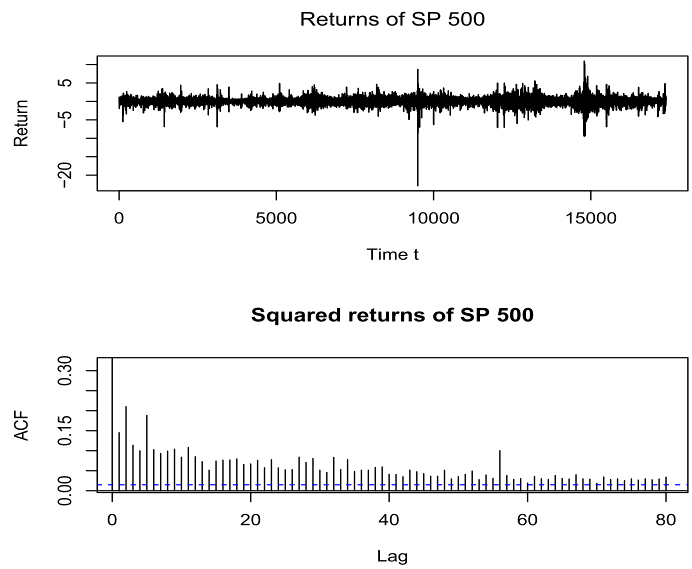

FIG 1. *Returns and the sample autocorrelations of squared returns of the S&P 500.*

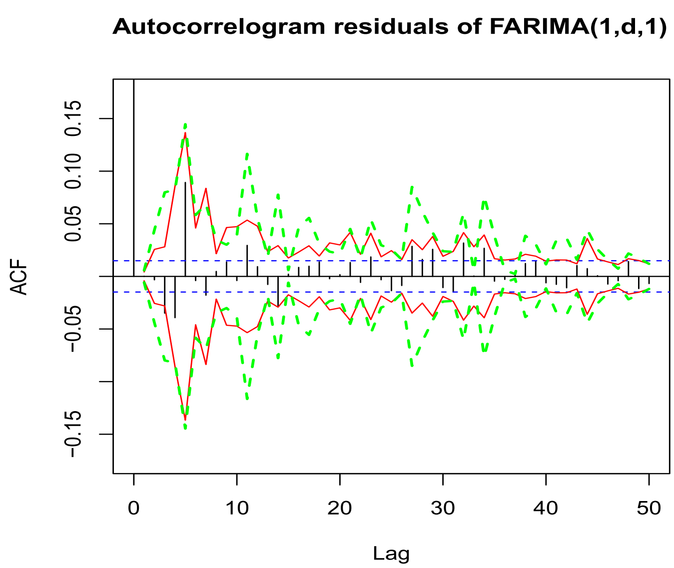

FIG 2. *Autocorrelation of the FARIMA$(1, 0.2338, 1)$ residuals for the squares of the S&P 500 returns. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong FARIMA assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak FARIMA assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 2. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 7.*

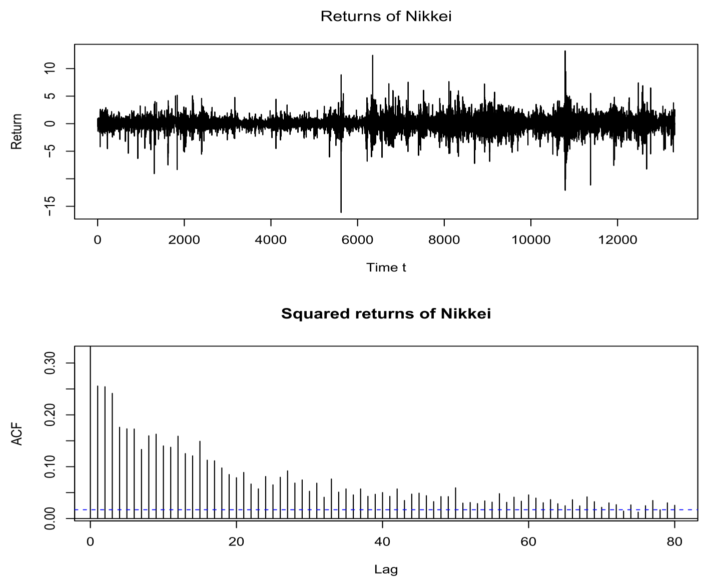

FIG 3. *Returns and the sample autocorrelations of squared returns of the Nikkei.*

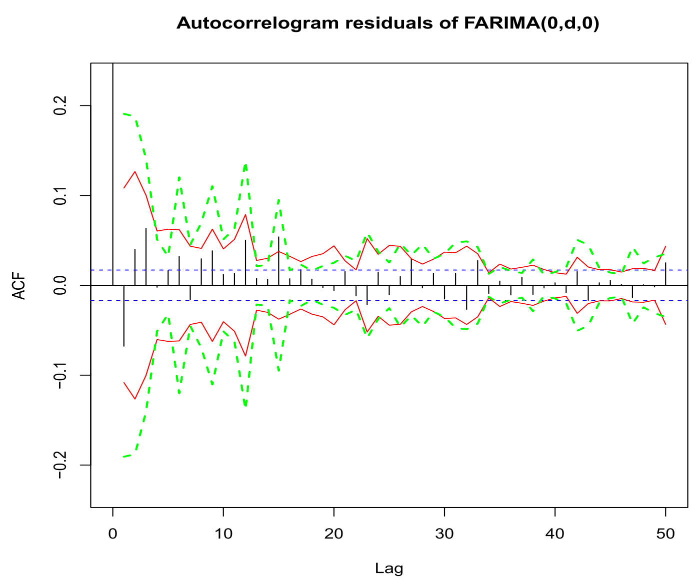

FIG 4. *Autocorrelation of the FARIMA$(0, 0.2132, 0)$ residuals for the squares of the Nikkei returns. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong FARIMA assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak FARIMA assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 2. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 7.*

into brackets (respectively in parentheses). Note that for these series, the estimated coefficients $|\hat{a}_n|$ and $|\hat{b}_n|$ are smaller than one. This is in accordance with the assumptions that the power series a_θ^{-1} and b_θ^{-1} are well defined (remind that the moving average polynomial is denoted b_θ and the autoregressive polynomials a_θ). We also observe that the estimated long-range dependence coefficients $\hat{d}_n$ is significant for any reasonable asymptotic level and is inside $(0, 0.5)$. So we think that the assumption **(A2)** is satisfied and thus our asymptotic normality theorem on the residual autocorrelations can be applied.

Concerning the S&P 500, the estimators of the parameters a and b are significant whereas it is not the case for the Nikkei (see (22)). In the Nikkei case, the coefficients could reasonably be set to zero. So we adjust a FARIMA$(0, d_0, 0)$ for the squares of Nikkei returns and (22) is reduced as

$$\hat{\theta}_n^{\text{Nikkei}} = \begin{pmatrix} 0.2132 & [0.0259] & (0.0000) \end{pmatrix} \text{ and } \hat{\sigma}_\epsilon^2 = 25.9793 \times 10^{-8}.$$

We thus apply Portmanteau tests to the residuals of FARIMA$(1, d_0, 1)$ (resp. FARIMA$(0, d_0, 0)$) model for the process $(X_t)_{t \geq 1}$ of S&P 500 (resp. of Nikkei). Table 6 (resp. Table 7) displays the statistics and the p-values of the standard and modified versions of BP and LB tests of model (19). From Tables 6 and 7, we draw the conclusion that:

- the strong FARIMA$(1, 0.2338, 1)$ and FARIMA$(0, 0.2132, 0)$ models are rejected
- but the weak FARIMA$(1, 0.2338, 1)$ and the weak FARIMA$(0, 0.2132, 0)$ models are not rejected.

Figure 2 (resp. Figure 4) displays the residual autocorrelations and their 5% significance limits under the strong FARIMA and weak FARIMA assumptions. In view of Figures 2 and 4, the diagnostic checking of residuals does not indicate any inadequacy for the proposed tests. All of the sample autocorrelations should lie between the bands (at 95%) shown as dashed lines (green color) and solid lines (red color) for the modified tests, while the horizontal dotted (blue color) for standard test indicate that strong FARIMA is not adequate. Figure 2 (resp. Figure 4) confirms the conclusions drawn from Table 6 (resp. Table 7).

5. Proofs

The following proofs are quite technical and are adaptations of the arguments used in [20], [19] and [7].

The results of [6] which will be needed for all the proofs are collected in the following Subsection 5.1 in order to have a self-contained paper.

In all our proofs, K is a positive constant that may vary from line to line.

5.1. Preliminary results

In this subsection, we shall give some results on estimations of the coefficients of formal power series that will arise in our study.

TABLE 6
Modified and standard versions of Portmanteau tests to check the null hypothesis that the S&P 500 squared returns follow a FARIMA$(1, 0.2338, 1)$ model (19).

Lag m	1	2	3	4	5	6	7
$\hat{\rho}(m)$	0.0002	−0.0033	−0.0350	−0.0393	0.0893	−0.0040	−0.0179
$\mathrm{LB}_{\mathrm{SN}}$	0.0653	18.150	41.924	58.057	186.72	313.78	341.38
$\mathrm{BP}_{\mathrm{SN}}$	0.0653	18.146	41.912	58.037	186.64	313.64	341.20
LB_{W}	0.0008	0.1885	21.445	48.248	186.95	187.23	192.77
BP_{W}	0.0008	0.1884	21.439	48.232	186.88	187.15	192.67
$\mathrm{p}_{\mathrm{W}}^{\mathrm{LB}}$	0.8525	0.6985	0.0916	0.3137	0.0678	0.0717	0.0752
$\mathrm{p}_{\mathrm{W}}^{\mathrm{BP}}$	0.8525	0.6986	0.0917	0.3138	0.0679	0.0718	0.0753
$\mathrm{p}_{\mathrm{S}}^{\mathrm{LB}}$	n.a.	n.a.	n.a.	0.0000	0.0000	0.0000	0.0000
$\mathrm{p}_{\mathrm{S}}^{\mathrm{BP}}$	n.a.	n.a.	n.a.	0.0000	0.0000	0.0000	0.0000
Lag m	8	9	10	11	12	13	14
$\hat{\rho}(m)$	0.0047	0.0137	−0.0040	0.0295	0.0093	−0.0077	−0.0286
$\mathrm{LB}_{\mathrm{SN}}$	397.27	397.38	415.22	465.52	468.76	567.87	573.02
$\mathrm{BP}_{\mathrm{SN}}$	397.04	397.13	414.93	465.17	468.33	567.38	572.49
LB_{W}	193.16	196.42	196.69	211.82	213.31	214.34	228.55
BP_{W}	193.09	196.34	196.61	211.74	213.22	214.25	228.45
$\mathrm{p}_{\mathrm{W}}^{\mathrm{LB}}$	0.0758	0.0786	0.0986	0.1053	0.1148	0.1226	0.1047
$\mathrm{p}_{\mathrm{W}}^{\mathrm{BP}}$	0.0758	0.0787	0.0987	0.1054	0.1150	0.1228	0.1048
$\mathrm{p}_{\mathrm{S}}^{\mathrm{LB}}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\mathrm{p}_{\mathrm{S}}^{\mathrm{BP}}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lag m	15	16	17	18	19	20	21
$\hat{\rho}(m)$	0.0021	0.0086	0.0097	0.0137	−0.0023	0.0016	0.0132
$\mathrm{LB}_{\mathrm{SN}}$	588.61	701.16	738.23	738.58	749.24	778.88	788.01
$\mathrm{BP}_{\mathrm{SN}}$	588.04	700.44	737.42	737.73	748.33	777.90	786.97
LB_{W}	228.63	229.91	231.54	234.83	234.92	234.97	238.00
BP_{W}	228.52	229.80	231.44	234.72	234.81	234.86	237.89
$\mathrm{p}_{\mathrm{W}}^{\mathrm{LB}}$	0.1079	0.1113	0.2212	0.2138	0.2127	0.2169	0.2324
$\mathrm{p}_{\mathrm{W}}^{\mathrm{BP}}$	0.1080	0.1114	0.2214	0.2140	0.2130	0.2171	0.2327
$\mathrm{p}_{\mathrm{S}}^{\mathrm{LB}}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\mathrm{p}_{\mathrm{S}}^{\mathrm{BP}}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

We begin by recalling the following properties on power series. If for $|z| \leq R$, the power series $f(z) = \sum_{i \geq 0} a_i z^i$ and $g(z) = \sum_{i \geq 0} b_i z^i$ are well defined, then one has $(fg)(z) = \sum_{i \geq 0} c_i z^i$ is also well defined for $|z| \leq R$ with the sequence $(c_i)_{i \geq 0}$ which is given by $c = a * b$ where $*$ denotes the convolution product between a and b defined by $c_i = \sum_{k=0}^{i} a_k b_{i-k} = \sum_{k=0}^{i} a_{i-k} b_k$. We will make use of the Young inequality that states that if the sequence $a \in \ell^{r_1}$ and $b \in \ell^{r_2}$ are such that $\frac{1}{r_1} + \frac{1}{r_2} = 1 + \frac{1}{r}$ with $1 \leq r_1, r_2, r \leq \infty$, then

$$\|a * b\|_{\ell^r} \leq \|a\|_{\ell^{r_1}} \times \|b\|_{\ell^{r_2}} .$$

Now we come back to the power series that arise in our context. Remind that for the true value of the parameter,

$$a_{\theta_0}(L)(1-L)^{d_0} X_t = b_{\theta_0}(L)\epsilon_t. \tag{23}$$

Table 7

Modified and standard versions of Portmanteau tests to check the null hypothesis that the Nikkei squared returns follow a FARIMA$(0, 0.2132, 0)$ model as in (19) with $a = b = 0$.

Lag m	1	2	3	4	5	6	7
$\hat{\rho}(m)$	−0.0678	0.0400	0.0634	−0.0022	0.0165	0.0320	−0.0158
$\mathrm{LB}_{\mathrm{SN}}$	5.7332	29.005	34.758	34.779	66.692	288.57	324.46
$\mathrm{BP}_{\mathrm{SN}}$	5.7319	28.997	34.745	34.764	66.657	288.40	324.24
LB_{W}	61.211	82.507	136.13	136.20	139.84	153.46	156.78
BP_{W}	61.198	82.487	136.09	136.16	139.76	153.41	156.73
$\mathrm{p}_{\mathrm{W}}^{\mathrm{LB}}$	0.1086	0.2186	0.1830	0.2551	0.3002	0.3519	0.3609
$\mathrm{p}_{\mathrm{W}}^{\mathrm{BP}}$	0.1086	0.2187	0.1831	0.2552	0.3003	0.3521	0.3611
$\mathrm{p}_{\mathrm{S}}^{\mathrm{LB}}$	n.a.	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\mathrm{p}_{\mathrm{S}}^{\mathrm{BP}}$	n.a.	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lag m	8	9	10	11	12	13	14
$\hat{\rho}(m)$	0.0295	0.0384	0.0121	0.0133	0.0503	0.0076	0.0068
$\mathrm{LB}_{\mathrm{SN}}$	387.88	512.70	575.09	600.81	791.67	808.20	808.27
$\mathrm{BP}_{\mathrm{SN}}$	387.59	512.28	574.57	600.22	790.83	807.29	807.30
LB_{W}	168.41	188.08	190.01	192.36	226.12	226.89	227.50
BP_{W}	168.35	187.10	189.93	192.29	225.10	226.76	227.39
$\mathrm{p}_{\mathrm{W}}^{\mathrm{LB}}$	0.3627	0.3757	0.3802	0.3825	0.3320	0.3447	0.3526
$\mathrm{p}_{\mathrm{W}}^{\mathrm{BP}}$	0.3629	0.3759	0.3804	0.3827	0.3323	0.3450	0.3529
$\mathrm{p}_{\mathrm{S}}^{\mathrm{LB}}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\mathrm{p}_{\mathrm{S}}^{\mathrm{BP}}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lag m	15	16	17	18	19	20	21
$\hat{\rho}(m)$	0.0538	0.0073	0.0173	0.0067	−0.0027	−0.0057	0.0153
$\mathrm{LB}_{\mathrm{SN}}$	839.87	842.24	842.31	845.36	885.74	935.70	946.03
$\mathrm{BP}_{\mathrm{SN}}$	838.80	841.10	841.11	844.10	884.35	934.15	944.40
LB_{W}	266.16	266.88	270.85	271.45	271.56	271.99	275.13
BP_{W}	265.99	266.71	270.68	271.28	271.38	271.82	274.94
$\mathrm{p}_{\mathrm{W}}^{\mathrm{LB}}$	0.3105	0.3163	0.3161	0.3264	0.3289	0.3329	0.3366
$\mathrm{p}_{\mathrm{W}}^{\mathrm{BP}}$	0.3108	0.3166	0.3165	0.3268	0.3293	0.3333	0.3369
$\mathrm{p}_{\mathrm{S}}^{\mathrm{LB}}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\mathrm{p}_{\mathrm{S}}^{\mathrm{BP}}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Thanks to the assumptions on the moving average polynomials b_θ and the autoregressive polynomials a_θ, the power series a_θ^{-1} and b_θ^{-1} are well defined.

Thus the functions $\epsilon_t(\theta)$ defined in (2) can be written as

$$\epsilon_t(\theta) = b_\theta^{-1}(L)a_\theta(L)(1-L)^d X_t \tag{24}$$

$$= b_\theta^{-1}(L)a_\theta(L)(1-L)^{d-d_0} a_{\theta_0}^{-1}(L)b_{\theta_0}(L)\epsilon_t \tag{25}$$

and if we denote $\gamma(\theta) = (\gamma_i(\theta))_{i\geq 0}$ the sequence of coefficients of the power series $b_\theta^{-1}(z)a_\theta(z)(1-z)^d$, we may write for all $t \in \mathbb{Z}$:

$$\epsilon_t(\theta) = \sum_{i\geq 0} \gamma_i(\theta) X_{t-i}. \tag{26}$$

In the same way, by (24) one has

$$X_t = (1-L)^{-d} a_\theta^{-1}(L) b_\theta(L) \epsilon_t(\theta)$$

and if we denote $\eta(\theta) = (\eta_i(\theta))_{i \geq 0}$ the coefficients of the power series $(1 - z)^{-d} a_\theta^{-1}(z) b_\theta(z)$ one has

$$X_t = \sum_{i \geq 0} \eta_i(\theta) \epsilon_{t-i}(\theta) \ . \tag{27}$$

We strength the fact that $\gamma_0(\theta) = \eta_0(\theta) = 1$ for all θ.

For large j, [26] have shown that uniformly in θ the sequences $\gamma(\theta)$ and $\eta(\theta)$ satisfy

$$\frac{\partial^k \gamma_j(\theta)}{\partial\theta_{i_1} \cdots \partial\theta_{i_k}} = \mathrm{O}\left(j^{-1-d} \left\{ \log(j) \right\}^k \right), \text{ for } k = 0, 1, 2, 3, \tag{28}$$

and

$$\frac{\partial^k \eta_j(\theta)}{\partial\theta_{i_1} \cdots \partial\theta_{i_k}} = \mathrm{O}\left(j^{-1+d} \left\{ \log(j) \right\}^k \right), \text{ for } k = 0, 1, 2, 3. \tag{29}$$

One difficulty that has to be addressed is that (26) includes the infinite past $(X_{t-i})_{i \geq 0}$ whereas only a finite number of observations $(X_t)_{1 \leq t \leq n}$ are available to compute the estimators defined in (4). The simplest solution is truncation which amounts to setting all unobserved values equal to zero. Thus, for all $\theta \in \Theta$ and $1 \leq t \leq n$ one defines

$$\tilde{\epsilon}_t(\theta) = \sum_{i=0}^{t-1} \gamma_i(\theta) X_{t-i} = \sum_{i \geq 0} \gamma_i^t(\theta) X_{t-i} \tag{30}$$

where the truncated sequence $\gamma^t(\theta) = (\gamma_i^t(\theta))_{i \geq 0}$ is defined by

$$\gamma_i^t(\theta) = \begin{cases} \gamma_i(\theta) & \text{if } \ 0 \leq i \leq t-1 \ , \\ 0 & \text{otherwise.} \end{cases}$$

In the following proposition, we show that the difference between $\epsilon_t(\theta)$ and $\tilde{\epsilon}_t(\theta)$ converges almost-surely to 0 as $t \to \infty$ and this uniformly in θ. This proposition shows that the convergence of the least squares estimator $\hat{\theta}_n$ in (4) studied in [6] is not only in probability but it is almost-sure when $d_0 \in (0, 1/2)$. This last confirmation can be easily demonstrated by following line by line the proof of Theorem 1 in [20].

Proposition 8. *Let $(X_t)_{t \in \mathbb{Z}}$ be the second-order stationary process given by (1). Under the standard assumptions of invertibility and identifiability on the autoregressive polynomial a and the moving-average polynomial b, we have almost-surely*

$$\lim_{t \to \infty} \sup_{\theta \in \Theta_\delta} \left| \epsilon_t(\theta) - \tilde{\epsilon}_t(\theta) \right| = 0. \tag{31}$$

Proof. From (26) and (30), it can be readily shown that for all $\theta \in \Theta_\delta$ and any $t \in \mathbb{Z}$,

$$\begin{aligned}\epsilon_t(\theta) - \tilde{\epsilon}_t(\theta) &= \sum_{j\geq 0} \gamma_j(\theta) X_{t-j} - \sum_{j=0}^{t-1} \gamma_j(\theta) X_{t-j} \\ &= \sum_{j\geq t} \gamma_j(\theta) X_{t-j} \\ &= \sum_{k\geq 0} \gamma_{t+k}(\theta) X_{-k}.\end{aligned}$$

Recall that for any sequence $(Y_n)_{n\geq 0}$ of random variables it holds that

$$Y_n \xrightarrow[n\to\infty]{\text{a.s.}} Y \Leftrightarrow \sup_{k\geq n} |Y_k - Y| \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Hence $\sup_{\theta\in\Theta_\delta} |\epsilon_t(\theta) - \tilde{\epsilon}_t(\theta)|$ converges almost-surely to 0 as soon as

$$\sup_{k\geq t} \sup_{\theta\in\Theta_\delta} |\epsilon_k(\theta) - \tilde{\epsilon}_k(\theta)|$$

converges in probability to 0. By (28), one has for all $\beta > 0$ and for large t,

$$\begin{aligned}\mathbb{P}\left(\sup_{k\geq t} \sup_{\theta\in\Theta_\delta} |\epsilon_k(\theta) - \tilde{\epsilon}_k(\theta)| > \beta\right) &= \mathbb{P}\left(\sup_{k\geq t} \sup_{\theta\in\Theta_\delta} \left|\sum_{j\geq 0} \gamma_{k+j}(\theta) X_{-j}\right| > \beta\right) \\ &\leq \mathbb{P}\left(\sum_{j\geq 0} \sup_{k\geq t} \sup_{\theta\in\Theta_\delta} |\gamma_{k+j}(\theta)|\, |X_{-j}| > \beta\right) \\ &\leq \frac{K}{\beta}\left(\sup_{t\in\mathbb{Z}} \mathbb{E}\,|X_t|\right) \sum_{j\geq 0} \left(\frac{1}{t+j}\right)^{1+d_1} \\ &\leq \frac{K \mathrm{Var}(X_1)}{\beta d_1} (t-1)^{-d_1} \xrightarrow[t\to\infty]{} 0,\end{aligned}$$

which completes the proof of the convergence in (31). □

Since our assumptions are made on the noise in (1), it will be useful to express the random variables $\epsilon_t(\theta)$ and its partial derivatives with respect to θ, as a function of $(\epsilon_{t-i})_{i\geq 0}$.

From (25), there exists a sequence $\lambda(\theta) = (\lambda_i(\theta))_{i\geq 0}$ such that

$$\epsilon_t(\theta) = \sum_{i=0}^{\infty} \lambda_i(\theta)\, \epsilon_{t-i} \tag{32}$$

where the sequence $\lambda(\theta)$ is given by the sequence of the coefficients of the power series $b_\theta^{-1}(z) a_\theta(z)(1-z)^{d-d_0} a_{\theta_0}^{-1}(z) b_{\theta_0}(z)$. Consequently $\lambda(\theta) = \gamma(\theta) * \eta(\theta_0)$ or, equivalently,

$$\lambda_i(\theta) = \sum_{j=0}^{i} \gamma_j(\theta)\eta_{i-j}(\theta_0). \tag{33}$$

As in [30], it can be shown using Stirling's approximation that there exists a positive constant K such that

$$\sup_{\theta\in\Theta_\delta} |\lambda_i(\theta)| \le K \sup_{d\in[d_1,d_2]} i^{-1-(d-d_0)} \le K i^{-1-(d_1-d_0)} \; . \tag{34}$$

Equation (32) and Inequality (34) imply that for all $\theta \in \Theta$ the random variable $\epsilon_t(\theta)$ belongs to $\mathbb{L}^2$, that the sequence $(\epsilon_t(\theta))_t$ is an ergodic sequence and that for all $t \in \mathbb{Z}$ the function $\epsilon_t(\cdot)$ is a continuous function. We proceed in the same way as regard to the derivatives of $\epsilon_t(\theta)$. More precisely, for any $\theta \in \Theta$, $t \in \mathbb{Z}$ and $1 \le k, l \le p+q+1$ there exists sequences $\dot{\lambda}_k(\theta) = (\dot{\lambda}_{i,k}(\theta))_{i\ge1}$ and $\ddot{\lambda}_{k,l}(\theta) = (\ddot{\lambda}_{i,k,l}(\theta))_{i\ge1}$ such that

$$\frac{\partial\epsilon_t(\theta)}{\partial\theta_k} = \sum_{i=1}^{\infty} \dot{\lambda}_{i,k}\,(\theta)\,\epsilon_{t-i} \tag{35}$$

$$\frac{\partial^2\epsilon_t(\theta)}{\partial\theta_k\partial\theta_l} = \sum_{i=1}^{\infty} \ddot{\lambda}_{i,k,l}\,(\theta)\,\epsilon_{t-i}. \tag{36}$$

Of course it holds that $\dot{\lambda}_k(\theta) = \frac{\partial\gamma(\theta)}{\partial\theta_k} * \eta(\theta_0)$ and $\ddot{\lambda}_{k,l}(\theta) = \frac{\partial^2\gamma(\theta)}{\partial\theta_k\partial\theta_l} * \eta(\theta_0)$.

Similarly we have

$$\tilde{\epsilon}_t(\theta) = \sum_{i=0}^{\infty} \lambda_i^t\,(\theta)\,\epsilon_{t-i}, \tag{37}$$

$$\frac{\partial\tilde{\epsilon}_t(\theta)}{\partial\theta_k} = \sum_{i=1}^{\infty} \dot{\lambda}_{i,k}^t\,(\theta)\,\epsilon_{t-i}, \tag{38}$$

$$\frac{\partial^2\tilde{\epsilon}_t(\theta)}{\partial\theta_k\partial\theta_l} = \sum_{i=1}^{\infty} \ddot{\lambda}_{i,k,l}^t\,(\theta)\,\epsilon_{t-i}, \tag{39}$$

where $\lambda^t(\theta) = \gamma^t(\theta) * \eta(\theta_0)$, $\dot{\lambda}_k^t(\theta) = \frac{\partial\gamma^t(\theta)}{\partial\theta_k} * \eta(\theta_0)$ and $\ddot{\lambda}_{k,l}^t(\theta) = \frac{\partial^2\gamma^t(\theta)}{\partial\theta_k\partial\theta_l} * \eta(\theta_0)$.

In order to handle the truncation error $\epsilon_t(\theta) - \tilde{\epsilon}_t(\theta)$, one needs some information on the sequence $\lambda(\theta) - \lambda^t(\theta)$. This is the purpose on the following lemma.

Lemma 1. *For $2 \le r \le \infty$, $1 \le k \le p+q+1$ and $\theta \in \Theta$, we have*

$$\|\, \lambda\,(\theta) - \lambda^t\,(\theta)\, \|_{\ell^r} = \mathrm{O}\left(t^{-1+\frac{1}{r}-(d-d_0)}\right)$$

$$\|\, \dot{\lambda}_k\,(\theta) - \dot{\lambda}_k^t\,(\theta)\, \|_{\ell^r} = \mathrm{O}\left(t^{-1+\frac{1}{r}-(d-d_0)}\right),$$

and

$$\|\, \ddot{\lambda}_{k,l}\,(\theta) - \ddot{\lambda}_{k,l}^t\,(\theta)\, \|_{\ell^r} = \mathrm{O}\left(t^{-1+\frac{1}{r}-(d-d_0)}\right).$$

Proof. We have

$$\lambda(\theta) - \lambda^t(\theta) = \big(\gamma(\theta) - \gamma^t(\theta)\big) * \eta(\theta_0).$$

In view of (29), the sequence $\eta(\theta_0)$ belongs to ℓ^q for any $q > 1/(1-d_0)$. Young's inequality for convolution yields that for all $r \geq 2$,

$$\| \lambda(\theta) - \lambda^t(\theta) \|_{\ell^r} \leq \| \gamma(\theta) - \gamma^t(\theta) \|_{\ell^p} \| \eta(\theta_0) \|_{\ell^q} \tag{40}$$

with $q = (1-(d_0+\beta))^{-1} > 1/(1-d_0)$ and $p = r/(1+r(d_0+\beta))$, for some $\beta > 0$ arbitrary and sufficiently small. Thus there exists K such that $\| \eta(\theta_0) \|_{\ell^q} \leq K$. Since for any $j \geq 0$,

$$\gamma_j(\theta) - \gamma_j^t(\theta) = \begin{cases} 0 & \text{if } 0 \leq j \leq t-1 \\ \gamma_j(\theta) & \text{otherwise,} \end{cases}$$

we obtain using (28) that

$$\begin{aligned} \| \lambda(\theta) - \lambda^t(\theta) \|_{\ell^r} &\leq K \left(\sum_{k=0}^{\infty} \left| \gamma_k(\theta) - \gamma_k^t(\theta) \right|^p \right)^{1/p} \\ &\leq K \left(\sum_{k=t}^{\infty} |\gamma_k(\theta)|^p \right)^{1/p} \\ &\leq K \left(\sum_{k=t}^{\infty} \frac{1}{k^{p+pd}} \right)^{1/p} \\ &\leq K \left(\int_t^{\infty} \frac{1}{x^{p+pd}} \mathrm{dx} \right)^{1/p} \\ &\leq K t^{-1-d+\frac{1}{p}} \\ &\leq K t^{-1+\frac{1}{r}-(d-d_0)+\beta}, \end{aligned}$$

where the constant K varies from line to line. The conclusion follows by tending β to 0.

The other two points of the lemma are shown in the same way as the first. This is because from (28), the coefficients $\partial\gamma_j(\theta)/\partial\theta_k$ and $\partial^2\gamma_j(\theta)/\partial\theta_{k_1}\partial\theta_{k_2}$ are equal to $\mathrm{O}(j^{-1-d+\zeta})$ for any small enough $\zeta > 0$. The proof of the lemma is then complete. □

Remark 5. The above lemma implies that the sequence $\dot{\lambda}_k(\theta_0) - \dot{\lambda}^t{}_k(\theta_0)$ is bounded and more precisely there exists K such that

$$\sup_{j \geq 1} \left| \dot{\lambda}_{j,k}(\theta_0) - \dot{\lambda}^t{}_{j,k}(\theta_0) \right| \leq \frac{K}{t}, \tag{41}$$

for any $t \geq 1$ and any $1 \leq k \leq p+q+1$.

Lemma 2. *For any* $2 \leq r \leq \infty$, $1 \leq k \leq p+q+1$ *and* $\theta \in \Theta$, *there exists a constant* K *such that*

$$\| \lambda^t(\theta) \|_{\ell^r} \leq K \quad \textit{and} \quad \| \dot{\lambda}_k^t(\theta) \|_{\ell^r} \leq K.$$

5.2. Proof of Proposition 1

First we remark that the asymptotic normality of the joint distribution of $\sqrt{n}(\hat{\theta}'_n - \theta'_0, \gamma'_m)'$ can be established along the same lines as the proof of Theorem 2 in [6]. The detailed proof is omitted. From (6) and (11) we have

$$\sqrt{n}\begin{pmatrix}\hat{\theta}_n - \theta_0 \\ \gamma_m\end{pmatrix} = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\begin{pmatrix}-2J^{-1}(\theta_0)\epsilon_t\frac{\partial}{\partial\theta}\epsilon_t(\theta_0) \\ (\epsilon_{t-1},\dots,\epsilon_{t-m})'\epsilon_t\end{pmatrix} + \begin{pmatrix}\mathrm{o}_{\mathbb{P}}(1) \\ \mathbf{0}_m\end{pmatrix}$$
$$= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}U_t + \mathrm{o}_{\mathbb{P}}(1),$$

where $\mathbf{0}_m$ is the vector of $\mathbb{R}^{m\times 1}$ with zero components. It is clear that U_t is a measurable function of $\epsilon_t, \epsilon_{t-1}, \dots$ Thus by using the same arguments as in [6] (see proof of Theorem 2), the central limit theorem (CLT) for strongly mixing processes $(U_t)_{t\in\mathbb{Z}}$ of [27] implies that $(1/\sqrt{n})\sum_{t=1}^{n}U_t$ has a limiting normal distribution with mean 0 and covariance matrix Ξ.

For $i \geq 1$, we denote $\Lambda_i(\theta_0) = (\dot{\lambda}_{i,1}(\theta_0), \dots, \dot{\lambda}_{i,p+q+1}(\theta_0))'$. From (35) we deduce that

$$\frac{\partial\epsilon_t(\theta_0)}{\partial\theta} = \sum_{i=1}^{\infty}\Lambda_i(\theta_0)\epsilon_{t-i}. \tag{42}$$

In view of (11) and (42), by applying the CLT for mixing processes we directly obtain

$$\Sigma_{\hat{\theta}} = \lim_{n\to\infty}\mathrm{Var}\left(2J^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\epsilon_t\frac{\partial}{\partial\theta}\epsilon_t(\theta_0)\right) := J^{-1}IJ^{-1}$$
$$= 4J^{-1}\sum_{\ell,\ell'=1}^{\infty}\Lambda_\ell(\theta_0)\Lambda'_{\ell'}(\theta_0)\sum_{h=-\infty}^{\infty}\mathbb{E}\left(\epsilon_t\epsilon_{t-\ell}\epsilon_{t-h}\epsilon_{t-\ell'-h}\right)J^{-1}$$
$$= 4J^{-1}\sum_{\ell,\ell'=1}^{\infty}\Lambda_\ell(\theta_0)\Lambda'_{\ell'}(\theta_0)\Gamma(\ell,\ell')J^{-1},$$

which gives the first block of the asymptotic covariance matrix of Proposition 1.

By the stationarity of $(\epsilon_t)_{t\in\mathbb{Z}}$ and Lebesgue's dominated convergence theorem, we obtain the (ℓ,ℓ')-th entry of the matrix $\Gamma_{m,m}$:

$$\lim_{n\to\infty}\mathrm{Cov}(\sqrt{n}\gamma(\ell), \sqrt{n}\gamma(\ell')) = \lim_{n\to\infty}\frac{1}{n}\sum_{t=\ell+1}^{n}\sum_{s=\ell'+1}^{n}\mathbb{E}\left[\epsilon_t\epsilon_{t-\ell}\epsilon_s\epsilon_{s-\ell'}\right]$$
$$= \sum_{h=-\infty}^{\infty}\mathbb{E}\left[\epsilon_t\epsilon_{t-\ell}\epsilon_{t-h}\epsilon_{t-h-\ell'}\right] := \Gamma(\ell,\ell').$$

We thus have $\Gamma_{m,m} = [\Gamma(\ell,\ell')]_{1\leq\ell,\ell'\leq m}$.

Finally, by the stationarity of $(\epsilon_t)_{t\in\mathbb{Z}}$ and $(\epsilon_t\partial\epsilon_t(\theta_0)/\partial\theta)_{t\in\mathbb{Z}}$ we have

$$\begin{aligned}\mathrm{Cov}&\left(-2J^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\epsilon_t\frac{\partial}{\partial\theta}\epsilon_t(\theta_0),\sqrt{n}\gamma(\ell')\right)\\&=-2J^{-1}\frac{1}{n}\sum_{t=1}^{n}\sum_{t'=\ell'+1}^{n}\mathrm{Cov}\left(\epsilon_t\frac{\partial}{\partial\theta}\epsilon_t(\theta_0),\epsilon_{t'}\epsilon_{t'-\ell'}\right)\\&=-2J^{-1}\frac{1}{n}\sum_{h=-n+1}^{n-1}(n-|h|)\mathrm{Cov}\left(\epsilon_t\frac{\partial\epsilon_t(\theta_0)}{\partial\theta},\epsilon_{t-h}\epsilon_{t-\ell'-h}\right).\end{aligned}$$

By the dominated convergence theorem and from (42), it follows that

$$\begin{aligned}\lim_{n\to\infty}\mathrm{Cov}&\left(-2J^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\epsilon_t\frac{\partial}{\partial\theta}\epsilon_t(\theta_0),\sqrt{n}\gamma(\ell')\right)\\&=-2J^{-1}\sum_{h=-\infty}^{\infty}\mathrm{Cov}\left(\epsilon_t\frac{\partial}{\partial\theta}\epsilon_t(\theta_0),\epsilon_{t-h}\epsilon_{t-\ell'-h}\right)\\&=-2J^{-1}\sum_{j\geq1}\Lambda_j\left(\theta_0\right)\sum_{h=-\infty}^{\infty}\mathbb{E}\left(\epsilon_t\epsilon_{t-j}\epsilon_{t-h}\epsilon_{t-\ell'-h}\right)\\&=-2J^{-1}\sum_{j\geq1}\Lambda_j\left(\theta_0\right)\Gamma(j,\ell'):=\Sigma_{\hat{\theta},\gamma_m}(\cdot,\ell').\end{aligned}$$

It is clear that the existence of the above matrices is ensured by the existence of $\Gamma(\ell,\ell')$ and $\sum_{\ell,\ell'=1}^{\infty}\|\Lambda_\ell(\theta_0){\Lambda'}_{\ell'}(\theta_0)\Gamma(\ell,\ell')\|$. The proof will thus follow from Lemma 3 below. □

We now justify the existence of the $\Gamma(\ell,\ell')$ and $\sum_{\ell,\ell'=1}^{\infty}\|\Lambda_\ell(\theta_0){\Lambda'}_{\ell'}(\theta_0)\Gamma(\ell,\ell')\|$ in the following result.

Lemma 3. *Under the assumptions **(A0)** and **(A3')** with* $\tau=4$*, we have for* $(\ell,\ell')\neq(0,0)$

$$\Gamma(\ell,\ell')=\sum_{h=-\infty}^{\infty}\mathbb{E}\left(\epsilon_t\epsilon_{t-\ell}\epsilon_{t-h}\epsilon_{t-h-\ell'}\right)<\infty \tag{43}$$

$$\textit{and}\quad\sum_{\ell,\ell'=1}^{\infty}\left\|\Lambda_\ell\left(\theta_0\right){\Lambda'}_{\ell'}\left(\theta_0\right)\Gamma(\ell,\ell')\right\|<\infty. \tag{44}$$

Proof. Note that, for all $h\in\mathbb{Z}$ and all $(\ell,\ell')\neq(0,0)$ we have

$$\begin{aligned}&\left|\mathbb{E}\left[\epsilon_t\epsilon_{t-\ell}\epsilon_{t-h}\epsilon_{t-h-\ell'}\right]\right|\\&\quad\leq\left|\mathrm{cum}\left(\epsilon_t,\epsilon_{t-\ell},\epsilon_{t-h},\epsilon_{t-h-\ell'}\right)\right|+\left|\mathbb{E}\left[\epsilon_t\epsilon_{t-\ell}\right]\right|\left|\mathbb{E}\left[\epsilon_{t-h}\epsilon_{t-h-\ell'}\right]\right|\\&\qquad+\left|\mathbb{E}\left[\epsilon_t\epsilon_{t-h}\right]\right|\left|\mathbb{E}\left[\epsilon_{t-\ell}\epsilon_{t-h-\ell'}\right]\right|+\left|\mathbb{E}\left[\epsilon_t\epsilon_{t-h-\ell'}\right]\right|\left|\mathbb{E}\left[\epsilon_{t-\ell}\epsilon_{t-h}\right]\right|.\end{aligned}$$

Then, using the stationarity of $(\epsilon_t)_{t\in\mathbb{Z}}$, and under the assumptions **(A0)** and **(A3')** with $\tau = 4$ it follows that

$$\Gamma(\ell,\ell') \leq \left[\mathbb{E}\left(\epsilon_t^2\right)\right]^2 + \sum_{h=-\infty}^{\infty} \left|\mathrm{cum}\left(\epsilon_0, \epsilon_{-\ell}, \epsilon_{-h}, \epsilon_{-h-\ell'}\right)\right| \leq K$$

which proves (43). Similarly, we obtain

$$\sum_{\ell,\ell'=1}^{\infty} \left\|\Lambda_\ell\left(\theta_0\right)\Lambda'_{\ell'}\left(\theta_0\right)\Gamma(\ell,\ell')\right\| \leq \sum_{h=-\infty}^{\infty}\sum_{\ell,\ell'=1}^{\infty} \left|\mathrm{cum}\left(\epsilon_0, \epsilon_{-\ell}, \epsilon_{-h}, \epsilon_{-h-\ell'}\right)\right| + \left[\mathbb{E}\left(\epsilon_t^2\right)\right]^2 \sum_{\ell=1}^{\infty} \left\|\Lambda_\ell\left(\theta_0\right)\right\|^2 \leq K$$

where we have used Lemma 2. The conclusion follows. □

5.3. Proof of Theorem 2

The proof is divided in two steps.

5.3.1. Step 1: Taylor's expansion of $\sqrt{n}\hat{\gamma}_m$ and $\sqrt{n}\hat{\rho}_m$

The aim of this step is to prove (8) and (9). First we prove that for $h = 1,\dots,m$,

$$\sqrt{n}\hat{\gamma}(h) = \sqrt{n}\gamma(h) + \left(\mathbb{E}\left[\epsilon_{t-h}\frac{\partial}{\partial\theta'}\epsilon_t(\theta_0)\right]\right)\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) + \mathrm{o}_{\mathbb{P}}(1). \qquad (45)$$

A Taylor expansion of $(1/\sqrt{n})\sum_{t=1+h}^n \tilde{\epsilon}_t(\cdot)\tilde{\epsilon}_{t-h}(\cdot)$ around θ_0 gives

$$\begin{aligned}\sqrt{n}\hat{\gamma}(h) &= \frac{1}{\sqrt{n}}\sum_{t=1+h}^{n}\tilde{\epsilon}_t(\theta_0)\tilde{\epsilon}_{t-h}(\theta_0) + \left(\frac{1}{n}\sum_{t=1+h}^{n}\tilde{D}_t(\theta_n^*)\right)\sqrt{n}\left(\hat{\theta}_n - \theta_0\right)\\ &= \sqrt{n}\gamma(h) + \left(\mathbb{E}\left[D_t(\theta_0)\right]\right)\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) + R_{n,h,1} + R_{n,h,2} + R_{n,h,3},\end{aligned}$$

where

$$\begin{aligned}\tilde{D}_t(\theta) &= \frac{\partial\tilde{\epsilon}_t(\theta)}{\partial\theta'}\tilde{\epsilon}_{t-h}(\theta) + \tilde{\epsilon}_t(\theta)\frac{\partial\tilde{\epsilon}_{t-h}(\theta)}{\partial\theta'},\\ D_t(\theta_0) &= \frac{\partial\epsilon_t(\theta_0)}{\partial\theta'}\epsilon_{t-h} + \epsilon_t\frac{\partial\epsilon_{t-h}(\theta_0)}{\partial\theta'},\\ R_{n,h,1} &= \frac{1}{\sqrt{n}}\sum_{t=1+h}^{n}\left\{\tilde{\epsilon}_t(\theta_0)\tilde{\epsilon}_{t-h}(\theta_0) - \epsilon_t(\theta_0)\epsilon_{t-h}(\theta_0)\right\},\\ R_{n,h,2} &= \left(\frac{1}{n}\sum_{t=1+h}^{n}\left(\tilde{D}_t(\theta_n^*) - D_t(\theta_0)\right)\right)\sqrt{n}\left(\hat{\theta}_n - \theta_0\right),\end{aligned}$$

$$R_{n,h,3} = \left(\frac{1}{n} \sum_{t=1+h}^{n} D_t(\theta_0) - \mathbb{E}\left[D_t(\theta_0)\right] \right) \sqrt{n} \left(\hat{\theta}_n - \theta_0 \right),$$

and where θ_n^* is between $\hat{\theta}_n$ and θ_0. Using the orthogonality between ϵ_t and any linear combination of the past values of ϵ_t (in particular $\partial \epsilon_{t-h}/\partial \theta$), we have

$$\sqrt{n}\hat{\gamma}(h) = \sqrt{n}\gamma(h) + \left(\mathbb{E}\left[\epsilon_{t-h} \frac{\partial}{\partial \theta'} \epsilon_t(\theta_0) \right] \right) \sqrt{n} \left(\hat{\theta}_n - \theta_0 \right) + R_{n,h,1} + R_{n,h,2} + R_{n,h,3}. \tag{46}$$

Thus, to obtain (45), we just need to prove that in (46) the sequences of random variables $(R_{n,h,1})_{n\geq 1}$, $(R_{n,h,2})_{n\geq 1}$ and $(R_{n,h,3})_{n\geq 1}$ converge in probability to 0.

One of the three above term is easy to handle. Indeed, by the ergodic theorem, we have $n^{-1} \sum_{t=1+h}^{n} D_t(\theta_0) - \mathbb{E}\left[D_t(\theta_0)\right] \to 0$ almost-surely as $n \to \infty$. Thus using the tightness of the sequence $(\sqrt{n}(\hat{\theta}_n - \theta_0))_{n\geq 1}$, we deduce that $R_{n,h,3} = \mathrm{o}_{\mathbb{P}}(1)$.

The proof of (45) will thus follow from Lemmas 4 and 5 in which the two others terms $R_{n,h,1}$ and $R_{n,h,2}$ are discussed. These lemmas are stated and proved hereafter (see Subsections 5.3.3 and 5.3.4).

We now remark that in Equation (45), $\mathbb{E}[\epsilon_{t-h}(\partial \epsilon_t(\theta_0)/\partial \theta')]$ is the line h of the matrix $\Psi_m \in \mathbb{R}^{m\times(p+q+1)}$ defined by (7). So for $h = 1, \dots, m$, Equation (45) becomes

$$\sqrt{n}\hat{\gamma}_m = \left(\sqrt{n}\hat{\gamma}(1), \dots, \sqrt{n}\hat{\gamma}(m)\right)' = \sqrt{n}\gamma_m + \Psi_m \sqrt{n} \left(\hat{\theta}_n - \theta_0 \right) + \mathrm{o}_{\mathbb{P}}(1).$$

Therefore the Taylor expansion (8) of $\hat{\gamma}_m$ is proved.

Now, it is clear that the asymptotic distribution of the residual autocovariances $\sqrt{n}\hat{\gamma}_m$ is related to the asymptotic behavior of $\sqrt{n}(\hat{\theta}_n' - \theta_0', \gamma_m')'$ obtained in Subsection 5.2. We come back to the vector $\hat{\rho}_m = (\hat{\rho}(1), \dots, \hat{\rho}(m))'$. Note that from (45), we have $\sqrt{n}(\hat{\gamma}(0) - \gamma(0)) = \mathrm{o}_{\mathbb{P}}(1)$. Applying the CLT for mixing processes (see [27]) to the process $(\epsilon_t^2)_{t\in\mathbb{Z}}$, we obtain

$$\sqrt{n}\left(\hat{\sigma}_\epsilon^2 - \sigma_\epsilon^2\right) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(\epsilon_t^2 - \mathbb{E}[\epsilon_t^2]\right) + \mathrm{o}_{\mathbb{P}}(1) \xrightarrow[n\to\infty]{\text{in law}} \mathcal{N}\left(0, \sum_{h=-\infty}^{\infty} \mathrm{Cov}\left(\epsilon_t^2, \epsilon_{t-h}^2\right)\right).$$

So we have $\sqrt{n}(\hat{\sigma}_\epsilon^2 - \sigma_\epsilon^2) = \mathrm{O}_{\mathbb{P}}(1)$ and $\sqrt{n}(\gamma(0) - \sigma_\epsilon^2) = \mathrm{O}_{\mathbb{P}}(1)$. Now, using (14) and the ergodic theorem, we have

$$n\left(\frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} - \frac{\hat{\gamma}(h)}{\sigma_\epsilon^2}\right) = \sqrt{n}\hat{\gamma}(h) \frac{\sqrt{n}\left(\sigma_\epsilon^2 - \hat{\gamma}(0)\right)}{\sigma_\epsilon^2 \hat{\gamma}(0)} = \mathrm{O}_{\mathbb{P}}(1),$$

which means that $\sqrt{n}\hat{\rho}(h) = \sqrt{n}\hat{\gamma}(h)/\sigma_\epsilon^2 + \mathrm{O}_{\mathbb{P}}(n^{-1/2})$. For $h = 1, \dots, m$, it follows that

$$\sqrt{n}\hat{\rho}_m = \frac{\sqrt{n}\hat{\gamma}_m}{\sigma_\epsilon^2} + \mathrm{o}_{\mathbb{P}}(1),$$

and the Taylor expansion (9) of $\hat{\rho}_m$ is proved. This ends our first step.

The next step deals with the asymptotic distributions of $\sqrt{n}\hat{\gamma}_m$ and $\sqrt{n}\hat{\rho}_m$.

5.3.2. Step 2: asymptotic distributions of $\sqrt{n}\hat{\gamma}_m$ and $\sqrt{n}\hat{\rho}_m$

The joint asymptotic distribution of $\sqrt{n}\gamma_m$ and $\sqrt{n}(\hat{\theta}_n - \theta_0)$ shows that $\sqrt{n}\hat{\gamma}_m$ has a limiting normal distribution with mean zero and covariance matrix

$$\begin{aligned}\lim_{n\to\infty} \mathrm{Var}\left(\sqrt{n}\hat{\gamma}_m\right) &= \lim_{n\to\infty} \mathrm{Var}\left(\sqrt{n}\gamma_m\right) + \Psi_m \lim_{n\to\infty} \mathrm{Var}\left(\sqrt{n}(\hat{\theta}_n - \theta_0)\right)\Psi_m' \\ &+ \Psi_m \lim_{n\to\infty} \mathrm{Cov}\left(\sqrt{n}(\hat{\theta}_n - \theta_0), \sqrt{n}\gamma_m\right) \\ &+ \lim_{n\to\infty} \mathrm{Cov}\left(\sqrt{n}\gamma_m, \sqrt{n}(\hat{\theta}_n - \theta_0)\right)\Psi_m' \\ &= \Gamma_{m,m} + \Psi_m\Sigma_{\hat{\theta}}\Psi_m' + \Psi_m\Sigma_{\hat{\theta},\gamma_m} + \Sigma'_{\hat{\theta},\gamma_m}\Psi_m'.\end{aligned}$$

Consequently, we have

$$\lim_{n\to\infty} \mathrm{Var}\left(\sqrt{n}\hat{\rho}_m\right) = \lim_{n\to\infty} \mathrm{Var}\left(\sqrt{n}\frac{\hat{\gamma}_m}{\sigma_\epsilon^2}\right) = \frac{1}{\sigma_\epsilon^4}\Sigma_{\hat{\gamma}_m}.$$

This ends our second step and the proof is completed. □

In the following, we justify the convergence of $R_{n,h,1}$, $R_{n,h,2}$.

5.3.3. Step 3: convergence of $R_{n,h,1}$

Lemma 4. *Under the assumptions of Theorem 2, the sequence of random variables*

$$R_{n,h,1} = \frac{1}{\sqrt{n}}\sum_{t=1+h}^{n}\left\{\tilde{\epsilon}_t(\theta_0)\tilde{\epsilon}_{t-h}(\theta_0) - \epsilon_t(\theta_0)\epsilon_{t-h}(\theta_0)\right\} \tag{47}$$

tends to zero in probability as $n \to \infty$.

Proof. Throughout this proof, $\theta = (\theta_1, \dots, \theta_{p+q}, d)' \in \Theta_\delta$ is such that $d_0 < d \leq d_2$ where d_2 is the upper bound of the support of the long-range parameter d_0. Let

$$R^1_{n,h,1} = \frac{1}{\sqrt{n}}\sum_{t=1+h}^{n}\left\{\tilde{\epsilon}_t(\theta_0) - \epsilon_t(\theta_0)\right\}\tilde{\epsilon}_{t-h}(\theta_0) \tag{48}$$

and

$$R^2_{n,h,1} = \frac{1}{\sqrt{n}}\sum_{t=1+h}^{n}\epsilon_t(\theta_0)\left\{\tilde{\epsilon}_{t-h}(\theta_0) - \epsilon_{t-h}(\theta_0)\right\}. \tag{49}$$

The lemma will be proved as soon as we show that $R^1_{n,h,1}$ and $R^2_{n,h,1}$ tend to zero in probability when $n \to \infty$.

Proof of the convergence in probability of $R^1_{n,h,1}$ The arguments follow the one of Lemma 5 in [6] in a simpler context. The proof is quite long so we divide it in four steps.

$\diamond$ *Step 1: preliminaries* We have

$$\begin{aligned}
R^1_{n,h,1} &= \frac{1}{\sqrt{n}} \sum_{t=1+h}^{n} \left\{\tilde{\epsilon}_t(\theta_0) - \tilde{\epsilon}_t(\theta)\right\} \tilde{\epsilon}_{t-h}(\theta_0) \\
&\quad + \frac{1}{\sqrt{n}} \sum_{t=1+h}^{n} \left\{\tilde{\epsilon}_t(\theta) - \epsilon_t(\theta)\right\} \tilde{\epsilon}_{t-h}(\theta_0) \\
&\quad + \frac{1}{\sqrt{n}} \sum_{t=1+h}^{n} \left\{\epsilon_t(\theta) - \epsilon_t(\theta_0)\right\} \tilde{\epsilon}_{t-h}(\theta_0) \\
&= \omega_{n,h,1}(\theta) + \omega_{n,h,2}(\theta) + \omega_{n,h,3}(\theta),
\end{aligned}$$

where

$$\begin{aligned}
\omega_{n,h,1}(\theta) &= \frac{1}{\sqrt{n}} \sum_{t=1+h}^{n} \left\{\tilde{\epsilon}_t(\theta_0) - \tilde{\epsilon}_t(\theta)\right\} \tilde{\epsilon}_{t-h}(\theta_0), \\
\omega_{n,h,2}(\theta) &= \frac{1}{\sqrt{n}} \sum_{t=1+h}^{n} \left\{\tilde{\epsilon}_t(\theta) - \epsilon_t(\theta)\right\} \tilde{\epsilon}_{t-h}(\theta_0)
\end{aligned}$$

and

$$\omega_{n,h,3}(\theta) = \frac{1}{\sqrt{n}} \sum_{t=1+h}^{n} \left\{\epsilon_t(\theta) - \epsilon_t(\theta_0)\right\} \tilde{\epsilon}_{t-h}(\theta_0).$$

Therefore, if we prove that the two sequences of random variables $(\omega_{n,h,2}(\theta))_{n\geq 1}$ and $(\omega_{n,h,1}(\theta) + \omega_{n,h,3}(\theta))_{n\geq 1}$ converge in probability towards 0, then the convergence in probability of $R^1_{n,h,1}$ to zero will be true.

$\diamond$ *Step 2: convergence in probability of* $(\omega_{n,h,2}(\theta))_{n\geq 1}$ *to* 0 For all $\beta > 0$, we have

$$\begin{aligned}
\mathbb{P}\left(|\omega_{n,h,2}| \geq \beta\right) &\leq \frac{1}{\sqrt{n}\beta} \sum_{t=1+h}^{n} \mathbb{E}\left[|\tilde{\epsilon}_t(\theta) - \epsilon_t(\theta)|\, |\tilde{\epsilon}_{t-h}(\theta_0)|\right] \\
&\leq \frac{1}{\sqrt{n}\beta} \sum_{t=1+h}^{n} \left\|\tilde{\epsilon}_t(\theta) - \epsilon_t(\theta)\right\|_{\mathbb{L}^2} \left\|\tilde{\epsilon}_{t-h}(\theta_0)\right\|_{\mathbb{L}^2}.
\end{aligned}$$

First, from (37) and using Lemma 2, we have

$$\begin{aligned}
\left\|\tilde{\epsilon}_{t-h}(\theta_0)\right\|^2_{\mathbb{L}^2} &= \mathbb{E}\left[\left(\sum_{i=0}^{\infty} \lambda_i^t\left(\theta_0\right) \epsilon_{t-i-h}\right)^2\right] \\
&= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \lambda_i^t\left(\theta_0\right) \lambda_j^t\left(\theta_0\right) \mathbb{E}\left[\epsilon_{t-i-h}\epsilon_{t-j-h}\right] + \sigma_\epsilon^2 \left\{\lambda_0^t\left(\theta_0\right)\right\}^2
\end{aligned}$$

$$= \sigma_\epsilon^2 \sum_{i=1}^{\infty} \left\{ \lambda_i^t(\theta_0) \right\}^2 + \sigma_\epsilon^2$$
$$\leq K. \tag{50}$$

In view of (32), (37) and (50), we may write

$$\begin{aligned}
&\mathbb{P}\left(|\omega_{n,h,2}(\theta)| \geq \beta\right) \\
&\quad \leq \frac{K}{\beta\sqrt{n}} \sum_{t=1+h}^{n} \left(\mathbb{E}\left[\left(\tilde{\epsilon}_{t-h}(\theta) - \epsilon_{t-h}(\theta) \right)^2 \right] \right)^{1/2} \\
&\quad \leq \frac{K}{\beta\sqrt{n}} \sum_{t=1+h}^{n} \left(\sum_{i\geq 0} \sum_{j \geq 0} \left(\lambda_i^t(\theta) - \lambda_i(\theta) \right) \left(\lambda_j^t(\theta) - \lambda_j(\theta) \right) \mathbb{E}\left[\epsilon_{t-i-h}\epsilon_{t-j-h} \right] \right)^{1/2} \\
&\quad \leq \frac{\sigma_\epsilon K}{\beta\sqrt{n}} \sum_{t=1}^{n} \left(\sum_{i\geq 0} \left(\lambda_i^t(\theta) - \lambda_i(\theta) \right)^2 \right)^{1/2} \\
&\quad \leq \frac{\sigma_\epsilon K}{\beta\sqrt{n}} \sum_{t=1}^{n} \left\| \lambda(\theta) - \lambda^t(\theta) \right\|_{\ell^2}.
\end{aligned}$$

We use Lemma 1, the fact that $d > d_0$ and the fractional version of Cesàro's Lemma[2] to obtain

$$\mathbb{P}\left(|\omega_{n,h,2}(\theta)| \geq \beta\right) \leq \frac{\sigma_\epsilon K}{\beta} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{1}{t^{1/2+(d-d_0)}} \xrightarrow[n\to\infty]{} 0.$$

This proves the expected convergence in probability.

⋄ *Step 3: convergence of* $(\omega_{n,h,1}(\theta)+\omega_{n,h,3}(\theta))_{n\geq 1}$ Note now that, for all $n \geq 1$, we have

$$\omega_{n,h,1}(\theta) + \omega_{n,h,3}(\theta) = \frac{1}{\sqrt{n}} \sum_{t=1+h}^{n} \Big\{ \left(\epsilon_t(\theta) - \tilde{\epsilon}_t(\theta) \right) - \left(\epsilon_t(\theta_0) - \tilde{\epsilon}_t(\theta_0) \right) \Big\} \tilde{\epsilon}_{t-h}(\theta_0).$$

By the mean value theorem, there exists $0 < c_\omega < 1$ such that

$$\begin{aligned}
&\Big| \left(\epsilon_t(\theta) - \tilde{\epsilon}_t(\theta) \right) - \left(\epsilon_t(\theta_0) - \tilde{\epsilon}_t(\theta_0) \right) \Big| \\
&\qquad \leq \left\| \frac{\partial(\epsilon_t - \tilde{\epsilon}_t)}{\partial\theta} ((1-c_\omega)\theta + c_\omega\theta_0) \right\|_{\mathbb{R}^{p+q+1}} \|\theta - \theta_0\|_{\mathbb{R}^{p+q+1}}.
\end{aligned} \tag{51}$$

[2]Recall that the fractional version of Cesàro's Lemma states that for $(h_t)_t$ a sequence of positive real numbers, $\kappa > 0$ and $c \geq 0$ we have

$$\lim_{t\to\infty} h_t t^{1-\kappa} = |\kappa|\, c \Rightarrow \lim_{n\to\infty} \frac{1}{n^\kappa} \sum_{t=0}^{n} h_t = c.$$

Following the same method as the one in Step 2, we obtain

$$\begin{aligned}
&\mathbb{E}\Big((\epsilon_t(\theta) - \tilde{\epsilon}_t(\theta)) - (\epsilon_t(\theta_0) - \tilde{\epsilon}_t(\theta_0)) \Big)^2 \\
&\qquad \le \|\theta - \theta_0\|^2_{\mathbb{R}^{p+q+1}} \sum_{k=1}^{p+q+1} \mathbb{E}\left[\left| \frac{\partial(\epsilon_t - \tilde{\epsilon}_t)}{\partial\theta_k}((1-c_\omega)\theta + c_\omega\theta_0) \right|^2 \right] \\
&\qquad \le \|\theta - \theta_0\|^2_{\mathbb{R}^{p+q+1}} \sum_{k=1}^{p+q+1} \sup_\theta \mathbb{E}\left[\left| \frac{\partial(\epsilon_t - \tilde{\epsilon}_t)}{\partial\theta_k}(\theta) \right|^2 \right] \\
&\qquad \le \|\theta - \theta_0\|^2_{\mathbb{R}^{p+q+1}} \sum_{k=1}^{p+q+1} \sigma_\epsilon^2 \sup_\theta \left\| (\dot{\lambda}_k - \dot{\lambda}_k^t)(\theta) \right\|^2_{\ell^2} \\
&\qquad \le K \|\theta - \theta_0\|^2_{\mathbb{R}^{p+q+1}} \sup_{d;d_0 \le d \le d_2} \left(\frac{1}{t^{1/2+(d-d_0)}} \right)^2 \\
&\qquad \le K \|\theta - \theta_0\|^2_{\mathbb{R}^{p+q+1}} \frac{1}{t}, \qquad (52)
\end{aligned}$$

where we have used the fact that the function

$$\theta \mapsto \mathbb{E}\left[\left| \frac{\partial(\epsilon_t - \tilde{\epsilon}_t)}{\partial\theta_k}(\theta) \right|^2 \right]$$

is bounded and continuous. By (50) and (52), it follows that

$$\mathbb{P}\left(|\omega_{n,h,1}(\theta) + \omega_{n,h,3}(\theta)| \ge \beta \right) \le \frac{K}{\beta} \|\theta - \theta_0\|_{\mathbb{R}^{p+q+1}} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{1}{t^{1/2}}$$

and the fractional version of Cesàro's Lemma implies

$$\lim_{n\to\infty} \mathbb{P}\left(|\omega_{n,h,1}(\theta) + \omega_{n,h,3}(\theta)| \ge \beta \right) \le \frac{K}{\beta} \|\theta - \theta_0\|_{\mathbb{R}^{p+q+1}}. \qquad (53)$$

⋄ *Step 4: end of the proof of the convergence in probability of* $R^1_{n,h,1}$ *to* 0 For any $\varepsilon > 0$, we choose θ such that $(K/\beta)\|\theta - \theta_0\|_{\mathbb{R}^{p+q+1}} \le \varepsilon$. Then, from (53), there exists n_0 such that for all $n \ge n_0$,

$$\mathbb{P}\left(|\omega_{n,h,1}(\theta) + \omega_{n,h,3}(\theta)| \ge \beta \right) \le \varepsilon.$$

By Step 2, one also has for $n \ge n_0$

$$\mathbb{P}\left(|\omega_{n,h,2}(\theta)| \ge \beta \right) \le \varepsilon.$$

Therefore, for all $n \ge n_0$,

$$\mathbb{P}\left(\left| R^1_{n,h,1} \right| \ge 2\beta \right) \le \mathbb{P}\left(|\omega_{n,h,1}(\theta) + \omega_{n,h,3}(\theta)| \ge \beta \right) + \mathbb{P}\left(|\omega_{n,h,2}(\theta)| \ge \beta \right) \le \varepsilon$$

and the expected convergence is proved.

Proof of the convergence in probability of $R^2_{n,h,1}$ Under Assumption **(A3)** with $\tau = 2$ it follows that $\epsilon_t(\theta_0)$ belongs to $\mathbb{L}^2$. Thus the proof of the convergence in probability of $R^2_{n,h,1}$ to zero is shown in the same way as the proof of the convergence in probability of $R^1_{n,h,1}$ to 0.

Conclusion: convergence in probability of $R_{n,h,1}$ The conclusion is a consequence of the above convergences. □

5.3.4. Step 4: convergence of $R_{n,h,2}$

Lemma 5. *Under the assumptions of Theorem 2, the sequence of random variables*

$$R_{n,h,2} = \left(\frac{1}{n} \sum_{t=1+h}^{n} \left(\tilde{D}_t(\theta_n^*) - D_t(\theta_0) \right) \right) \sqrt{n} \left(\hat{\theta}_n - \theta_0 \right) \tag{54}$$

tends to zero in probability as $n \to \infty$.

Proof. Since $(\sqrt{n}(\hat{\theta}_n - \theta_0))_{n \geq 1}$ is a tight sequence, we have $\sqrt{n}(\hat{\theta}_n - \theta_0) = \mathrm{O}_{\mathbb{P}}(1)$. Hence, to prove the convergence in probability of $(R_{n,h,2})_{n\geq 1}$ to 0, it suffices to show that

$$\frac{1}{n} \sum_{t=1+h}^{n} \left(\tilde{D}_t(\theta_n^*) - D_t(\theta_0) \right) = \mathrm{o}_{\mathbb{P}}(1). \tag{55}$$

This will be proved using Lemma 1 and Cesàro's Lemma. Nevertheless, the proof is quite long so we divide it in four steps.

◇ **Step 1: preliminaries** We have

$$\frac{1}{n} \sum_{t=1+h}^{n} \left(\tilde{D}_t(\theta_n^*) - D_t(\theta_0) \right) = T_{n,h,1}(\theta_n^*) + T_{n,h,2}(\theta_n^*) + T_{n,h,3}(\theta_n^*) + T_{n,h,4}(\theta_n^*) + T_{n,h,5}(\theta_n^*),$$

where

$$T_{n,h,1}(\theta) = \frac{1}{n} \sum_{t=1+h}^{n} \frac{\partial \tilde{\epsilon}_t(\theta)}{\partial \theta'} \left(\tilde{\epsilon}_{t-h}(\theta) - \epsilon_{t-h}(\theta) \right),$$

$$T_{n,h,2}(\theta) = \frac{1}{n} \sum_{t=1+h}^{n} \left(\tilde{\epsilon}_t(\theta) - \epsilon_t(\theta) \right) \frac{\partial \tilde{\epsilon}_{t-h}(\theta)}{\partial \theta'},$$

$$T_{n,h,3}(\theta) = \frac{1}{n} \sum_{t=1+h}^{n} \left(\frac{\partial \tilde{\epsilon}_t(\theta)}{\partial \theta'} - \frac{\partial \epsilon_t(\theta)}{\partial \theta'} \right) \epsilon_{t-h}(\theta),$$

$$T_{n,h,4}(\theta) = \frac{1}{n} \sum_{t=1+h}^{n} \epsilon_t(\theta) \left(\frac{\partial \tilde{\epsilon}_{t-h}(\theta)}{\partial \theta'} - \frac{\partial \epsilon_{t-h}(\theta)}{\partial \theta'} \right) \quad and$$

$$T_{n,h,5}(\theta) = \frac{1}{n} \sum_{t=1+h}^{n} \left(D_t(\theta) - D_t(\theta_0)\right).$$

Therefore, if we prove that the five sequences of random variables $(T_{n,h,i}(\theta_n^*))_{n\geq 1}$ (for $i = 1, \dots, 5$) converge in probability towards 0, then (55) will be true.

◇ **Step 2: convergence in probability of** $(T_{n,h,1}(\theta_n^*))_{n\geq 1}$ **to** 0 For all $\beta > 0$, we have

$$\begin{aligned}
\mathbb{P}\left(\|T_{n,h,1}(\theta_n^*)\| \geq \beta\right) &\leq \frac{1}{n\beta} \sum_{t=1+h}^{n} \mathbb{E}\left[\left\|\frac{\partial \tilde{\epsilon}_t(\theta_n^*)}{\partial \theta'}\right\| \left|\tilde{\epsilon}_{t-h}(\theta_n^*) - \epsilon_{t-h}(\theta_n^*)\right|\right] \\
&\leq \frac{1}{n\beta} \sup_{\theta\in\Theta_\delta} \sum_{t=1+h}^{n} \mathbb{E}\left[\left\|\frac{\partial \tilde{\epsilon}_t(\theta)}{\partial \theta'}\right\| \left|\tilde{\epsilon}_{t-h}(\theta) - \epsilon_{t-h}(\theta)\right|\right] \\
&\leq \frac{1}{n\beta} \sup_{\theta\in\Theta_\delta} \sum_{t=1+h}^{n} \left\|\tilde{\epsilon}_{t-h}(\theta) - \epsilon_{t-h}(\theta)\right\|_{\mathbb{L}^2} \left\|\frac{\partial \tilde{\epsilon}_t(\theta)}{\partial \theta'}\right\|_{\mathbb{L}^2}.
\end{aligned}$$

First, from (37) and using Lemma 2 we have for $1 \leq k \leq p+q+1$

$$\begin{aligned}
\left\|\frac{\partial}{\partial \theta_k}\tilde{\epsilon}_t(\theta)\right\|_{\mathbb{L}^2}^2 &= \mathbb{E}\left[\left(\sum_{i=1}^{\infty} \dot{\lambda}_{i,k}^{t}(\theta)\epsilon_{t-i}\right)^2\right] \\
&= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \dot{\lambda}_{i,k}^{t}(\theta)\dot{\lambda}_{j,k}^{t}(\theta)\mathbb{E}\left[\epsilon_{t-i}\epsilon_{t-j}\right] \\
&= \sigma_\epsilon^2 \sum_{i=1}^{\infty} \left\{\dot{\lambda}_{i,k}^{t}(\theta)\right\}^2 \\
&\leq K. \qquad (56)
\end{aligned}$$

In view of (32), (37), (56) and following the same approach used in Step 2 of Lemma 4 we have

$$\begin{aligned}
&\mathbb{P}\left(|T_{n,h,1}(\theta_n^*)| \geq \beta\right) \\
&\quad\leq \frac{K}{\beta n} \sup_{\theta\in\Theta_\delta} \sum_{t=1+h}^{n} \left(\mathbb{E}\left[\left(\tilde{\epsilon}_{t-h}(\theta) - \epsilon_{t-h}(\theta)\right)^2\right]\right)^{1/2} \\
&\quad\leq \frac{K}{\beta n} \sup_{\theta\in\Theta_\delta} \sum_{t=1+h}^{n} \left(\sum_{i\geq 0}\sum_{j\geq 0} \left(\lambda_i^t(\theta) - \lambda_i(\theta)\right)\left(\lambda_j^t(\theta) - \lambda_j(\theta)\right)\mathbb{E}\left[\epsilon_{t-i-h}\epsilon_{t-j-h}\right]\right)^{1/2} \\
&\quad\leq \frac{\sigma_\epsilon K}{\beta n} \sup_{\theta\in\Theta_\delta} \sum_{t=1}^{n} \left(\sum_{i\geq 0} \left(\lambda_i^t(\theta) - \lambda_i(\theta)\right)^2\right)^{1/2} \\
&\quad\leq \frac{\sigma_\epsilon K}{\beta n} \sup_{\theta\in\Theta_\delta} \sum_{t=1}^{n} \left\|\lambda(\theta) - \lambda^t(\theta)\right\|_{\ell^2}.
\end{aligned}$$

We use Lemma 1, the fact that $|d_2 - d_1| < 1/2$ and the Cesàro Lemma to obtain

$$\mathbb{P}\left(|T_{n,h,1}(\theta_n^*)| \geq \beta\right) \leq \frac{\sigma_\epsilon K}{\beta} \frac{1}{n} \sum_{t=1}^{n} \frac{1}{t^{1/2+(d_2-d_1)}} \xrightarrow[n\to\infty]{} 0.$$

This proves the expected convergence in probability of $T_{n,h,1}(\theta_n^*)$.

The same calculations holds for the sequences of random variables $(T_{n,h,2}(\theta_n^*))_{n\geq 1}$, $(T_{n,h,3}(\theta_n^*))_{n\geq 1}$ and $(T_{n,h,4}(\theta_n^*))_{n\geq 1}$.

$\diamond$ **Step 3: convergence in probability of** $(T_{n,h,5}(\theta_n^*))_{n\geq 1}$ **to** 0 For $1 \leq i, j \leq p+q+1$ and in view of (26), (28), we have

$$\begin{aligned}
\sup_{\theta\in\Theta_\delta} \left| \frac{\partial}{\partial\theta_i}\epsilon_t(\theta) \frac{\partial}{\partial\theta_j}\epsilon_t(\theta) \right| &= \sup_{\theta\in\Theta_\delta} \left| \sum_{k_1,k_2\geq 1} \frac{\partial}{\partial\theta_i}\gamma_{k_1}(\theta) \frac{\partial}{\partial\theta_j}\gamma_{k_2}(\theta) X_{t-k_1} X_{t-k_2} \right| \\
&\leq \sum_{k_1,k_2\geq 1} \sup_{\theta\in\Theta_\delta} \left| \frac{\partial}{\partial\theta_i}\gamma_{k_1}(\theta) \right| \sup_{\theta\in\Theta_\delta} \left| \frac{\partial}{\partial\theta_j}\gamma_{k_2}(\theta) \right| |X_{t-k_1}| \, |X_{t-k_2}| \\
&\leq K \sum_{k_1,k_2\geq 1} \log(k_1) k_1^{-1-d_1} \log(k_2) k_2^{-1-d_1} |X_{t-k_1}| \, |X_{t-k_2}| \, .
\end{aligned}$$

Consequently, we obtain

$$\begin{aligned}
&\mathbb{E}_{\theta_0}\left[\sup_{\theta\in\Theta_\delta} \left| \frac{\partial}{\partial\theta_i}\epsilon_t(\theta) \frac{\partial}{\partial\theta_j}\epsilon_t(\theta) \right| \right] \\
&\qquad \leq K \sum_{k_1,k_2\geq 1} \log(k_1) k_1^{-1-d_1} \log(k_2) k_2^{-1-d_1} \sup_{t\in\mathbb{Z}} \mathbb{E}_{\theta_0} |X_t|^2 \leq K. \qquad (57)
\end{aligned}$$

Following the same approach used to obtain (57), we have

$$\mathbb{E}_{\theta_0}\left[\sup_{\theta\in\Theta_\delta} \left| \epsilon_t(\theta) \frac{\partial^2}{\partial\theta_i\partial\theta_j}\epsilon_t(\theta) \right| \right] < \infty. \qquad (58)$$

A Taylor expansion of $D_t(\cdot)$ around θ_0 implies that

$$\begin{aligned}
\|T_{n,h,5}(\theta_n^*)\| &\leq \frac{1}{n} \sum_{t=1}^{n} \sup_{\theta\in\Theta_\delta} \left\| \frac{\partial}{\partial\theta} D_t(\theta) \right\| \|\theta_n^{**} - \theta_0\| \\
&\leq \|\theta_n^{**} - \theta_0\| \times \frac{1}{n} \sum_{t=1}^{n} \sup_{\theta\in\Theta_\delta} \left\| \frac{\partial}{\partial\theta} D_t(\theta) \right\|
\end{aligned}$$

with θ_n^{**} between θ_0 and θ_n^*. The almost-sure convergence of $(\hat{\theta}_n - \theta_0)_{n\geq 1}$ implies taht $\theta_n^{**} - \theta_0$ tends to 0 almost-surely. From (57) and (58), it follows that

$$\begin{aligned}\mathbb{E}\left[\sup_{\theta\in\Theta_\delta}\left\|\frac{\partial}{\partial\theta}D_t(\theta)\right\|\right] &= \mathbb{E}\left[\sup_{\theta\in\Theta_\delta}\left\|\epsilon_{t-h}(\theta)\frac{\partial^2}{\partial\theta\partial\theta'}\epsilon_t(\theta)+\frac{\partial}{\partial\theta}\epsilon_{t-h}(\theta)\frac{\partial}{\partial\theta'}\epsilon_t(\theta)\right.\right.\\ &\qquad\left.\left.+\frac{\partial}{\partial\theta}\epsilon_t(\theta)\frac{\partial}{\partial\theta'}\epsilon_{t-h}(\theta)+\epsilon_t(\theta)\frac{\partial^2}{\partial\theta\partial\theta'}\epsilon_{t-h}(\theta)\right\|\right]\\ &\le K. \qquad (59)\end{aligned}$$

Equation (59), the ergodic theorem and the almost-sure convergence of $(\hat{\theta}_n - \theta_0)_{n\geq 1}$ to 0 imply that $T_{n,h,5}(\theta_n^*)$ tends to 0 almost-surely.

◇ **Step 4: end of the proof of the convergence in probability of $R_{n,h,2}$ to zero** By Step 2 and 3 we deduce that

$$R_{n,h,2} = \mathrm{o}_{\mathbb{P}}(1)$$

and the convergence in probability is proved.

The proof of the lemma is completed. □

5.4. Proof of Remark 4

We suppose that **(H1)** holds true. One may rewrite the above arguments in order to prove that there exists a nonsingular matrix D^* such that

$$\sqrt{n}(\hat{\rho}_m - \rho_m) = \frac{1}{\sigma_\epsilon^2}\sqrt{n}(\hat{\gamma}_m - \gamma_m^0) + \mathrm{o}_{\mathbb{P}}(1) \xrightarrow[n\to\infty]{\mathrm{d}} \mathcal{N}\left(0, D^*\right) . \qquad (60)$$

The matrix D^* is given by

$$D^* = \left[\Sigma_{\gamma_m^0} + C_m^* J^{-1} I J^{-1} C_m^{*\,\prime} + C_m^* \Sigma_{\hat{\theta}_n,\gamma_m^0} + \Sigma'_{\hat{\theta}_n,\gamma_m^0} C_m^{*\,\prime}\right]/\sigma_\epsilon^4,$$

where the matrices $\Sigma_{\gamma_m^0}$ and $\Sigma_{\hat{\theta}_n,\gamma_m^0}$ are obtained from the asymptotic distribution of

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{t=1}^n U_t^* &:= \frac{1}{\sqrt{n}}\sum_{t=1}^n \begin{pmatrix} -2J^{-1}\epsilon_t\frac{\partial}{\partial\theta}\epsilon_t \\ (\epsilon_t\epsilon_{t-1}-\gamma^0(1),\dots,\epsilon_t\epsilon_{t-m}-\gamma^0(m))' \end{pmatrix}\\ &\xrightarrow[n\to\infty]{\mathrm{d}} \mathcal{N}(0, \mathbb{E}\left[U_t^* U_t^{*\prime}\right]),\end{aligned}$$

with

$$\mathbb{E}\left[U_t^* U_t^{*\prime}\right] =: \begin{pmatrix} J^{-1}IJ^{-1} & \Sigma_{\hat{\theta}_n,\gamma_m^0} \\ \Sigma'_{\hat{\theta}_n,\gamma_m^0} & \Sigma_{\gamma_m^0}\end{pmatrix}.$$

For $h = 1, \dots, m$ the row h of the matrix C_m^* is given by

$$c_h^* := \mathbb{E}\left[\epsilon_{t-h}\frac{\partial\epsilon_t}{\partial\theta} + \epsilon_t\frac{\partial\epsilon_{t-h}}{\partial\theta}\right].$$

We point out the fact that under **(H1)**,

$$\mathbb{E}\left[\epsilon_t \frac{\partial \epsilon_{t-h}}{\partial \theta}\right] \neq 0$$

whereas it vanishes under **(H0)**. Thus we have

$$\frac{\partial \gamma_m}{\partial \theta} \underset{n\to\infty}{\longrightarrow} C_m^* := \begin{pmatrix} c_1^{*\prime} \\ \vdots \\ c_m^{*\prime} \end{pmatrix}.$$

Now we write

$$\begin{aligned} \sqrt{n}\hat{\rho}_m &= \sqrt{n}(\hat{\rho}_m - \rho_m) + \sqrt{n}\rho_m \\ &= \frac{1}{\sigma_\epsilon^2}\left[\sqrt{n}(\hat{\gamma}_m - \gamma_m^0) + \sqrt{n}\gamma_m^0\right] + \mathrm{o}_{\mathbb{P}}(1) \ . \end{aligned}$$

Then it holds that

$$\begin{aligned} n\hat{\rho}_m'\hat{\rho}_m &= \frac{n\hat{\gamma}_m'\hat{\gamma}_m}{\sigma_\epsilon^4} + \mathrm{o}_{\mathbb{P}}(1) \\ &= \frac{1}{\sigma_\epsilon^4}\left[n(\hat{\gamma}_m - \gamma_m^0)'(\hat{\gamma}_m - \gamma_m^0) + 2n(\hat{\gamma}_m - \gamma_m^0)'\gamma_m^0 + n{\gamma_m^0}'\gamma_m^0\right] + \mathrm{o}_{\mathbb{P}}(1) \end{aligned} \tag{61}$$

By the ergodic theroem, $(\hat{\gamma}_m - \gamma_m^0)'\gamma_m^0 = \mathrm{o}_{\mathbb{P}}(1)$. By [50, Lemma 17.1], the convergence (60) implies that

$$n(\hat{\gamma}_m - \gamma_m^0)'(\hat{\gamma}_m - \gamma_m^0) \xrightarrow[n\to\infty]{\mathrm{d}} \sum_{i=1}^m \lambda_i Z_i^2$$

where $(Z_i)_{1\le i\le m}$ are i.i.d. with $\mathcal{N}(0,1)$ laws and the λ_i's are the eigenvalues of the matrix $\sigma_\epsilon^4 D^*$. Reporting these convergences in (61), we deduce that

$$\begin{aligned} \hat{\gamma}_m'\hat{\gamma}_m &= (\hat{\gamma}_m - \gamma_m^0)'(\hat{\gamma}_m - \gamma_m^0) + 2(\hat{\gamma}_m - \gamma_m^0)'\gamma_m^0 + {\gamma_m^0}'\gamma_m^0 + \mathrm{o}_{\mathbb{P}}(1) \\ &= {\gamma_m^0}'\gamma_m^0 + \mathrm{o}_{\mathbb{P}}(1) \end{aligned}$$

and the remark is proved. □

5.5. Proof of Proposition 5

The following proofs are quite technical and are adaptations of the arguments used in [7].

To prove the invertibility of the normalized matrix C_m, we need to introduce the following notation.

Let $S_t(i)$ be the i-th component of the vector $S_t = \sum_{j=1}^{t} (\Lambda U_j - \gamma_m) \in \mathbb{R}^m$. We remark that

$$S_{t-1}(i) = S_t(i) - \sum_{k=1}^{p+q+1} \delta_{i,k}\epsilon_t \frac{\partial}{\partial\theta_k}\epsilon_t(\theta_0) - \epsilon_t\epsilon_{t-i} + \gamma(i), \tag{62}$$

where $\delta_{i,k}$ is the (i,k)-th entry of the $m \times (p+q+1)$ matrix $\Delta := -2\Psi_m J^{-1}$.

If the matrix C_m is not invertible, there exists some real constants $c_1, \dots, c_m$ not all equal to zero, such that we have

$$\sum_{i=1}^{m}\sum_{j=1}^{m} c_j C_m(j,i) c_i = \frac{1}{n^2}\sum_{t=1}^{n}\sum_{i=1}^{m}\sum_{j=1}^{m} c_j S_t(j) S_t(i) c_i = \frac{1}{n^2}\sum_{t=1}^{n}\left(\sum_{i=1}^{m} c_i S_t(i)\right)^2 = 0,$$

which implies that $\sum_{i=1}^{m} c_i S_t(i) = 0$ for all $t \geq 1$.

Then by (62), it would imply that

$$\sum_{i=1}^{m}\sum_{k=1}^{p+q+1} c_i\delta_{i,k}\epsilon_t \frac{\partial}{\partial\theta_k}\epsilon_t(\theta_0) + \sum_{i=1}^{m} c_i\epsilon_t\epsilon_{t-i} = \sum_{i=1}^{m} c_i\gamma(i). \tag{63}$$

By the ergodic Theorem, we also have $\sum_{i=1}^{m} c_i\gamma(i) \to 0$ almost-surely as n goes to infinity.

Consequently replacing this convergence in (63) implies that for all $t \geq 1$

$$\sum_{i=1}^{m}\sum_{k=1}^{p+q+1} c_i\delta_{i,k}\epsilon_t \frac{\partial}{\partial\theta_k}\epsilon_t(\theta_0) + \sum_{i=1}^{m} c_i\epsilon_t\epsilon_{t-i} = 0, \quad \text{a.s.}$$

Using (32), it yields that

$$\epsilon_t\left\{\sum_{\ell\geq 1}\left(\sum_{i=1}^{m}\sum_{k=1}^{p+q+1} c_i\delta_{i,k}\dot{\lambda}_{\ell,k}(\theta_0)\right)\epsilon_{t-\ell} + \sum_{\ell=1}^{m} c_\ell\epsilon_{t-\ell}\right\} = 0, \quad \text{a.s.}$$

Or equivalently,

$$\epsilon_t\left\{\sum_{\ell=1}^{m}\left(\sum_{i=1}^{m} c_i \sum_{k=1}^{p+q+1} \delta_{i,k}\dot{\lambda}_{\ell,k}(\theta_0) + c_\ell\right)\epsilon_{t-\ell} + \sum_{\ell\geq m+1}\left(\sum_{i=1}^{m} c_i \sum_{k=1}^{p+q+1}\delta_{i,k}\dot{\lambda}_{\ell,k}(\theta_0)\right)\epsilon_{t-\ell}\right\} = 0, \quad \text{a.s.}$$

Thanks to Assumption **(A4)**, ϵ_t has a positive density in some neighborhood of zero and then $\epsilon_t \neq 0$ almost-surely. Hence we obtain

$$\sum_{\ell=1}^{m}\left(\sum_{i=1}^{m} c_i \sum_{k=1}^{p+q+1} \delta_{i,k}\dot{\lambda}_{\ell,k}(\theta_0) + c_\ell\right)\epsilon_{t-\ell} + \sum_{\ell\geq m+1}\left(\sum_{i=1}^{m} c_i \sum_{k=1}^{p+q+1}\delta_{i,k}\dot{\lambda}_{\ell,k}(\theta_0)\right)\epsilon_{t-\ell} = 0, \quad \text{a.s.}$$

Since the variance of the innovation process in not equal to zero, we deduce that

$$\begin{cases} \sum_{i=1}^{m} c_i \sum_{k=1}^{p+q+1} \delta_{i,k} \dot{\lambda}_{\ell,k}(\theta_0) + c_\ell = 0 & \text{for all } \ell \in \{1,\dots,m\} \\ \sum_{i=1}^{m} c_i \sum_{k=1}^{p+q+1} \delta_{i,k} \dot{\lambda}_{\ell,k}(\theta_0) = 0 & \text{for all } \ell \in \{m+1,\dots\}. \end{cases}$$

Then we would have $c_1 = \cdots = c_m = 0$ which is impossible. Thus we have a contradiction and the matrix $C_m \in \mathbb{R}^{m\times m}$ is non singular. □

5.6. Proof of Theorem 6

We recall that the Skorokhod space $\mathbb{D}^{\ell}[0,1]$ is the set of $\mathbb{R}^{\ell}$-valued functions on [0,1] which are right-continuous and have left limits everywhere. It is endowed with the Skorokhod topology and the weak convergence on $\mathbb{D}^{\ell}[0,1]$ is mentioned by $\xrightarrow{\mathbb{D}^{\ell}}$. The integer part of x will be denoted by $\lfloor x \rfloor$.

The proof is divided in two steps.

5.6.1. Functional central limit theorem for $(\Lambda U_t)_{t\geq 1}$

In view of (8) and (13), we deduce that

$$\begin{aligned} \sqrt{n}\hat{\gamma}_m &= \sqrt{n}\gamma_m + \sqrt{n}\Psi_m\left(\hat{\theta}_n - \theta_0\right) + \mathrm{o}_{\mathbb{P}}(1) \\ &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_{2t} + \Psi_m\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_{1t} + \mathrm{o}_{\mathbb{P}}(1)\right) + \mathrm{o}_{\mathbb{P}}(1) \\ &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \Lambda U_t + \mathrm{o}_{\mathbb{P}}(1). \end{aligned} \tag{64}$$

Now, it is clear that the asymptotic behaviour of $\hat{\gamma}_m$ is related to the limit distribution of $U_t = (U_{1t}', U_{2t}')'$. Our first goal is to show that there exists a lower triangular matrix Π with nonnegative diagonal entries such that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{\lfloor nr \rfloor} \Lambda U_t \xrightarrow[n\to\infty]{\mathbb{D}^m} \left(\Pi\Pi'\right)^{1/2} B_m(r), \tag{65}$$

where $(B_m(r))_{r\geq 0}$ is a m-dimensional standard Brownian motion. Using (32), U_t can be rewritten as

$$U_t = \Big(-2\left\{\sum_{i=1}^{\infty} \dot{\lambda}_{i,1}(\theta_0)\,\epsilon_t\epsilon_{t-i}, \dots, \sum_{i=1}^{\infty} \dot{\lambda}_{i,p+q+1}(\theta_0)\,\epsilon_t\epsilon_{t-i}\right\} J^{-1'}, \\ \epsilon_t\epsilon_{t-1}, \dots, \epsilon_t\epsilon_{t-m}\Big)'.$$

The non-correlation between ϵ_t's implies that the process $(U_t)_{t\in\mathbb{Z}}$ of $\mathbb{R}^{p+q+1+m}$ is centered. In order to apply the functional central limit theorem for strongly mixing process (see [27]), we need to identify the asymptotic covariance matrix in the classical central limit theorem for the sequence $(U_t)_{t\in\mathbb{Z}}$. It is proved in Proposition 1 that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_t \xrightarrow[n\to\infty]{\text{in law}} \mathcal{N}\left(0, \Xi := 2\pi f_U(0)\right),$$

where $f_U(0)$ is the spectral density of the stationary process $(U_t)_{t\in\mathbb{Z}}$ evaluated at frequency 0. The existence of the matrix Ξ has already been discussed in Lemma 3.

Since the matrix Ξ is positive definite, it can be factorized as $\Xi = \Upsilon\Upsilon'$, where the $(p+q+1+m)\times(p+q+1+m)$ lower triangular matrix Υ has nonnegative diagonal entries. Therefore, we have

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} \Lambda U_t \xrightarrow[n\to\infty]{\text{in law}} \mathcal{N}\left(0, \Lambda\Xi\Lambda'\right),$$

and the new variance matrix can also been factorized as $\Lambda\Xi\Lambda' = (\Lambda\Upsilon)(\Lambda\Upsilon)' := \Pi\Pi'$, where $\Pi \in \mathbb{R}^{m\times(p+q+1)}$. Thus

$$n^{-1/2}\sum_{t=1}^{n}(\Pi\Pi')^{-1/2}\Lambda U_t \xrightarrow[n\to\infty]{\text{in law}} \mathcal{N}(0, I_m),$$

where $(\Pi\Pi')^{-1/2}$ is the Moore-Penrose inverse (see [40], p. 36) of $(\Pi\Pi')^{1/2}$.

Using the same arguments as in the proof of Theorem 2 in [6], the asymptotic distribution of $n^{-1/2}\sum_{t=1}^{n} U_t$ when n tends to infinity is obtained by introducing the random vector U_t^k defined for any positive integer k by

$$U_t^k = \Big(-2\left\{\sum_{i=1}^{k}\dot{\lambda}_{i,1}(\theta_0)\,\epsilon_t\epsilon_{t-i}, \dots, \sum_{i=1}^{k}\dot{\lambda}_{i,p+q+1}(\theta_0)\,\epsilon_t\epsilon_{t-i}\right\} J^{-1'}, \\ \epsilon_t\epsilon_{t-1}, \dots, \epsilon_t\epsilon_{t-m}\Big)'.$$

Since U_t^k depends on a finite number of values of the noise-process $(\epsilon_t)_{t\in\mathbb{Z}}$, it also satisfies a mixing property (see Theorem 14.1 in [12], p. 210). Then applying the central limit theorem for strongly mixing process of [27] shows that its asymptotic distribution is normal with zero mean and variance matrix Ξ_k that converges when k tends to infinity to Ξ. More precisely we have

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_t^k \xrightarrow[n\to\infty]{\text{in law}} \mathcal{N}\left(0, \Xi_k\right).$$

The above arguments also apply to matrix Ξ_k with some matrix Π_k which is defined analogously as Π. Consequently we obtain

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} \Lambda U_t^k \xrightarrow[n\to\infty]{\text{in law}} \mathcal{N}\left(0, \Lambda\Xi_k\Lambda'\right)$$

and we also have $n^{-1/2}\sum_{t=1}^{n}(\Pi_k\Pi_k')^{-1/2}\Lambda U_t^k \xrightarrow[n\to\infty]{\text{in law}} \mathcal{N}(0, I_m)$.

Now we are able to apply the functional central limit theorem (see [27]) and we obtain that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{\lfloor nr\rfloor}(\Pi_k\Pi_k')^{-1/2}\Lambda U_t^k \xrightarrow[n\to\infty]{\mathbb{D}^m} B_m(r).$$

Since for all $t\in\{1,\dots,\lfloor nr\rfloor\}$ we write

$$(\Pi\Pi')^{-1/2}\Lambda U_t^k = \left((\Pi\Pi')^{-1/2}-(\Pi_k\Pi_k')^{-1/2}\right)\Lambda U_t^k + (\Pi_k\Pi_k')^{-1/2}\Lambda U_t^k,$$

we obtain the following weak convergence on $\mathbb{D}^m[0,1]$:

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{\lfloor nr\rfloor}(\Pi\Pi')^{-1/2}\Lambda U_t^k \xrightarrow[n\to\infty]{\mathbb{D}^m} B_m(r).$$

In order to conclude that (65) is true, it remains to observe that uniformly with respect to n

$$Y_n^k(r) := \frac{1}{\sqrt{n}}\sum_{t=1}^{\lfloor nr\rfloor}(\Pi\Pi')^{-1/2}\Lambda Z_t^k \xrightarrow[k\to\infty]{\mathbb{D}^m} 0, \tag{66}$$

where

$$Z_t^k = \Bigg(-2\left\{\sum_{i=k+1}^{\infty}\dot{\lambda}_{i,1}(\theta_0)\,\epsilon_t\epsilon_{t-i},\dots,\sum_{i=k+1}^{\infty}\dot{\lambda}_{i,p+q+1}(\theta_0)\,\epsilon_t\epsilon_{t-i}\right\}J^{-1'}, \\ \epsilon_t\epsilon_{t-1},\dots,\epsilon_t\epsilon_{t-m}\Bigg)'.$$

Using the same arguments as those used in the proof of Theorem 2 in [6], we have

$$\sup_n \operatorname{Var}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} Z_t^k\right) \xrightarrow[k\to\infty]{} 0$$

and since $\lfloor nr\rfloor \le n$,

$$\sup_{0\le r\le 1}\sup_n\left\{\left\|Y_n^k(r)\right\|\right\} \xrightarrow[k\to\infty]{} 0.$$

Thus (66) is true and the proof of (65) is achieved.

5.6.2. *Limit theorem*

To conclude the prove of Theorem 6, we follow the arguments developed in [7]. Note that the previous step ensures us that Assumption 1 in [37] is satisfied for the sequence $(\Lambda U_t)_{t\ge 1}$. Firstly from (65) we deduce that

$$\frac{1}{\sqrt{n}}S_{\lfloor nr\rfloor} = \frac{1}{\sqrt{n}}\sum_{t=1}^{\lfloor nr\rfloor}\Lambda U_t - \frac{\lfloor nr\rfloor}{n}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\Lambda U_t\right) \\ \xrightarrow[n\to\infty]{\mathbb{D}^m} (\Pi\Pi')^{1/2}B_m(r) - r(\Pi\Pi')^{1/2}B_m(1). \tag{67}$$

Observe now that the continuous mapping theorem implies

$$\begin{aligned} C_m &= \frac{1}{n}\sum_{t=1}^{n}\left(\frac{1}{\sqrt{n}}S_t\right)\left(\frac{1}{\sqrt{n}}S_t\right)' \\ &\xrightarrow[n\to\infty]{\mathbb{D}^m} (\Pi\Pi')^{1/2}\left[\int_0^1 \{B_m(r)-rB_m(1)\}\{B_m(r)-rB_m(1)\}'\,\mathrm{d}r\right](\Pi\Pi')^{1/2} \end{aligned}$$

and consequently

$$C_m \xrightarrow[n\to\infty]{\mathbb{D}^m} (\Pi\Pi')^{1/2}V_m(\Pi\Pi')^{1/2}.$$

Using (64), (67) and again the continuous mapping theorem on the Skorokhod space, one finally obtains

$$\begin{aligned} n\hat{\gamma}_m' C_m^{-1}\hat{\gamma}_m &\xrightarrow[n\to\infty]{\mathbb{D}^m} \left\{(\Pi\Pi')^{1/2}B_m(1)\right\}'\left\{(\Pi\Pi')^{1/2}V_m(\Pi\Pi')^{1/2}\right\}^{-1}\left\{(\Pi\Pi')^{1/2}B_m(1)\right\} \\ &= B_m'(1)V_m^{-1}B_m(1) := \mathcal{U}_m. \end{aligned}$$

Consequently, from (9) it follows that

$$n\sigma_\epsilon^4\hat{\rho}_m' C_m^{-1}\hat{\rho}_m \xrightarrow[n\to\infty]{\mathbb{D}^m} \mathcal{U}_m,$$

which completes the proof of Theorem 6. □

5.7. Proof of Theorem 7

The proof follows the same line as in the proof of Theorem 2 in [7] (see also the proof of in [6]).

Supplementary Material

Supplement A: Example of explicit calculation of $\Sigma_{\hat{\rho}_m}$ and C_m

The results of the previous subsections 3.2 and 3.3 are particularized in the FARIMA$(1,d_0,0)$ and FARIMA$(0,d_0,1)$ cases. First we consider the case of a FARIMA$(1,d_0,0)$ model of the form

$$(1-L)^{d_0}(X_t - aX_{t-1}) = \epsilon_t, \tag{68}$$

where the unknown parameter is $\theta_0 = (a,d_0)$. We assume that in (68) the innovation process $(\epsilon_t)_{t\in\mathbb{Z}}$ is a GARCH$(1,1)$ process given by (20). We also assume that in (20): $\alpha_1^2\kappa + \beta_1^2 + 2\alpha_1\beta_1 < 1$,[3] where $\kappa := \mathbb{E}\eta_1^4$ and we assume that $\kappa > 1$.

For the sake of simplicity we assume that the variables $(\eta_t)_{t\in\mathbb{Z}}$ involved in (20) have a symmetric distribution. More precisely, we have the following symmetry assumption

[3] This is a necessary and sufficient condition for the existence of a nonanticipative stationary solution process $(\epsilon_t)_{t\in\mathbb{Z}}$ with fourth-order moments (see [24, Example 2.3]).

$$\mathbb{E}[\epsilon_{t_1}\epsilon_{t_2}\epsilon_{t_3}\epsilon_{t_4}] = 0 \qquad \text{when} \qquad t_1 \neq t_2,\ t_1 \neq t_3 \text{ and } t_1 \neq t_4, \tag{69}$$

made in [23, 5]. For this particular GARCH$(1,1)$ model with fourth-order moments and symmetric innovations satisfying (69), it can be shown that

$$\mathbb{E}\left[\epsilon_t\epsilon_{t-\ell}\epsilon_{t-h}\epsilon_{t-h-\ell'}\right] = \begin{cases} \mathbb{E}\left[\epsilon_t^2\epsilon_{t-\ell}^2\right] & \text{if } h = 0 \text{ and } \ell = \ell' \\ \\ 0 & \text{otherwise.} \end{cases} \tag{70}$$

Now we need to compute the autocovariance structure of $(\epsilon_t^2)_{t\in\mathbb{Z}}$. We will use the fact that the GARCH process $(\epsilon_t)_{t\in\mathbb{Z}}$ is fourth-order stationary, then $(\epsilon_t^2)_{t\in\mathbb{Z}}$ is a solution of the following ARMA$(1,1)$ model

$$\epsilon_t^2 = \omega + (\alpha_1 + \beta_1)\epsilon_{t-1}^2 + \nu_t - \beta_1\nu_{t-1}, \quad t \in \mathbb{Z} \tag{71}$$

where $\nu_t = \epsilon_t^2 - \sigma_t^2$ is the innovation of $(\epsilon_t^2)_{t\in\mathbb{Z}}$. From (71) the autocovariances of $(\epsilon_t^2)_{t\in\mathbb{Z}}$ take the form

$$\gamma_{\epsilon^2}(\ell) := \mathrm{Cov}(\epsilon_t^2, \epsilon_{t-\ell}^2) = \gamma_{\epsilon^2}(1)(\alpha_1 + \beta_1)^{\ell-1}, \qquad \ell \geq 1, \tag{72}$$

where

$$\gamma_{\epsilon^2}(1) = \frac{(\kappa - 1)(\alpha_1 - \alpha_1\beta_1^2 - \alpha_1^2\beta_1)}{1 - \beta_1^2 - 2\alpha_1\beta_1 - \alpha_1^2\kappa}\sigma_\epsilon^4,$$

$$\gamma_{\epsilon^2}(0) := \mathrm{Var}(\epsilon_t^2) = \frac{(\kappa - 1)(1 - \beta_1^2 - 2\alpha_1\beta_1)}{1 - \beta_1^2 - 2\alpha_1\beta_1 - \alpha_1^2\kappa}\sigma_\epsilon^4,$$

$$\text{and } \sigma_\epsilon^2 := \frac{\omega}{1 - \alpha_1 - \beta_1}.$$

From (70) and (72) we deduce that for any $\ell \geq 1$

$$\begin{aligned} \Gamma(\ell, \ell) &= \mathbb{E}\left[\epsilon_t^2\epsilon_{t-\ell}^2\right] = \mathrm{Cov}(\epsilon_t^2, \epsilon_{t-\ell}^2) + \mathbb{E}\left[\epsilon_t^2\right]\mathbb{E}\left[\epsilon_{t-\ell}^2\right] \\ &= \left\{1 + \frac{1}{\sigma_\epsilon^4}\gamma_{\epsilon^2}(1)(\alpha_1 + \beta_1)^{\ell-1}\right\}\sigma_\epsilon^4. \end{aligned} \tag{73}$$

Examples of analytic and numerical computations of $\boldsymbol{\Sigma}_{\hat{\rho}_m}$

As mentioned before, the subject of this subsection is to give an explicit expression of the asymptotic variance of residual autocorrelations $\Sigma_{\hat{\rho}_m}$ defined in (15) in the particular case of model (68). For that sake, we need the following additional expressions. It is classical that the noise derivatives $(\partial\epsilon_t(\theta_0)/\partial a, \partial\epsilon_t(\theta_0)/\partial d)'$ in (68) can be represented as

$$\begin{pmatrix} \frac{\partial\epsilon_t(\theta_0)}{\partial a} \\ \frac{\partial\epsilon_t(\theta_0)}{\partial d} \end{pmatrix} = -\sum_{j\geq 1}\begin{pmatrix} a^{j-1} \\ \frac{1}{j} \end{pmatrix}\epsilon_{t-j}. \tag{74}$$

We compute the information matrices $J(\theta_0)$ and $I(\theta_0)$ by using (74). Then we have

$$J(\theta_0) = 2\sigma_\epsilon^2 \begin{pmatrix} \frac{1}{1-a^2} & -\frac{\ln(1-a)}{a} \\ -\frac{\ln(1-a)}{a} & \frac{\pi^2}{6} \end{pmatrix}. \tag{75}$$

A simple calculation implies that

$$J^{-1}(\theta_0) = \frac{1}{2\sigma_\epsilon^2 c(a)} \begin{pmatrix} \frac{\pi^2}{6} & \frac{\ln(1-a)}{a} \\ \frac{\ln(1-a)}{a} & \frac{1}{1-a^2} \end{pmatrix}, \tag{76}$$

where

$$c(a) = \frac{\pi^2}{6(1-a^2)} - \left(\frac{\ln(1-a)}{a}\right)^2. \tag{77}$$

We now investigate a similar tractable expression for $I(\theta_0)$. Using (74) and (69) we have

$$\begin{aligned} I(\theta_0) &= 2\sigma_\epsilon^2 J(\theta_0) \\ &\quad + 4\sigma_\epsilon^4 \frac{(\kappa-1)(\alpha_1 - \alpha_1\beta_1^2 - \alpha_1^2\beta_1)}{1-\beta_1^2 - 2\alpha_1\beta_1 - \alpha_1^2\kappa} \begin{pmatrix} \frac{1}{1-a^2(\alpha_1+\beta_1)} & -\frac{\ln[1-a(\alpha_1+\beta_1)]}{a(\alpha_1+\beta_1)} \\ -\frac{\ln[1-a(\alpha_1+\beta_1)]}{a(\alpha_1+\beta_1)} & \frac{\mathrm{Li}_2(\alpha_1+\beta_1)}{\alpha_1+\beta_1} \end{pmatrix}, \end{aligned} \tag{78}$$

where Li_2 is the Spence function defined by $\mathrm{Li}_2(z) = \sum_{k=1}^{\infty} z^k k^{-2}$. Note that we retrieve the well know result: $I(\theta_0) = 2\sigma_\epsilon^2 J(\theta_0)$ in the strong FARIMA case (*i.e.* when $\alpha_1 = \beta_1 = 0$ in (78)).

The matrix defined in (7) can be rewritten as

$$\Psi_m = -\sigma_\epsilon^2 \begin{pmatrix} 1 & a & \dots & a^{m-1} \\ 1 & \frac{1}{2} & \dots & \frac{1}{m} \end{pmatrix}'. \tag{79}$$

Using (73) and under the symmetry assumption (69), the matrix $\Gamma_{m,m}$ takes the simple following diagonal form

$$\begin{aligned} \Gamma_{m,m} &= \sigma_\epsilon^4 I_m \\ &\quad + \sigma_\epsilon^4 \frac{(\kappa-1)(\alpha_1 - \alpha_1\beta_1^2 - \alpha_1^2\beta_1)}{1-\beta_1^2 - 2\alpha_1\beta_1 - \alpha_1^2\kappa} \mathrm{diag}(1, (\alpha_1+\beta_1), \dots, (\alpha_1+\beta_1)^{m-1}). \end{aligned} \tag{80}$$

Using (69), (74) and (76), the matrix $\Sigma'_{\hat{\theta},\gamma_m}$ is given by

$$\Sigma'_{\hat{\theta},\gamma_m} = \frac{1}{\sigma_\epsilon^2 c(a)} \times$$
$$\begin{pmatrix} \left\{\frac{\pi^2}{6} + \frac{\ln(1-a)}{a}\right\} \Gamma_{m,m}(1,1) & \left\{\frac{1}{1-a^2} + \frac{\ln(1-a)}{a}\right\} \Gamma_{m,m}(1,1) \\ \left\{a\frac{\pi^2}{6} + \frac{\ln(1-a)}{2a}\right\} \Gamma_{m,m}(2,2) & \left\{\frac{1}{2(1-a^2)} + \ln(1-a)\right\} \Gamma_{m,m}(2,2) \\ \vdots & \vdots \\ \left\{a^{m-1}\frac{\pi^2}{6} + \frac{\ln(1-a)}{ma}\right\} \Gamma_{m,m}(m,m) & \left\{\frac{1}{m(1-a^2)} + a^{m-2}\ln(1-a)\right\} \Gamma_{m,m}(m,m) \end{pmatrix}, \tag{81}$$

where for any $1 \le i, j \le m$, $\Gamma_{m,m}(i,j)$ is given by (80).

From Remark 2, in the strong FARIMA case the asymptotic variance of residual autocorrelations takes a simpler form

$$\Sigma^{\mathrm{S}}_{\hat{\rho}_m} = I_m - \frac{1}{c(a)}\left[\frac{\pi^2}{6}\left(a^{i+j-2}\right) + \frac{1}{1-a^2}\left(\frac{1}{ij}\right) + \frac{\ln(1-a)}{a}\left(\frac{a^{j-1}}{i}+\frac{a^{i-1}}{j}\right)\right]_{1\leq i,j\leq m}$$

where $c(a)$ is the constant given in (77).

From the above explicit expressions we deduce that the asymptotic variance of residual autocorrelations for this model is in the form

$$\Sigma_{\hat{\rho}_m} = \Sigma^{\mathrm{S}}_{\hat{\rho}_m} + \frac{(\kappa-1)(\alpha_1-\alpha_1\beta_1^2-\alpha_1^2\beta_1)}{1-\beta_1^2-2\alpha_1\beta_1-\alpha_1^2\kappa}\Bigg[(\alpha_1+\beta_1)^{i-1}\mathbf{1}_{\{i=j\}} + \frac{1}{c(a)}M(i,j) - \left\{(\alpha_1+\beta_1)^{i-1}+(\alpha_1+\beta_1)^{j-1}\right\}\frac{1}{c(a)}\left\{\frac{\pi^2}{6}\left(a^{i+j-2}\right)+\frac{1}{1-a^2}\left(\frac{1}{ij}\right)+ \frac{\ln(1-a)}{a}\left(\frac{a^{j-1}}{i}+\frac{a^{i-1}}{j}\right)\right\}\Bigg]_{1\leq i,j\leq m},$$

where

$$\begin{aligned}
M(i,j) &= \left[\frac{\ln(1-a)}{a}\frac{1}{1-a^2(\alpha_1+\beta_1)} - \frac{1}{1-a^2}\frac{\ln(1-a(\alpha_1+\beta_1))}{a(\alpha_1+\beta_1)}\right]\left[\frac{\pi^2}{6}\frac{a^{j-1}}{i}+\frac{1}{ij}\frac{\ln(1-a)}{a}\right]\\
&+\left[\frac{\mathrm{Li}_2(\alpha_1+\beta_1)}{\alpha_1+\beta_1}\frac{1}{1-a^2} - \frac{\ln(1-a)}{a}\frac{\ln(1-a(\alpha_1+\beta_1))}{a(\alpha_1+\beta_1)}\right]\left[\frac{\ln(1-a)}{a}\frac{a^{j-1}}{i}+\frac{1}{ij}\frac{1}{1-a^2}\right]\\
&+\left[\frac{\pi^2}{6}\frac{1}{1-a^2(\alpha_1+\beta_1)} - \frac{\ln(1-a)}{a}\frac{\ln(1-a(\alpha_1+\beta_1))}{a(\alpha_1+\beta_1)}\right]\left[\frac{\pi^2}{6}a^{i+j-2}+\frac{a^{i-1}}{j}\frac{\ln(1-a)}{a}\right]\\
&+\left[\frac{\mathrm{Li}_2(\alpha_1+\beta_1)}{\alpha_1+\beta_1}\frac{\ln(1-a)}{a} - \frac{\pi^2}{6}\frac{\ln(1-a(\alpha_1+\beta_1))}{a(\alpha_1+\beta_1)}\right]\left[\frac{\ln(1-a)}{a}a^{i+j-2}+\frac{a^{i-1}}{j}\frac{1}{1-a^2}\right].
\end{aligned}$$

For simplicity, we take in the sequel $\beta_1 = 0$ to consider the case of an ARCH(1) model. For instance when $m=3$, $\kappa=3$, $\omega=1$ and $a=-0.55$ we have

	$\Sigma_{\hat{\rho}_3}$	Eigenvalues ξ_3	$Z_3(\xi_3)$
$\alpha_1 = 0$	$\begin{pmatrix} 0.14 & 0.09 & -0.27 \\ 0.09 & 0.25 & 0.01 \\ -0.27 & 0.06 & 0.91 \end{pmatrix}$	$(1.00, 0.28, 0.02)$	$\chi^2_1 + 0.28\chi^2_1 + 0.02\chi^2_1$
$\alpha_1 = 0.55$	$\begin{pmatrix} 0.7 & 0.38 & -1.60 \\ 0.38 & 0.94 & -0.23 \\ -1.60 & -0.23 & 4.8 \end{pmatrix}$	$(5.38, 1.00, 0.05)$	$5.38\chi^2_1 + 1.00\chi^2_1 + 0.05\chi^2_1$

It is clear that for $\alpha_1 = 0.55$, the [33] approximation by a χ^2_1 distribution will be disastrous. The eigenvalues ξ_m can be very different from those of strong FARIMA models which are close to 1 or 0 when the lag m is large enough (see Remark 3). More precisely, for instance for $\alpha_1 = 0$ and $m = 12$ we obtain

$$\xi_{12} = (1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 0.07, 0.00)',$$

In this weak FARIMA$(1, d, 0)$ with $\alpha_1 = 0.55$ and $m = 12$ we also obtain

$$\xi_{12} = (5.46, 3.75, 2.32, 1.79, 1.41, 1.24, 1.13, 1.07, 1.04, 1.02, 0.080.00)'.$$

The same result holds for FARIMA$(0, d, 1)$ model with a replaced by b in θ_0.

Explicit form of the matrix C_m

The following example gives an explicit form of the normalization matrix C_m for the model given in (68). For reading convenience, we restrict ourselves to the case $m = 3$. Using the expression of $J^{-1}(\theta_0)$ given in (76) and Equation (74), we obtain that for all $1 \le j \le n$

$$-2J^{-1}(\theta_0)\epsilon_j \begin{pmatrix} \frac{\partial \epsilon_j(\theta_0)}{\partial a} \\ \frac{\partial \epsilon_j(\theta_0)}{\partial d} \end{pmatrix} = \begin{pmatrix} v_j^{(1)}(a) \\ v_j^{(2)}(a) \end{pmatrix},$$

where

$$v_j^{(1)}(a) = \frac{1}{\sigma_\epsilon^2 c(a)} \sum_{k \ge 1} \left\{ \frac{\pi^2}{6} a^{k-1} + \frac{\ln(1-a)}{a} \frac{1}{k} \right\} \epsilon_j \epsilon_{j-k}$$

and

$$v_j^{(2)}(a) = \frac{1}{\sigma_\epsilon^2 c(a)} \sum_{k \ge 1} \left\{ \frac{\ln(1-a)}{a} a^{k-1} + \frac{1}{1-a^2} \frac{1}{k} \right\} \epsilon_j \epsilon_{j-k}.$$

Thus, the vector ΛU_j is given by

$$\Lambda U_j = \begin{pmatrix} -\sigma_\epsilon^2 v_j^{(1)}(a) - \sigma_\epsilon^2 v_j^{(2)}(a) + \epsilon_j \epsilon_{j-1} \\ -\sigma_\epsilon^2 a v_j^{(1)}(a) - \sigma_\epsilon^2 v_j^{(2)}(a)/2 + \epsilon_j \epsilon_{j-2} \\ -\sigma_\epsilon^2 a^2 v_j^{(1)}(a) - \sigma_\epsilon^2 v_j^{(2)}(a)/3 + \epsilon_j \epsilon_{j-3} \end{pmatrix}.$$

A simple calculation shows that, for any $1 \le j_1, j_2 \le n$,

$$(\Lambda U_{j_1})(\Lambda U_{j_2})' = \begin{pmatrix} K_{j_1}^{(1)}(a) K_{j_2}^{(1)}(a) & K_{j_1}^{(1)}(a) K_{j_2}^{(2)}(a) & K_{j_1}^{(1)}(a) K_{j_2}^{(3)}(a) \\ K_{j_1}^{(2)}(a) K_{j_2}^{(1)}(a) & K_{j_1}^{(2)}(a) K_{j_2}^{(2)}(a) & K_{j_1}^{(2)}(a) K_{j_2}^{(3)}(a) \\ K_{j_1}^{(3)}(a) K_{j_2}^{(1)}(a) & K_{j_1}^{(3)}(a) K_{j_2}^{(2)}(a) & K_{j_1}^{(3)}(a) K_{j_2}^{(3)}(a) \end{pmatrix},$$

where

$$\begin{aligned} K_j^{(1)}(a) &= -\sigma_\epsilon^2 v_j^{(1)}(a) - \sigma_\epsilon^2 v_j^{(2)}(a) + \epsilon_j \epsilon_{j-1}, \\ K_j^{(2)}(a) &= -\sigma_\epsilon^2 a v_j^{(1)}(a) - \sigma_\epsilon^2 v_j^{(2)}(a)/2 + \epsilon_j \epsilon_{j-2} \\ \text{and} \quad K_j^{(3)}(a) &= -\sigma_\epsilon^2 a^2 v_j^{(1)}(a) - \sigma_\epsilon^2 v_j^{(2)}(a)/3 + \epsilon_j \epsilon_{j-3}. \end{aligned}$$

Therefore we deduce that for all positive integer t

$$S_t = \sum_{j=1}^{t} (\Lambda U_j - \gamma_3)$$
$$= \sum_{j=1}^{t} \begin{pmatrix} -\sigma_\epsilon^2 v_j^{(1)}(a) - \sigma_\epsilon^2 v_j^{(2)}(a) + \epsilon_j \epsilon_{j-1} \\ -\sigma_\epsilon^2 a v_j^{(1)}(a) - \sigma_\epsilon^2 v_j^{(2)}(a)/2 + \epsilon_j \epsilon_{j-2} \\ -\sigma_\epsilon^2 a^2 v_j^{(1)}(a) - \sigma_\epsilon^2 v_j^{(2)}(a)/3 + \epsilon_j \epsilon_{j-3} \end{pmatrix} - \frac{t}{n} \begin{pmatrix} \sum_{j=2}^{n} \epsilon_j \epsilon_{j-1} \\ \sum_{j=3}^{n} \epsilon_j \epsilon_{j-2} \\ \sum_{j=4}^{n} \epsilon_j \epsilon_{j-3} \end{pmatrix}.$$

The same result holds for FARIMA$(0, d_0, 1)$ model with a replaced by b in θ_0.

Supplement B: Additional Monte Carlo experiments

For the nominal level $\alpha = 5\%$, the empirical size over the N independent replications should vary between the significant limits 3.6% and 6.4% with probability 95%. When the relative rejection frequencies are outside the 95% significant limits, they are displayed in bold type in Tables.

FARIMA models with $a \neq 0$ and $b \neq 0$

Table 8 displays the relative rejection frequencies of the null hypothesis **(H0)** that the DGP follows a strong FARIMA model (19), over the N independent replications. When $p = q = 1$ for all tests, the percentages of rejection belong to the confident interval with probabilities 95%, except for $\mathrm{LB_s}$ and $\mathrm{BP_s}$ (see Table 8). Consequently all these tests well control the error of first kind.

We draw the conclusion that in these strong FARIMA cases the proposed modified version may be clearly preferable to the standard ones.

Now, we repeat the same experiments on two weak FARIMA models. As expected Tables 9 and 10 show that the standard $\mathrm{LB_s}$ or $\mathrm{BP_s}$ test poorly performs in assessing the adequacy of these particular weak FARIMA models. Indeed, we observe that

- the observed relative rejection frequencies of $\mathrm{LB_s}$ and $\mathrm{BP_s}$ are definitely outside the significant limits,
- the errors of the first kind are only globally well controlled by the proposed tests when n is large.

We also investigate the case where the GARCH model (20) have infinite fourth moments. As showing in Figures 5, ..., 10 the results are qualitatively similar to what we observe here in Tables 9 and 10.

Table 8

Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of a strong FARIMA$(1, d_0, 1)$ defined by (19) with $\theta_0 = (0.9, 0.2, d_0)$. The nominal asymptotic level of the tests is $\alpha = 5\%$. The number of replications is $N = 1,000$.

d_0	Length n	Lag m	LB_{SN}	BP_{SN}	LB_W	BP_W	LB_S	BP_S
		1	5.8	5.7	**7.4**	**7.3**	n.a.	n.a.
		2	5.0	5.0	**7.4**	**7.3**	n.a.	n.a.
0.05	$n = 1,000$	3	4.3	4.3	5.8	5.8	n.a.	n.a.
		6	4.1	4.1	5.6	5.5	**10.9**	**10.9**
		12	5.1	4.6	4.7	4.5	**6.9**	**6.6**
		15	5.0	4.7	5.0	4.8	**6.9**	5.9
		1	6.0	6.0	**7.4**	**7.4**	n.a.	n.a.
		2	**6.5**	**6.5**	**7.9**	**7.9**	n.a.	n.a.
0.05	$n = 5,000$	3	4.7	4.7	**6.7**	**6.7**	n.a.	n.a.
		6	**3.5**	**3.5**	5.2	5.1	**11.0**	**10.9**
		12	5.3	5.3	5.8	5.8	**7.9**	**7.6**
		15	4.5	4.5	5.8	5.5	**7.0**	**6.9**
		1	4.2	4.2	6.1	6.1	n.a.	n.a.
		2	4.2	4.2	6.3	6.4	n.a.	n.a.
0.05	$n = 10,000$	3	3.8	3.8	5.9	5.9	n.a.	n.a.
		6	**3.5**	**3.5**	4.7	4.7	**10.4**	**10.4**
		12	4.2	4.2	6.1	6.1	**7.6**	**7.6**
		15	4.0	3.8	5.7	5.7	**7.4**	**7.4**
		1	5.8	5.8	**9.2**	**9.1**	n.a.	n.a.
		2	4.9	4.9	**7.5**	**7.5**	n.a.	n.a.
0.20	$n = 1,000$	3	4.6	4.5	5.9	5.9	n.a.	n.a.
		6	4.2	4.1	5.6	5.4	**10.3**	**10.2**
		12	5.4	4.9	4.7	4.4	6.4	5.9
		15	5.5	4.9	5.1	4.4	6.8	6.2
		1	6.4	6.4	6.1	6.2	n.a.	n.a.
		2	**6.8**	**6.8**	**6.9**	**6.9**	n.a.	n.a.
0.20	$n = 5,000$	3	4.3	4.3	5.9	5.8	n.a.	n.a.
		6	3.8	3.8	4.6	4.6	**10.0**	**10.0**
		12	5.2	5.2	5.7	5.6	**7.6**	**7.5**
		15	4.5	4.5	5.6	5.3	**6.8**	**6.7**
		1	4.5	4.5	5.5	5.5	n.a.	n.a.
		2	4.1	4.1	5.8	5.8	n.a.	n.a.
0.20	$n = 10,000$	3	**3.1**	**3.1**	5.3	5.3	n.a.	n.a.
		6	3.7	3.6	4.3	4.3	**10.1**	**10.1**
		12	3.8	3.8	6.1	6.1	**7.5**	**7.5**
		15	3.7	3.7	5.8	5.7	**7.0**	**6.9**
		1	4.3	4.3	**8.7**	**8.7**	n.a.	n.a.
		2	**3.0**	**3.0**	5.9	5.9	n.a.	n.a.
0.45	$n = 1,000$	3	3.7	3.7	4.4	4.4	n.a.	n.a.
		6	3.8	3.8	4.7	4.5	**8.1**	**7.8**
		12	5.1	4.6	4.3	4.2	5.1	4.9
		15	4.6	4.5	4.7	4.3	5.0	4.7
		1	5.6	5.5	6.0	6.0	n.a.	n.a.
		2	5.2	5.2	6.4	6.4	n.a.	n.a.
0.45	$n = 5,000$	3	4.0	4.0	5.9	5.9	n.a.	n.a.
		6	3.8	3.8	4.6	4.6	**10.1**	**9.9**
		12	5.2	5.2	5.4	5.4	**7.2**	**7.1**
		15	4.6	4.6	5.0	4.9	**6.7**	**6.6**
		1	4.3	4.3	5.3	5.3	n.a.	n.a.
		2	3.2	3.2	5.7	5.7	n.a.	n.a.
0.45	$n = 10,000$	3	**3.1**	**3.0**	5.4	5.4	n.a.	n.a.
		6	3.7	3.7	4.3	4.3	**9.8**	**9.8**
		12	4.3	4.3	5.8	5.8	**7.2**	**7.0**
		15	3.6	3.3	5.7	5.7	**6.8**	**6.8**

TABLE 9

Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of a weak FARIMA$(1, d_0, 1)$ *defined by* (19) *with* $\theta_0 = (0.9, 0.2, d_0)$ *and where* $\omega = 0.4$, $\alpha_1 = 0.3$ *and* $\beta_1 = 0.3$ *in* (20). *The nominal asymptotic level of the tests is* $\alpha = 5\%$. *The number of replications is* $N = 1,000$.

d_0	Length n	Lag m	$\mathrm{LB_{SN}}$	$\mathrm{BP_{SN}}$	$\mathrm{LB_W}$	$\mathrm{BP_W}$	$\mathrm{LB_S}$	$\mathrm{BP_S}$
		1	4.9	4.9	**6.7**	**6.7**	n.a.	n.a.
		2	3.8	3.8	6.3	6.3	n.a.	n.a.
0.05	$n = 1,000$	3	**3.2**	**3.2**	5.2	5.2	n.a.	n.a.
		6	3.9	3.8	4.9	4.8	**18.5**	**18.3**
		12	**2.3**	**2.3**	4.1	4.0	**10.2**	**9.7**
		15	**2.7**	**2.3**	4.4	4.2	**9.7**	**9.3**
		1	5.1	5.1	5.6	5.6	n.a.	n.a.
		2	4.9	4.9	5.4	5.4	n.a.	n.a.
0.05	$n = 5,000$	3	**2.6**	**2.6**	5.0	5.0	n.a.	n.a.
		6	**3.5**	**3.5**	4.4	4.4	**19.6**	**19.6**
		12	**2.7**	**2.7**	**3.3**	**3.2**	**11.4**	**11.4**
		15	**3.4**	**3.4**	4.2	4.1	**10.8**	**10.7**
		1	4.8	4.8	**6.9**	**6.9**	n.a.	n.a.
		2	4.8	4.8	**6.7**	**6.7**	n.a.	n.a.
0.05	$n = 10,000$	3	4.7	4.7	5.5	5.5	n.a.	n.a.
		6	**3.3**	**3.3**	6.4	6.4	**20.2**	**20.2**
		12	4.2	4.2	6.3	6.3	**12.4**	**12.3**
		15	3.6	3.6	5.5	5.5	**11.6**	**11.6**
		1	5.3	5.3	**7.8**	**7.7**	n.a.	n.a.
		2	3.6	**3.4**	5.7	5.7	n.a.	n.a.
0.20	$n = 1,000$	3	**3.1**	**3.1**	4.9	4.8	n.a.	n.a.
		6	**3.3**	**3.2**	4.5	4.5	**17.6**	**17.4**
		12	**2.3**	**2.0**	4.1	4.1	**9.4**	**8.9**
		15	**2.4**	**2.1**	4.4	4.2	**9.0**	**8.1**
		1	4.6	4.6	4.3	4.3	n.a.	n.a.
		2	4.3	4.3	4.4	4.4	n.a.	n.a.
0.20	$n = 5,000$	3	**3.1**	**3.1**	4.4	4.3	n.a.	n.a.
		6	4.1	4.1	3.9	3.9	**19.0**	**19.0**
		12	**2.6**	**2.6**	**2.9**	**2.9**	**10.9**	**10.6**
		15	**3.4**	**3.3**	4.0	4.0	**10.0**	**9.9**
		1	4.8	4.8	5.1	5.1	n.a.	n.a.
		2	4.7	4.7	5.0	5.0	n.a.	n.a.
0.20	$n = 10,000$	3	4.5	4.5	4.8	4.8	n.a.	n.a.
		6	**3.5**	**3.5**	5.6	5.6	**19.1**	**19.1**
		12	4.1	4.1	5.9	5.9	**12.1**	**12.1**
		15	3.7	3.7	5.3	5.3	**11.3**	**11.3**
		1	4.4	4.4	**11.1**	**11.0**	n.a.	n.a.
		2	**3.4**	**3.4**	5.4	5.3	n.a.	n.a.
0.45	$n = 1,000$	3	**3.1**	**3.1**	4.9	4.9	n.a.	n.a.
		6	**3.1**	**2.9**	4.5	4.4	**15.3**	**15.1**
		12	**2.2**	**2.1**	4.0	4.0	**7.9**	**7.5**
		15	**2.1**	**2.0**	4.4	4.3	**7.0**	**6.5**
		1	3.9	3.9	4.2	4.2	n.a.	n.a.
		2	**3.4**	**3.4**	4.2	4.2	n.a.	n.a.
0.45	$n = 5,000$	3	**2.9**	**2.9**	4.4	4.4	n.a.	n.a.
		6	**3.5**	**3.5**	3.9	3.9	**18.4**	**18.4**
		12	**2.4**	**2.4**	**2.8**	**2.7**	**9.9**	**9.8**
		15	**3.2**	**3.2**	3.9	3.8	**9.2**	**9.2**
		1	4.6	4.6	5.3	5.3	n.a.	n.a.
		2	4.3	4.3	5.1	5.0	n.a.	n.a.
0.45	$n = 10,000$	3	**3.5**	**3.5**	5.0	5.0	n.a.	n.a.
		6	**2.8**	**2.8**	5.3	5.3	**19.3**	**19.3**
		12	4.2	4.2	5.5	5.5	**12.2**	**12.2**
		15	3.6	**3.5**	5.5	5.5	**11.4**	**11.4**

Table 10

Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of a weak FARIMA$(1, d_0, 1)$ defined by (19)–(21) with $\theta_0 = (0.9, 0.2, d_0)$. The nominal asymptotic level of the tests is $\alpha = 5\%$. The number of replications is $N = 1,000$.

d_0	Length n	Lag m	LB_{SN}	BP_{SN}	LB_W	BP_W	LB_S	BP_S
0.05	$n = 1,000$	1	5.1	5.1	**7.3**	**7.3**	n.a.	n.a.
		2	3.6	3.6	**6.9**	**6.9**	n.a.	n.a.
		3	**2.9**	**2.9**	4.3	4.1	n.a.	n.a.
		6	**2.6**	**2.5**	**3.1**	**3.0**	**10.3**	**10.3**
		12	**0.9**	**0.9**	**1.2**	**1.1**	**8.7**	**8.3**
		15	**0.4**	**0.4**	**1.0**	**0.8**	**8.0**	**7.3**
0.05	$n = 5,000$	1	3.9	3.9	5.4	5.4	n.a.	n.a.
		2	3.9	3.9	5.9	5.9	n.a.	n.a.
		3	3.9	3.9	5.5	5.5	n.a.	n.a.
		6	**3.2**	**3.1**	3.8	3.8	**10.6**	**10.6**
		12	**2.4**	**2.4**	**3.5**	**3.4**	**8.3**	**8.2**
		15	**2.7**	**2.7**	**3.3**	**3.3**	**8.4**	**8.3**
0.05	$n = 10,000$	1	5.0	5.0	5.2	5.2	n.a.	n.a.
		2	4.9	4.9	4.5	4.5	n.a.	n.a.
		3	3.8	3.8	5.6	5.6	n.a.	n.a.
		6	3.6	3.6	4.5	4.5	**10.4**	**10.4**
		12	**3.3**	**3.3**	4.3	4.3	**8.5**	**8.4**
		15	4.7	4.7	3.8	3.8	**7.7**	**7.4**
0.20	$n = 1,000$	1	5.7	5.6	**10.1**	**10.0**	n.a.	n.a.
		2	**3.4**	**3.4**	5.5	5.5	n.a.	n.a.
		3	3.7	3.7	4.0	4.0	n.a.	n.a.
		6	**2.9**	**2.8**	**2.5**	**2.4**	**10.2**	**9.7**
		12	**0.9**	**0.9**	**1.1**	**1.1**	**7.9**	**7.2**
		15	**0.5**	**0.5**	**0.8**	**0.8**	**7.5**	**6.9**
0.20	$n = 5,000$	1	**3.5**	**3.5**	4.0	3.9	n.a.	n.a.
		2	3.7	3.7	4.3	4.3	n.a.	n.a.
		3	4.1	4.1	5.0	5.0	n.a.	n.a.
		6	**3.1**	**3.1**	**3.5**	**3.5**	**10.0**	**10.0**
		12	**2.8**	**2.8**	**3.3**	**3.3**	**8.2**	**8.2**
		15	**2.4**	**2.4**	**3.1**	**3.1**	**7.9**	**7.8**
0.20	$n = 10,000$	1	5.1	5.1	4.8	4.8	n.a.	n.a.
		2	4.7	4.7	4.2	4.2	n.a.	n.a.
		3	3.8	3.8	4.7	4.7	n.a.	n.a.
		6	3.8	3.8	4.1	4.1	**10.1**	**10.1**
		12	**3.4**	**3.4**	4.0	4.0	**8.0**	**8.0**
		15	4.8	4.8	3.6	3.6	**7.5**	**7.4**
0.45	$n = 1,000$	1	3.8	3.8	**12.1**	**12.0**	n.a.	n.a.
		2	**2.4**	**2.4**	4.4	4.4	n.a.	n.a.
		3	**2.7**	**2.6**	3.8	3.7	n.a.	n.a.
		6	**3.2**	**3.0**	**2.3**	**2.3**	**8.3**	**7.9**
		12	**1.1**	**0.9**	**1.0**	**0.9**	6.4	6.3
		15	**0.3**	**0.3**	**1.4**	**1.1**	**6.8**	6.4
0.45	$n = 5,000$	1	**3.1**	**3.1**	4.4	4.4	n.a.	n.a.
		2	**2.7**	**2.7**	4.5	4.5	n.a.	n.a.
		3	**3.2**	**3.2**	4.9	4.9	n.a.	n.a.
		6	**3.2**	**3.1**	**3.4**	**3.4**	**9.7**	**9.7**
		12	**3.3**	**3.3**	**3.3**	**3.3**	**7.3**	**7.3**
		15	**2.4**	**2.4**	**3.2**	**3.1**	**7.2**	**7.0**
0.45	$n = 10,000$	1	5.1	5.1	4.8	4.8	n.a.	n.a.
		2	4.9	4.9	4.3	4.3	n.a.	n.a.
		3	3.6	3.6	4.9	4.9	n.a.	n.a.
		6	**3.5**	**3.5**	4.3	4.2	**10.2**	**10.2**
		12	3.7	3.7	3.7	3.7	**7.7**	**7.6**
		15	4.8	4.8	3.9	3.9	**7.2**	**7.1**

Power of the tests

In this section, we repeat the same experiments as in Section 4.1 to examine the power of the tests for the null hypothesis of Model (19) against the following FARIMA alternative defined by

$$(1-L)^d (X_t - aX_{t-1}) = \epsilon_t - b_1\epsilon_{t-1} - b_2\epsilon_{t-2}, \tag{82}$$

with $\theta_0 = (a, b_1, b_2, d_0)$ and where the innovation process $(\epsilon_t)_{t\in\mathbb{Z}}$ follows a strong or weak white noise introduced in Section 4.1.

For each of these N replications we fit a FARIMA$(1, d, 1)$ model (19) and perform standard and modified tests based on $m = 1, 2, 3$, 6, 12 and 15 residual autocorrelations.

Tables 11, 12 and 13 compare the empirical powers of Model (82) with $\theta_0 = (0.9, 1, -0.2, d_0)$ over the N independent replications. For these particular strong and weak FARIMA models, we notice that the standard BP_{S} and LB_{S} and our proposed tests have very similar powers except for $\mathrm{BP}_{\mathrm{SN}}$ and $\mathrm{LB}_{\mathrm{SN}}$ when $n = 5,000$.

Table 11

Empirical power (in %) of the modified and standard versions of the LB and BP tests in the case of a strong FARIMA$(1, d_0, 2)$ defined by (82) with $\theta_0 = (0.9, 1, -0.2, d_0)$. The nominal asymptotic level of the tests is $\alpha = 5\%$. The number of replications is $N = 1{,}000$.

d_0	Length n	Lag m	LB_{SN}	BP_{SN}	LB_W	BP_W	LB_S	BP_S
		1	24.5	24.5	37.9	37.9	n.a.	n.a.
		2	28.8	28.8	46.1	46.1	n.a.	n.a.
0.05	$n = 5,000$	3	36.7	36.7	22.2	22.1	n.a.	n.a.
		6	55.7	55.7	40.6	40.3	47.6	47.6
		12	54.9	54.7	27.2	27.2	28.3	28.0
		15	54.0	53.6	18.0	17.8	27.9	27.7
		1	44.9	44.9	62.8	62.7	n.a.	n.a.
		2	51.4	51.3	76.1	76.0	n.a.	n.a.
0.05	$n = 10,000$	3	62.8	62.8	39.9	39.9	n.a.	n.a.
		6	86.5	86.5	80.9	80.8	84.7	84.7
		12	85.8	85.8	64.9	64.8	66.4	66.2
		15	82.0	82.0	43.2	43.2	60.8	60.8
		1	14.0	14.0	58.0	57.9	n.a.	n.a.
		2	22.2	22.2	71.1	71.1	n.a.	n.a.
0.20	$n = 5,000$	3	24.1	23.8	40.7	40.7	n.a.	n.a.
		6	32.1	32.0	74.4	74.4	78.5	78.5
		12	52.3	52.2	62.4	62.2	67.7	67.6
		15	51.6	51.3	14.1	14.0	62.1	61.7
		1	21.4	21.4	84.9	85.0	n.a.	n.a.
		2	30.6	30.6	93.1	93.1	n.a.	n.a.
0.20	$n = 10,000$	3	35.6	35.6	65.9	65.7	n.a.	n.a.
		6	44.1	44.1	96.9	96.9	97.8	97.8
		12	76.3	76.2	93.2	93.2	94.3	94.3
		15	73.7	73.7	43.9	43.9	91.6	91.6
		1	0.0	0.	100.0	100.0	n.a.	n.a.
		2	49.1	49.1	100.0	100.0	n.a.	n.a.
0.45	$n = 5,000$	3	69.0	69.0	100.0	100.0	n.a.	n.a.
		6	76.7	76.6	100.0	100.0	100.0	100.0
		12	86.8	86.7	100.0	100.0	100.0	100.0
		15	90.9	90.7	100.0	100.0	100.0	100.0
		1	0.0	0.0	100.0	100.0	n.a.	n.a.
		2	77.9	77.9	100.0	100.0	n.a.	n.a.
0.45	$n = 10,000$	3	90.3	90.2	100.0	100.0	n.a.	n.a.
		6	94.2	94.2	100.0	100.0	100.0	100.0
		12	98.9	98.9	100.0	100.0	100.0	100.0
		15	99.5	99.4	100.0	100.0	100.0	100.0

TABLE 12
Empirical power (in %) of the modified and standard versions of the LB and BP tests in the case of a weak FARIMA$(1, d_0, 2)$ defined by (82) with $\theta_0 = (0.9, 1, -0.2, d_0)$ and where $\omega = 0.4$, $\alpha_1 = 0.3$ and $\beta_1 = 0.3$ in (20). The nominal asymptotic level of the tests is $\alpha = 5\%$. The number of replications is $N = 1,000$.

d_0	Length n	Lag m	LB_{SN}	BP_{SN}	LB_W	BP_W	LB_S	BP_S
		1	22.5	22.5	32.8	32.7	n.a.	n.a.
		2	27.3	27.3	41.7	41.8	n.a.	n.a.
0.05	$n = 5,000$	3	32.4	32.3	20.1	20.0	n.a.	n.a.
		6	52.1	52.0	34.0	34.0	55.8	55.7
		12	54.1	54.1	23.5	23.5	34.2	34.1
		15	53.9	53.4	17.1	16.9	31.9	31.8
		1	36.1	36.1	53.2	53.2	n.a.	n.a.
		2	44.9	44.9	64.5	64.5	n.a.	n.a.
0.05	$n = 10,000$	3	56.5	56.5	33.1	33.1	n.a.	n.a.
		6	83.1	83.1	71.2	71.2	86.4	86.2
		12	84.0	83.9	59.0	59.0	70.4	70.2
		15	80.6	80.5	40.1	40.1	67.4	67.2
		1	14.6	14.5	51.0	50.9	n.a.	n.a.
		2	21.8	21.8	67.1	67.1	n.a.	n.a.
0.20	$n = 5,000$	3	22.4	22.3	37.7	37.7	n.a.	n.a.
		6	32.3	32.3	68.3	68.3	81.9	81.9
		12	51.6	51.5	55.9	55.8	68.7	68.5
		15	51.7	51.6	64.2	64.1	64.8	64.6
		1	22.8	22.8	74.1	74.0	n.a.	n.a.
		2	29.6	29.6	86.2	86.2	n.a.	n.a.
0.20	$n = 10,000$	3	32.9	32.9	56.6	56.5	n.a.	n.a.
		6	43.1	43.1	92.3	92.3	97.1	97.1
		12	72.9	72.8	88.3	88.3	93.8	93.8
		15	71.2	71.1	89.1	88.9	92.0	92.0
		1	30.1	30.1	99.8	99.8	n.a.	n.a.
		2	40.1	40.1	100.0	100.0	n.a.	n.a.
0.45	$n = 5,000$	3	57.9	57.9	100.0	100.0	n.a.	n.a.
		6	65.7	65.7	100.0	100.0	100.0	100.0
		12	78.8	78.5	100.0	100.0	100.0	100.0
		15	84.7	84.6	100.0	100.0	100.0	100.0
		1	62.2	62.2	99.9	99.9	n.a.	n.a.
		2	72.2	72.2	100.0	99.9	n.a.	n.a.
0.45	$n = 10,000$	3	84.8	84.8	100.0	100.0	n.a.	n.a.
		6	89.8	89.7	100.0	100.0	100.0	100.0
		12	97.7	97.7	100.0	100.0	100.0	100.0
		15	99.0	99.0	100.0	100.0	100.0	100.0

Table 13
Empirical power (in %) of the modified and standard versions of the LB and BP tests in the case of a weak FARIMA$(1, d_0, 2)$ defined by (82)–(21) with $\theta_0 = (0.9, 1, -0.2, d_0)$. The nominal asymptotic level of the tests is $\alpha = 5\%$. The number of replications is $N = 1,000$.

d_0	Length n	Lag m	LB_{SN}	BP_{SN}	LB_W	BP_W	LB_S	BP_S
0.05	$n = 5,000$	1	27.6	27.6	42.6	42.7	n.a.	n.a.
		2	32.7	32.6	51.4	51.3	n.a.	n.a.
		3	36.9	36.9	23.7	23.7	n.a.	n.a.
		6	53.3	53.0	39.7	39.7	46.0	45.9
		12	49.6	49.3	23.7	23.7	29.3	29.2
		15	44.4	44.2	17.5	17.4	28.5	28.1
0.05	$n = 10,000$	1	48.5	48.5	68.3	68.3	n.a.	n.a.
		2	58.7	58.6	76.6	76.5	n.a.	n.a.
		3	66.8	66.8	42.5	42.5	n.a.	n.a.
		6	84.2	84.0	77.0	76.9	83.2	83.2
		12	79.9	79.9	62.7	62.6	66.0	66.0
		15	75.8	75.8	40.5	40.5	61.4	61.3
0.20	$n = 5,000$	1	15.3	15.3	62.4	62.5	n.a.	n.a.
		2	23.5	23.4	74.6	74.6	n.a.	n.a.
		3	25.9	25.9	45.3	45.2	n.a.	n.a.
		6	34.0	34.0	73.1	72.9	78.5	78.4
		12	51.3	50.8	56.8	56.6	64.5	64.4
		15	46.3	45.8	15.0	14.9	60.1	60.1
0.20	$n = 10,000$	1	23.0	23.0	85.2	85.2	n.a.	n.a.
		2	33.8	33.8	93.6	93.6	n.a.	n.a.
		3	36.5	36.5	68.3	68.3	n.a.	n.a.
		6	46.8	46.7	95.4	95.4	97.1	97.1
		12	81.7	81.7	90.8	90.8	93.7	93.6
		15	79.0	78.7	44.2	44.0	91.7	91.7
0.45	$n = 5,000$	1	41.9	41.9	99.9	99.9	n.a.	n.a.
		2	51.9	51.9	100.0	100.0	n.a.	n.a.
		3	66.7	66.7	100.0	100.0	n.a.	n.a.
		6	73.6	73.6	100.0	100.0	100.0	100.0
		12	83.1	83.0	100.0	100.0	100.0	100.0
		15	85.5	85.4	100.0	100.0	100.0	100.0
0.45	$n = 10,000$	1	69.2	69.2	100.0	99.9	n.a.	n.a.
		2	79.2	79.2	100.0	100.0	n.a.	n.a.
		3	90.8	90.8	100.0	100.0	n.a.	n.a.
		6	93.6	93.6	100.0	100.0	100.0	100.0
		12	97.8	97.8	100.0	100.0	100.0	100.0
		15	99.1	99.1	100.0	100.0	100.0	100.0

Small sample size

The following tables deal with the same numerical experiments that in Section 4 when the sample sizes are less than 500.

Table 14

Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of a strong FARIMA$(0, d_0, 0)$ defined by (19) with $\theta_0 = (0, 0, d_0)$. The nominal asymptotic level of the tests is $\alpha = 5\%$. The number of replications is $N = 1,000$.

d_0	Length n	Lag m	LB_{SN}	BP_{SN}	LB_W	BP_W	LB_S	BP_S
		1	3.9	3.6	**10.1**	**9.6**	n.a.	n.a.
		2	**3.3**	**3.2**	**8.1**	**7.4**	**7.6**	**7.1**
0.05	$n = 100$	3	3.8	**3.1**	5.9	5.2	**8.1**	**6.8**
		6	**3.1**	**2.7**	5.0	3.9	**6.9**	5.9
		12	**2.4**	**1.3**	3.9	**2.1**	5.8	3.8
		15	**2.8**	**1.0**	4.5	**2.3**	**6.9**	4.3
		1	5.3	5.2	**7.6**	**7.3**	n.a.	n.a.
		2	5.0	4.7	5.4	5.3	6.1	6.0
0.05	$n = 250$	3	4.7	4.5	5.6	5.5	5.8	5.6
		6	5.2	4.8	6.4	6.1	6.7	6.3
		12	5.0	3.8	4.4	3.7	6.2	5.3
		15	4.6	**3.2**	4.4	**3.5**	6.0	4.9
		1	5.0	5.0	5.6	5.6	n.a.	n.a.
		2	5.5	5.5	5.7	5.6	6.0	5.8
0.05	$n = 500$	3	5.9	5.7	5.9	5.7	**6.6**	**6.5**
		6	5.3	5.1	5.6	5.2	6.0	5.9
		12	5.1	4.3	5.0	4.7	5.9	5.0
		15	5.4	4.5	4.6	4.2	6.0	5.2
		1	4.5	4.0	5.9	5.3	n.a.	n.a.
		2	4.1	3.7	6.5	6.0	**6.5**	5.8
0.20	$n = 100$	3	4.1	**3.5**	5.3	4.9	6.4	6.1
		6	3.3	**2.9**	4.6	3.7	6.1	4.9
		12	3.6	**1.5**	4.1	**2.0**	5.5	**3.4**
		15	**2.9**	**0.9**	4.4	**2.0**	6.5	**3.5**
		1	5.8	5.7	5.8	5.7	n.a.	n.a.
		2	5.2	5.1	5.2	4.8	5.8	5.6
0.20	$n = 250$	3	5.1	5.0	5.5	5.4	5.4	5.1
		6	5.7	5.4	5.9	5.3	6.3	5.7
		12	5.6	4.0	4.2	3.8	5.8	5.1
		15	4.8	3.6	4.5	3.6	6.2	4.7
		1	5.7	5.5	5.0	5.0	n.a.	n.a.
		2	5.4	5.4	5.4	5.3	5.5	5.3
0.20	$n = 500$	3	6.2	6.1	5.7	5.6	6.3	6.2
		6	5.4	5.0	5.5	5.0	5.6	5.6
		12	5.1	4.4	5.0	4.7	6.0	5.0
		15	5.2	4.3	4.4	4.2	5.9	5.1
		1	4.3	4.1	**9.4**	**8.9**	n.a.	n.a.
		2	3.9	3.4	**8.3**	**7.5**	**7.7**	**7.3**
0.45	$n = 100$	3	4.0	**3.3**	**6.5**	5.7	**7.0**	**6.5**
		6	**3.3**	**2.4**	4.7	**3.5**	**6.5**	5.3
		12	**3.5**	**1.7**	**3.9**	**2.3**	5.5	**3.2**
		15	3.9	**1.4**	4.2	**2.2**	6.1	3.7
		1	5.4	5.4	**8.2**	**7.9**	n.a.	n.a.
		2	5.0	4.9	5.3	5.1	5.5	5.3
0.45	$n = 250$	3	5.1	5.0	5.8	5.3	5.3	5.0
		6	5.6	5.2	6.0	5.2	6.2	5.4
		12	5.4	3.9	4.6	3.9	5.8	5.2
		15	5.1	4.0	4.7	3.7	6.2	5.0
		1	5.4	5.2	5.6	5.6	n.a.	n.a.
		2	5.2	5.2	5.4	5.3	5.9	5.8
0.45	$n = 500$	3	5.9	5.8	6.3	6.1	6.4	6.4
		6	6.0	5.6	5.6	5.0	5.6	5.5
		12	4.9	3.9	5.6	4.8	5.7	5.1
		15	5.2	4.3	4.6	4.2	6.1	4.9

Table 15

Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of a weak FARIMA$(0,d_0,0)$ defined by (19) with $\theta_0=(0,0,d_0)$ with $\omega=0.4$, $\alpha_1=0.3$ and $\beta_1=0.3$ in (21). The nominal asymptotic level of the tests is $\alpha=5\%$. The number of replications is $N=1,000$.

d_0	Length n	Lag m	$\mathrm{LB_{SN}}$	$\mathrm{BP_{SN}}$	$\mathrm{LB_{W}}$	$\mathrm{BP_{W}}$	$\mathrm{LB_{S}}$	$\mathrm{BP_{S}}$
		1	**2.3**	**2.3**	**10.1**	**9.6**	n.a.	n.a.
		2	**2.6**	**2.6**	5.9	5.3	**13.1**	**12.4**
0.05	$n=100$	3	**1.9**	**1.6**	4.0	**3.1**	**11.1**	**9.9**
		6	**1.4**	**1.1**	**3.0**	**2.5**	**12.8**	**11.2**
		12	**1.0**	**0.3**	**3.5**	**2.0**	**14.5**	**10.8**
		15	**0.8**	**0.1**	**2.6**	**0.8**	**16.1**	**11.0**
		1	**3.0**	**3.0**	**8.1**	**8.1**	n.a.	n.a.
		2	**2.6**	**2.4**	5.3	5.2	**16.4**	**16.4**
0.05	$n=250$	3	**1.9**	**1.8**	4.3	3.9	**16.2**	**15.6**
		6	**0.7**	**0.4**	4.3	4.1	**20.1**	**18.8**
		12	**0.6**	**0.5**	3.6	**2.6**	**24.6**	**22.4**
		15	**0.2**	**0.2**	4.0	**2.9**	**25.7**	**22.4**
		1	**3.4**	**3.4**	**7.2**	**7.0**	n.a.	n.a.
		2	**2.0**	**2.0**	6.3	6.3	**20.4**	**20.3**
0.05	$n=500$	3	**1.5**	**1.5**	5.1	5.0	**21.1**	**20.7**
		6	**0.9**	**0.9**	4.6	4.6	**28.0**	**27.6**
		12	**0.4**	**0.4**	4.0	**3.2**	**34.2**	**32.8**
		15	**0.1**	**0.0**	**3.3**	**3.0**	**36.2**	**34.7**
		1	**2.8**	**2.7**	5.3	5.0	n.a.	n.a.
		2	**3.1**	**3.1**	4.9	4.2	**10.9**	**10.1**
0.20	$n=100$	3	**1.8**	**1.6**	3.8	**2.9**	**9.9**	**8.3**
		6	**1.9**	**1.1**	**2.9**	**2.0**	**10.8**	**9.0**
		12	**0.8**	**0.3**	**3.1**	**1.8**	**13.1**	**9.7**
		15	**0.7**	**0.1**	**2.3**	**0.7**	**14.7**	**9.6**
		1	**3.2**	**3.2**	5.5	5.4	n.a.	n.a.
		2	**3.0**	**3.0**	4.3	4.2	**14.4**	**14.3**
0.20	$n=250$	3	**2.4**	**2.3**	3.6	**3.4**	**14.9**	**14.2**
		6	**0.7**	**0.7**	4.3	3.8	**18.3**	**17.3**
		12	**0.6**	**0.4**	**3.5**	**2.6**	**23.6**	**21.2**
		15	**0.4**	**0.1**	3.8	**2.5**	**23.9**	**21.0**
		1	3.8	3.8	5.3	5.3	n.a.	n.a.
		2	**2.4**	**2.3**	6.1	6.1	**18.9**	**18.9**
0.20	$n=500$	3	**1.8**	**1.7**	4.9	4.6	**19.9**	**19.6**
		6	**0.9**	**0.9**	4.4	4.3	**26.5**	**26.2**
		12	**0.4**	**0.4**	3.7	**3.2**	**33.5**	**31.5**
		15	**0.1**	**0.1**	**3.3**	**3.0**	**35.4**	**33.8**
		1	**2.8**	**2.6**	**8.9**	**8.3**	n.a.	n.a.
		2	**2.5**	**2.2**	**6.9**	**6.5**	**12.1**	**11.4**
0.45	$n=100$	3	**1.6**	**1.5**	5.0	4.1	**11.4**	**10.0**
		6	**1.6**	**1.2**	3.4	**2.2**	**10.9**	**8.4**
		12	**0.9**	**0.5**	**3.2**	**1.9**	**13.5**	**10.0**
		15	**0.9**	**0.3**	**2.2**	**0.8**	**14.3**	**9.0**
		1	**3.3**	**3.1**	**8.7**	**8.6**	n.a.	n.a.
		2	**3.3**	**3.1**	6.1	6.1	**16.8**	**16.2**
0.45	$n=250$	3	**2.6**	**2.5**	4.3	4.2	**15.5**	**15.1**
		6	**1.0**	**0.9**	4.5	4.3	**19.0**	**18.0**
		12	**0.6**	**0.4**	3.9	**2.8**	**23.7**	**21.8**
		15	**0.4**	**0.3**	3.6	**2.5**	**24.5**	**21.6**
		1	3.6	**3.5**	**6.7**	**6.6**	n.a.	n.a.
		2	**2.4**	**2.3**	**6.9**	**6.8**	**20.0**	**20.0**
0.45	$n=500$	3	**1.7**	**1.7**	5.4	5.2	**21.3**	**21.2**
		6	**1.0**	**0.9**	4.8	4.5	**26.9**	**26.4**
		12	**0.5**	**0.4**	3.7	**3.5**	**33.2**	**32.0**
		15	**0.1**	**0.1**	**3.5**	**3.1**	**36.3**	**34.8**

TABLE 16

Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of a weak FARIMA$(0, d_0, 0)$ defined by (19)–(21) with $\theta_0 = (0, 0, d_0)$. The nominal asymptotic level of the tests is $\alpha = 5\%$. The number of replications is $N = 1,000$.

d_0	Length n	Lag m	LB_{SN}	BP_{SN}	LB_W	BP_W	LB_S	BP_S
		1	**2.2**	**2.1**	**20.0**	**19.5**	n.a.	n.a.
		2	**1.5**	**1.5**	**15.2**	**14.7**	**18.3**	**17.3**
0.05	$n = 100$	3	**1.1**	**0.9**	**10.7**	**10.1**	**15.3**	**14.4**
		6	**0.4**	**0.2**	6.0	5.2	**10.4**	**9.7**
		12	**0.0**	**0.0**	**3.2**	**2.5**	**8.2**	5.9
		15	**0.2**	**0.0**	**2.4**	**1.7**	**7.7**	5.0
		1	**3.2**	**2.9**	**14.4**	**14.2**	n.a.	n.a.
		2	**3.1**	**2.9**	**10.7**	**10.6**	**18.7**	**18.3**
0.05	$n = 250$	3	**1.9**	**1.8**	**7.8**	**7.6**	**16.3**	**16.0**
		6	**0.9**	**0.6**	4.5	4.2	**12.6**	**12.0**
		12	**0.4**	**0.3**	**2.0**	**1.5**	**10.6**	**8.8**
		15	**0.2**	**0.2**	**1.3**	**1.3**	**10.0**	**8.2**
		1	4.3	4.3	**11.7**	**11.6**	n.a.	n.a.
		2	3.7	3.7	**8.7**	**8.6**	**18.7**	**18.6**
0.05	$n = 500$	3	**2.9**	**2.7**	**6.5**	6.4	**16.7**	**16.6**
		6	**1.8**	**1.6**	**3.4**	**3.2**	**14.4**	**14.1**
		12	**0.3**	**0.2**	**2.2**	**1.7**	**10.9**	**10.4**
		15	**0.2**	**0.2**	**1.1**	**1.0**	**10.2**	**9.7**
		1	3.9	3.7	**11.9**	**11.3**	n.a.	n.a.
		2	**1.5**	**1.5**	**7.4**	**6.8**	**12.3**	**11.4**
0.20	$n = 100$	3	**1.4**	**1.4**	5.2	4.5	**10.7**	**9.6**
		6	**0.3**	**0.2**	**2.3**	**1.8**	**8.4**	**7.6**
		12	**0.1**	**0.0**	**1.1**	**0.8**	**6.5**	4.2
		15	**0.2**	**0.0**	**0.9**	**0.4**	5.8	**3.4**
		1	3.9	3.8	**7.1**	**6.9**	n.a.	n.a.
		2	3.6	**3.4**	6.1	5.7	**13.2**	**13.1**
0.20	$n = 250$	3	**1.9**	**1.8**	3.8	**3.4**	**11.7**	**11.3**
		6	**0.9**	**0.6**	**2.6**	**2.3**	**9.8**	**9.3**
		12	**0.3**	**0.3**	**1.0**	**0.6**	**8.8**	**7.6**
		15	**0.2**	**0.2**	**0.5**	**0.5**	**8.9**	**7.2**
		1	5.3	5.3	6.3	6.1	n.a.	n.a.
		2	4.0	3.9	5.4	5.3	**15.8**	**15.6**
0.20	$n = 500$	3	**3.3**	**3.3**	3.7	3.6	**12.9**	**12.9**
		6	**1.9**	**1.5**	**1.4**	**1.4**	**11.9**	**11.5**
		12	**0.2**	**0.1**	**1.2**	**0.9**	**9.8**	**9.2**
		15	**0.3**	**0.2**	**0.5**	**0.5**	**9.2**	**8.9**
		1	3.9	3.8	**21.5**	**20.2**	n.a.	n.a.
		2	**1.6**	**1.5**	**13.1**	**11.9**	**16.5**	**16.4**
0.45	$n = 100$	3	**1.2**	**0.9**	**7.5**	**7.2**	**13.7**	**12.7**
		6	**0.7**	**0.7**	**3.1**	**2.4**	**10.6**	**9.2**
		12	**0.1**	**0.0**	**1.3**	**0.8**	**6.9**	5.2
		15	**0.2**	**0.0**	**1.3**	**0.3**	6.2	3.8
		1	5.0	5.0	**15.7**	**15.5**	n.a.	n.a.
		2	**3.0**	**3.0**	**10.4**	**10.0**	**18.6**	**18.2**
0.45	$n = 250$	3	**2.3**	**2.3**	**7.5**	**7.3**	**16.1**	**15.9**
		6	**0.6**	**0.4**	3.6	3.6	**12.1**	**11.4**
		12	**0.4**	**0.3**	**1.5**	**1.1**	**9.7**	**8.6**
		15	**0.2**	**0.2**	**1.1**	**0.8**	**10.1**	**8.8**
		1	4.8	4.8	**12.5**	**12.5**	n.a.	n.a.
		2	4.2	4.0	**8.9**	**8.7**	**19.6**	**19.5**
0.45	$n = 500$	3	3.2	3.2	5.7	5.6	16.6	16.6
		6	**2.0**	**1.8**	**2.6**	**2.5**	**13.7**	**13.4**
		12	**0.1**	**0.1**	**1.5**	**1.1**	**10.8**	**10.3**
		15	0.3	0.2	0.6	0.6	10.4	10.1

TABLE 17

Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of a strong FARIMA$(1, d_0, 1)$ defined by (19) with $\theta_0 = (0.9, 0.2, d_0)$. The nominal asymptotic level of the tests is $\alpha = 5\%$. The number of replications is $N = 1,000$.

d_0	Length n	Lag m	LB_{SN}	BP_{SN}	LB_W	BP_W	LB_S	BP_S
0.05	$n = 100$	1	4.7	4.4	**23.2**	**22.9**	n.a.	n.a.
		2	3.9	3.6	**8.1**	**7.5**	n.a.	n.a.
		3	4.3	4.0	**6.9**	6.1	n.a.	n.a.
		6	4.7	3.6	5.4	3.7	**8.4**	5.9
		12	5.1	**2.9**	4.2	**2.3**	5.0	**2.6**
		15	6.2	3.5	4.9	**2.5**	5.8	**2.5**
0.05	$n = 250$	1	5.3	5.3	**10.8**	**10.7**	n.a.	n.a.
		2	3.6	**3.3**	**6.8**	**6.8**	n.a.	n.a.
		3	4.0	3.7	5.7	5.4	n.a.	n.a.
		6	4.2	3.7	5.4	5.1	**10.6**	**9.6**
		12	**3.1**	**2.2**	5.3	4.2	**6.5**	5.7
		15	**3.3**	**2.5**	5.6	4.3	6.4	5.2
0.05	$n = 500$	1	4.6	4.6	**6.9**	**6.8**	n.a.	n.a.
		2	4.3	4.2	5.8	5.6	n.a.	n.a.
		3	4.3	4.2	5.7	5.5	n.a.	n.a.
		6	5.0	4.8	**6.7**	**6.5**	**11.0**	**10.7**
		12	4.9	4.2	5.5	4.6	**7.1**	6.2
		15	5.6	4.3	5.7	4.5	**7.1**	6.2
0.20	$n = 100$	1	5.1	4.8	**27.1**	**25.9**	n.a.	n.a.
		2	4.0	3.8	**8.7**	**8.2**	n.a.	n.a.
		3	4.1	4.0	**7.5**	**6.9**	n.a.	n.a.
		6	5.5	3.9	5.3	3.9	**7.6**	6.2
		12	4.9	**3.0**	4.3	**2.6**	4.3	**2.9**
		15	**6.9**	**2.4**	5.1	**2.9**	5.2	**2.7**
0.20	$n = 250$	1	5.1	5.0	**14.0**	**13.9**	n.a.	n.a.
		2	3.4	**3.1**	**7.3**	**7.2**	n.a.	n.a.
		3	4.3	4.1	6.2	5.9	n.a.	n.a.
		6	4.7	4.3	6.0	5.5	**10.3**	**9.8**
		12	3.8	**2.6**	5.1	4.3	5.7	5.1
		15	3.9	**2.8**	5.9	4.4	5.7	5.0
0.20	$n = 500$	1	5.6	5.6	**12.1**	**12.1**	n.a.	n.a.
		2	4.9	4.9	**7.0**	**6.9**	n.a.	n.a.
		3	5.0	4.9	**6.7**	6.4	n.a.	n.a.
		6	5.5	5.2	6.2	5.7	**10.1**	**9.6**
		12	5.6	4.8	5.3	4.6	6.3	5.3
		15	5.7	4.4	5.4	4.5	5.9	5.1
0.45	$n = 100$	1	**3.2**	**3.1**	**32.0**	**31.6**	n.a.	n.a.
		2	**3.5**	**3.4**	**8.3**	**7.3**	n.a.	n.a.
		3	**2.9**	**2.5**	**6.9**	6.4	n.a.	n.a.
		6	3.8	**2.9**	3.6	**2.8**	4.6	**3.5**
		12	3.6	**1.3**	**2.7**	**1.8**	**2.1**	**1.2**
		15	4.1	**1.9**	3.7	**1.5**	**2.2**	**0.9**
0.45	$n = 250$	1	**3.4**	**3.3**	**18.3**	**18.0**	n.a.	n.a.
		2	**3.2**	**3.2**	6.4	6.1	n.a.	n.a.
		3	3.6	**3.4**	5.2	5.1	n.a.	n.a.
		6	3.8	**3.3**	4.8	4.4	**7.9**	**7.3**
		12	**3.1**	**2.3**	4.0	**3.2**	4.4	3.7
		15	**3.2**	**2.3**	4.7	**3.3**	4.0	**3.1**
0.45	$n = 500$	1	3.6	3.6	**14.5**	**14.4**	n.a.	n.a.
		2	**3.4**	**3.4**	5.3	5.3	n.a.	n.a.
		3	**3.4**	**3.4**	5.5	5.5	n.a.	n.a.
		6	5.0	4.7	4.9	4.6	**7.2**	**7.0**
		12	5.2	4.7	4.4	3.9	4.2	4.0
		15	5.0	4.3	4.4	3.6	4.2	3.7

Table 18

Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of a weak FARIMA$(1, d_0, 1)$ defined by (19) with $\theta_0 = (0.9, 0.2, d_0)$ and where $\omega = 0.4$, $\alpha_1 = 0.3$ and $\beta_1 = 0.3$ in (20). The nominal asymptotic level of the tests is $\alpha = 5\%$. The number of replications is $N = 1,000$.

d_0	Length n	Lag m	LB_{SN}	BP_{SN}	LB_W	BP_W	LB_S	BP_S
		1	**3.1**	**3.1**	**19.7**	**18.7**	n.a.	n.a.
		2	**2.0**	**1.7**	**7.8**	**7.3**	n.a.	n.a.
0.05	$n = 100$	3	**1.7**	**1.6**	**6.8**	6.2	n.a.	n.a.
		6	**1.4**	**0.9**	6.1	4.7	**15.6**	**12.5**
		12	**1.5**	**0.9**	5.1	3.7	**13.5**	**8.9**
		15	**2.0**	**1.2**	5.1	**2.6**	**13.1**	**8.9**
		1	**2.5**	**2.4**	**10.6**	**10.0**	n.a.	n.a.
		2	**2.1**	**1.7**	**6.6**	6.4	n.a.	n.a.
0.05	$n = 250$	3	**1.2**	**1.1**	5.7	5.2	n.a.	n.a.
		6	**0.8**	**0.8**	5.3	4.7	**25.0**	**24.2**
		12	**0.8**	**0.7**	3.7	**3.3**	**23.5**	**21.5**
		15	**1.1**	**1.1**	3.8	**3.0**	**24.7**	**21.8**
		1	**2.4**	**2.4**	**8.1**	**8.1**	n.a.	n.a.
		2	**1.7**	**1.7**	**7.1**	**7.0**	n.a.	n.a.
0.05	$n = 500$	3	**0.8**	**0.7**	6.1	6.0	n.a.	n.a.
		6	**0.7**	**0.6**	4.6	4.2	**31.5**	**31.0**
		12	**1.1**	**1.1**	3.9	3.8	**33.5**	**32.3**
		15	**1.0**	**0.9**	4.6	4.0	**35.0**	**33.4**
		1	**2.6**	**2.6**	**24.0**	**23.4**	n.a.	n.a.
		2	**1.7**	**1.6**	**9.0**	**8.4**	n.a.	n.a.
0.20	$n = 100$	3	**2.3**	**1.7**	**6.7**	6.2	n.a.	n.a.
		6	**1.5**	**0.8**	5.5	4.2	**15.2**	**12.3**
		12	**1.4**	**0.6**	4.5	**3.1**	**12.0**	**7.7**
		15	**2.0**	**0.8**	4.7	**2.8**	**11.2**	**7.5**
		1	**3.5**	**3.5**	**17.1**	**16.8**	n.a.	n.a.
		2	**1.9**	**1.9**	**8.5**	**8.0**	n.a.	n.a.
0.20	$n = 250$	3	**1.1**	**1.0**	5.5	5.0	n.a.	n.a.
		6	**0.7**	**0.7**	4.3	4.1	**24.2**	**23.4**
		12	**0.6**	**0.6**	**3.3**	**2.9**	**22.1**	**19.7**
		15	**0.6**	**0.5**	3.8	**3.1**	**22.9**	**20.1**
		1	**2.5**	**2.4**	**12.0**	**11.8**	n.a.	n.a.
		2	**2.0**	**2.0**	**7.7**	**7.7**	n.a.	n.a.
0.20	$n = 500$	3	**1.4**	**1.4**	6.1	5.6	n.a.	n.a.
		6	**0.8**	**0.8**	4.3	4.0	**30.2**	**29.6**
		12	**0.8**	**0.7**	**3.4**	**3.2**	**33.2**	**31.7**
		15	**0.7**	**0.6**	4.3	3.8	**34.3**	**32.7**
		1	**2.4**	**2.3**	**33.2**	**32.9**	n.a.	n.a.
		2	**1.4**	**1.3**	**8.5**	**7.8**	n.a.	n.a.
0.45	$n = 100$	3	**1.5**	**1.2**	6.3	5.4	n.a.	n.a.
		6	**1.4**	**0.8**	4.5	**3.5**	**10.5**	**8.3**
		12	**0.8**	**0.3**	4.3	**2.7**	**7.0**	5.0
		15	**1.5**	**0.4**	4.1	**2.4**	**7.5**	4.3
		1	**2.1**	**2.1**	**20.1**	**20.1**	n.a.	n.a.
		2	**1.7**	**1.7**	5.9	5.8	n.a.	n.a.
0.45	$n = 250$	3	**1.1**	**0.8**	5.2	4.9	n.a.	n.a.
		6	**0.9**	**0.9**	4.1	3.7	**18.8**	**18.0**
		12	**0.4**	**0.4**	**2.6**	**2.1**	**17.4**	**15.4**
		15	**0.2**	**0.2**	4.2	**3.0**	**18.4**	**15.7**
		1	**2.1**	**2.1**	**13.3**	**13.2**	n.a.	n.a.
		2	**1.2**	**1.2**	5.8	5.7	n.a.	n.a.
0.45	$n = 500$	3	**1.1**	**1.0**	4.9	4.9	n.a.	n.a.
		6	**0.6**	**0.6**	4.0	3.8	**27.3**	**26.4**
		12	**0.2**	**0.2**	**3.1**	**2.8**	**28.3**	**27.0**
		15	**0.2**	**0.1**	4.3	3.8	**28.4**	**26.8**

TABLE 19

Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of a weak FARIMA$(1, d_0, 1)$ defined by (19) with $\theta_0 = (0.9, 0.2, d_0)$ and where $\omega = 0.04$, $\alpha_1 = 0.12$ and $\beta_1 = 0.85$ in (20). The nominal asymptotic level of the tests is $\alpha = 5\%$. The number of replications is $N = 1,000$.

d_0	Length n	Lag m	$\mathrm{LB_{SN}}$	$\mathrm{BP_{SN}}$	$\mathrm{LB_W}$	$\mathrm{BP_W}$	$\mathrm{LB_S}$	$\mathrm{BP_S}$
0.05	$n = 100$	1	**3.1**	**3.1**	**19.7**	**18.7**	n.a.	n.a.
		2	**2.0**	**1.7**	**7.8**	**7.3**	n.a.	n.a.
		3	**1.7**	**1.6**	**6.8**	6.2	n.a.	n.a.
		6	**1.4**	**0.9**	6.1	4.7	**15.6**	**12.5**
		12	**1.5**	**0.9**	5.1	3.7	**13.5**	**8.9**
		15	**2.0**	**1.2**	5.1	**2.6**	**13.1**	**8.9**
0.05	$n = 250$	1	**2.5**	**2.4**	**10.6**	**10.0**	n.a.	n.a.
		2	**2.1**	**1.7**	**6.6**	6.4	n.a.	n.a.
		3	**1.2**	**1.1**	5.7	5.2	n.a.	n.a.
		6	**0.8**	**0.8**	5.3	4.7	**25.0**	**24.2**
		12	**0.8**	**0.7**	3.7	**3.3**	**23.5**	**21.5**
		15	**1.1**	**1.1**	3.8	**3.0**	**24.7**	**21.8**
0.05	$n = 500$	1	**2.4**	**2.4**	**8.1**	**8.1**	n.a.	n.a.
		2	**1.7**	**1.7**	**7.1**	**7.0**	n.a.	n.a.
		3	**0.8**	**0.7**	6.1	6.0	n.a.	n.a.
		6	**0.7**	**0.6**	4.6	4.2	**31.5**	**31.0**
		12	**1.1**	**1.1**	3.9	3.8	**33.5**	**32.3**
		15	**1.0**	**0.9**	4.6	4.0	**35.0**	**33.4**
0.20	$n = 100$	1	**2.6**	**2.6**	**24.0**	**23.4**	n.a.	n.a.
		2	**1.7**	**1.6**	**9.0**	**8.4**	n.a.	n.a.
		3	**2.3**	**1.7**	**6.7**	6.2	n.a.	n.a.
		6	**1.5**	**0.8**	5.5	4.2	**15.2**	**12.3**
		12	**1.4**	**0.6**	4.5	**3.1**	**12.0**	**7.7**
		15	**2.0**	**0.8**	4.7	**2.8**	**11.2**	**7.5**
0.20	$n = 250$	1	**3.5**	**3.5**	**17.1**	**16.8**	n.a.	n.a.
		2	**1.9**	**1.9**	**8.5**	**8.0**	n.a.	n.a.
		3	**1.1**	**1.0**	5.5	5.0	n.a.	n.a.
		6	**0.7**	**0.7**	4.3	4.1	**24.2**	**23.4**
		12	**0.6**	**0.6**	**3.3**	**2.9**	**22.1**	**19.7**
		15	**0.6**	**0.5**	3.8	**3.1**	**22.9**	**20.1**
0.20	$n = 500$	1	**2.5**	**2.4**	**12.0**	**11.8**	n.a.	n.a.
		2	**2.0**	**2.0**	**7.7**	**7.7**	n.a.	n.a.
		3	**1.4**	**1.4**	6.1	5.6	n.a.	n.a.
		6	**0.8**	**0.8**	4.3	4.0	**30.2**	**29.6**
		12	**0.8**	**0.7**	**3.4**	**3.2**	**33.2**	**31.7**
		15	**0.7**	**0.6**	4.3	3.8	**34.3**	**32.7**
0.45	$n = 100$	1	**2.4**	**2.3**	**33.2**	**32.9**	n.a.	n.a.
		2	**1.4**	**1.3**	**8.5**	**7.8**	n.a.	n.a.
		3	**1.5**	**1.2**	6.3	5.4	n.a.	n.a.
		6	**1.4**	**0.8**	4.5	**3.5**	**10.5**	**8.3**
		12	**0.8**	**0.3**	4.3	**2.7**	**7.0**	5.0
		15	**1.5**	**0.4**	4.1	**2.4**	**7.5**	4.3
0.45	$n = 250$	1	**2.1**	**2.1**	**20.1**	**20.1**	n.a.	n.a.
		2	**1.7**	**1.7**	5.9	5.8	n.a.	n.a.
		3	**1.1**	**0.8**	5.2	4.9	n.a.	n.a.
		6	**0.9**	**0.9**	4.1	3.7	**18.8**	**18.0**
		12	**0.4**	**0.4**	**2.6**	**2.1**	**17.4**	**15.4**
		15	**0.2**	**0.2**	4.2	**3.0**	**18.4**	**15.7**
0.45	$n = 500$	1	**2.1**	**2.1**	**13.3**	**13.2**	n.a.	n.a.
		2	**1.2**	**1.2**	5.8	5.7	n.a.	n.a.
		3	**1.1**	**1.0**	4.9	4.9	n.a.	n.a.
		6	**0.6**	**0.6**	4.0	3.8	**27.3**	**26.4**
		12	**0.2**	**0.2**	**3.1**	**2.8**	**28.3**	**27.0**
		15	**0.2**	**0.1**	4.3	3.8	**28.4**	**26.8**

TABLE 20

Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of a weak FARIMA$(1, d_0, 1)$ defined by (19) with $\theta_0 = (0.9, 0.2, d_0)$ and where $\omega = 0.04$, $\alpha_1 = 0.12$ and $\beta_1 = 0.85$ in (20). The nominal asymptotic level of the tests is $\alpha = 5\%$. The number of replications is $N = 1,000$.

d_0	Length n	Lag m	LB_{SN}	BP_{SN}	LB_W	BP_W	LB_S	BP_S
		1	**3.3**	**3.3**	**8.9**	**8.9**	n.a.	n.a.
		2	**2.5**	**2.5**	**7.5**	**7.5**	n.a.	n.a.
0.05	$n = 1,000$	3	**2.1**	**2.1**	5.4	5.3	n.a.	n.a.
		6	**1.1**	**1.0**	4.3	4.1	**38.4**	**38.1**
		12	**0.6**	**0.6**	3.7	**3.2**	**43.3**	**42.7**
		15	**0.2**	**0.2**	**3.5**	**3.5**	**45.7**	**44.6**
		1	4.4	4.4	5.4	5.4	n.a.	n.a.
		2	**3.5**	**3.5**	5.0	5.0	n.a.	n.a.
0.05	$n = 10,000$	3	**3.3**	**3.3**	6.3	6.3	n.a.	n.a.
		6	**2.3**	**2.2**	4.5	4.4	**57.7**	**57.6**
		12	**1.6**	**1.6**	4.0	3.9	**68.9**	**68.8**
		15	**1.3**	**1.3**	4.6	4.6	**73.8**	**73.8**
		1	4.6	4.6	5.2	5.2	n.a.	n.a.
		2	4.3	4.3	5.0	5.0	n.a.	n.a.
0.05	$n = 20,000$	3	3.9	3.9	4.8	4.8	n.a.	n.a.
		6	**2.5**	**2.5**	5.0	4.8	**41.7**	**41.7**
		12	3.8	3.8	4.2	4.2	**51.1**	**51.1**
		15	**3.3**	**3.2**	3.7	3.7	**52.8**	**52.7**
		1	**3.5**	**3.4**	**9.9**	**9.7**	n.a.	n.a.
		2	**2.4**	**2.4**	6.4	6.4	n.a.	n.a.
0.20	$n = 1,000$	3	**2.2**	**2.1**	4.9	4.9	n.a.	n.a.
		6	**1.1**	**0.8**	**3.5**	**3.4**	**37.4**	**37.2**
		12	**0.3**	**0.3**	**3.5**	**3.3**	**42.9**	**42.4**
		15	**0.0**	**0.0**	3.6	**3.5**	**44.4**	**43.2**
		1	4.2	4.2	4.0	4.0	n.a.	n.a.
		2	**3.4**	**3.4**	4.1	4.1	n.a.	n.a.
0.20	$n = 10,000$	3	**3.3**	**3.3**	5.3	5.3	n.a.	n.a.
		6	**2.2**	**2.2**	4.3	4.3	**55.8**	**55.8**
		12	**1.6**	**1.6**	3.9	3.9	**67.7**	**67.7**
		15	**1.3**	**1.3**	4.1	4.1	**72.9**	**72.9**
		1	5.0	5.0	4.3	4.3	n.a.	n.a.
		2	4.6	4.6	4.4	4.4	n.a.	n.a.
0.20	$n = 20,000$	3	3.9	3.9	4.7	4.7	n.a.	n.a.
		6	**2.7**	**2.7**	4.7	4.7	**41.0**	**41.0**
		12	3.7	3.7	4.0	4.0	**50.3**	**50.3**
		15	**3.4**	**3.4**	3.6	**3.5**	**51.9**	**51.8**
		1	**3.0**	**3.0**	**12.1**	**12.2**	n.a.	n.a.
		2	**1.8**	**1.8**	5.5	5.4	n.a.	n.a.
0.45	$n = 1,000$	3	**1.7**	**1.6**	4.4	4.4	n.a.	n.a.
		6	**0.6**	**0.6**	**3.4**	**3.3**	**34.7**	**34.4**
		12	**0.4**	**0.4**	**3.3**	**3.0**	**38.6**	**38.0**
		15	**0.2**	**0.2**	3.6	**3.5**	**40.0**	**38.9**
		1	3.7	3.6	3.7	3.7	n.a.	n.a.
		2	**3.0**	**3.0**	3.7	3.8	n.a.	n.a.
0.45	$n = 10,000$	3	**3.0**	**3.0**	5.1	5.1	n.a.	n.a.
		6	**2.0**	**2.0**	4.7	4.7	**55.3**	**55.3**
		12	**1.7**	**1.7**	3.8	3.8	**67.6**	**67.4**
		15	**1.3**	**1.3**	4.1	4.1	**72.0**	**71.8**
		1	5.0	5.0	4.1	4.1	n.a.	n.a.
		2	4.5	4.5	4.1	4.1	n.a.	n.a.
0.45	$n = 20,000$	3	3.7	3.7	4.8	4.8	n.a.	n.a.
		6	**2.9**	**2.9**	4.5	4.4	**40.5**	**40.5**
		12	3.7	3.7	3.8	3.8	**49.8**	**49.7**
		15	**3.5**	**3.5**	3.6	3.6	**51.3**	**51.3**

GARCH process with infinite moment

In order to see if the test procedures remain reliable for GARCH process with infinite moment (for $\alpha_1 + \beta_1 \geq 1$), we replicate the numerical experiments made on Model (19)–(20) with $\omega = 0.04$, $\alpha_1 = 0.13$ and $\beta_1 = 0.88$.

Figures 5, ..., 10 indicate that the results are qualitatively similar to what we observe here in Tables 2, 3 9 and 10.

Figures 5, ..., 9 display the residual autocorrelations of a realization of size $n = 2,000$ for weak FARIMA models (19)–(20) with $\omega = 0.04$, $\alpha_1 = 0.13$, $\beta_1 = 0.88$ and three values of d_0, and their 5% significance limits under the strong FARIMA and weak FARIMA assumptions. These figures confirm clearly the conclusions drawn in Subsection 4.1. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong FARIMA assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak FARIMA assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 2. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 7.

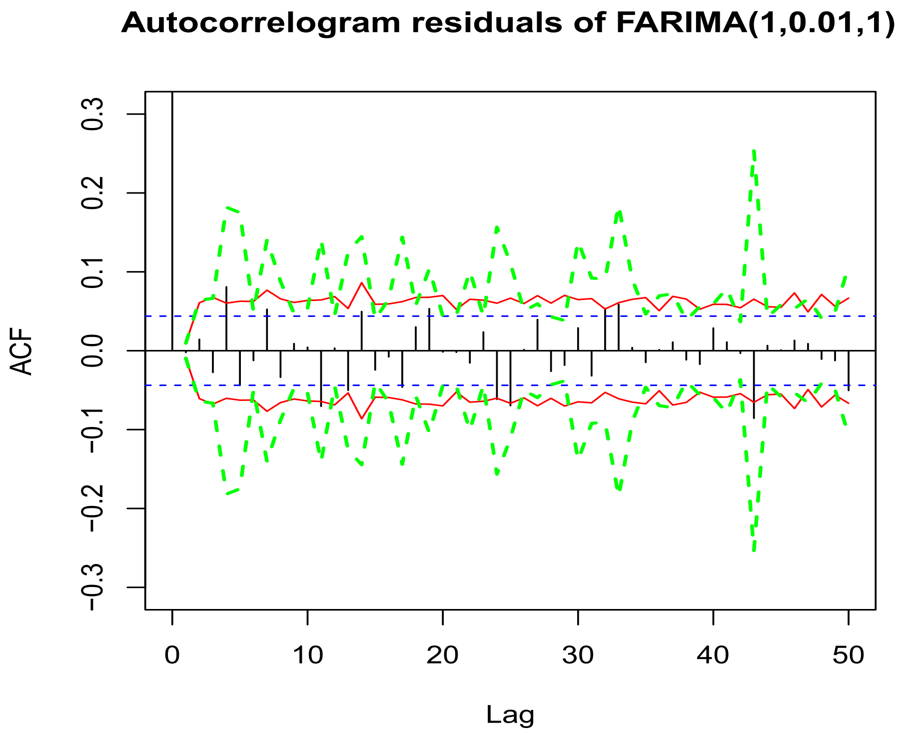

Fig 5. *Autocorrelation of a realization of size* $n = 2,000$ *for a weak FARIMA*$(1, 0.01, 1)$ *model* (19)–(20) *with* $\theta_0 = (0.9, 0.2, 0.01)$ *and where* $\omega = 0.04$, $\alpha_1 = 0.13$ *and* $\beta_1 = 0.88$. *The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong FARIMA assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak FARIMA assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 2. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 7.*

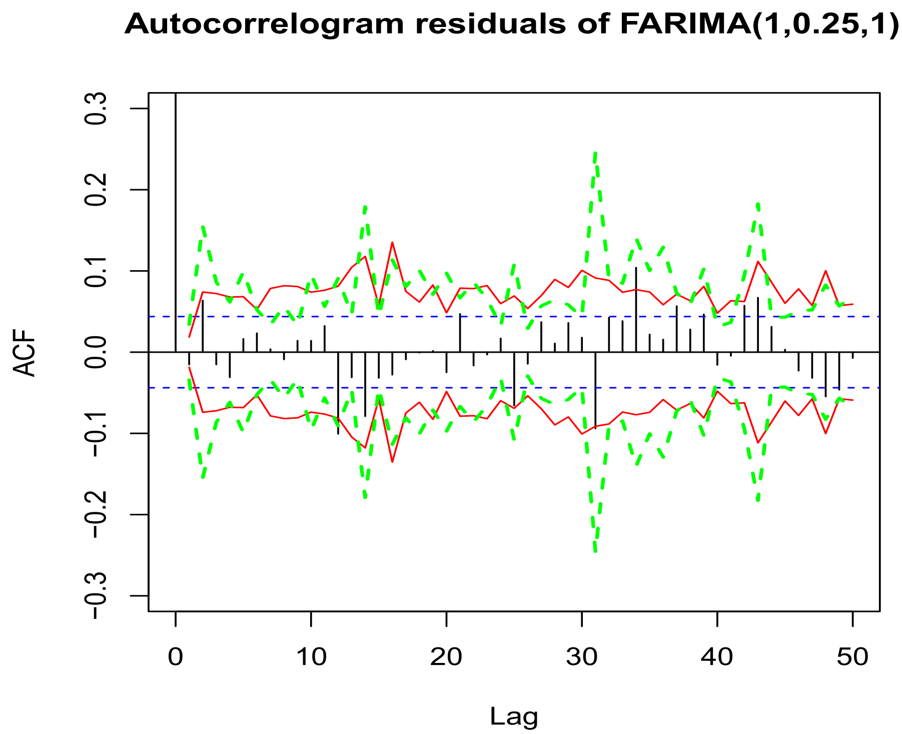

FIG 6. *Autocorrelation of a realization of size* $n = 2,000$ *for a weak FARIMA*$(1, 0.25, 1)$ *model* (19)–(20) *with* $\theta_0 = (0.9, 0.2, 0.25)$ *and where* $\omega = 0.04$, $\alpha_1 = 0.13$ *and* $\beta_1 = 0.88$. *The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong FARIMA assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak FARIMA assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 2. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 7.*

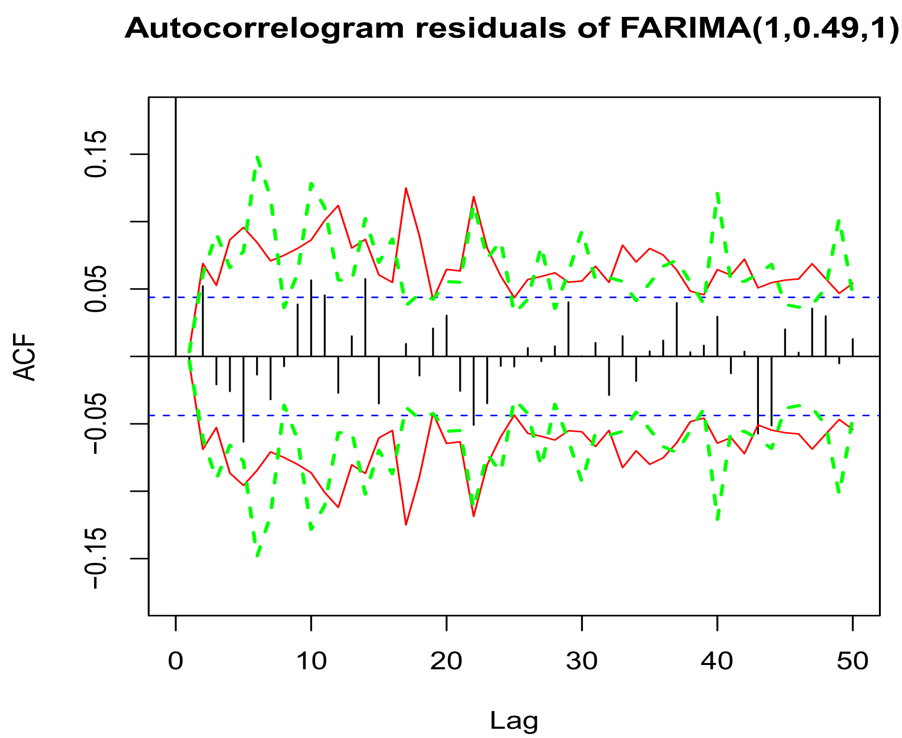

FIG 7. *Autocorrelation of a realization of size $n = 2,000$ for a weak FARIMA$(1, 0.49, 1)$ model (19)–(20) with $\theta_0 = (0.9, 0.2, 0.49)$ and where $\omega = 0.04$, $\alpha_1 = 0.13$ and $\beta_1 = 0.88$. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong FARIMA assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak FARIMA assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 2. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 7.*

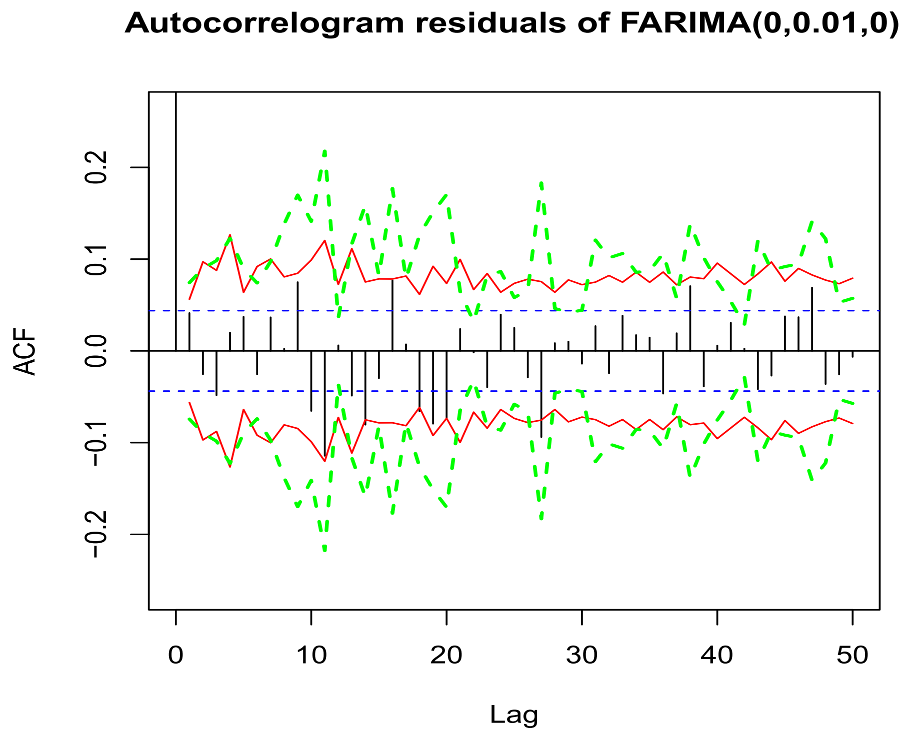

Fig 8. *Autocorrelation of a realization of size* $n = 2,000$ *for a weak FARIMA*$(0, 0.01, 0)$ *model* (19)–(20) *with* $\theta_0 = (0, 0, 0.01)$ *and where* $\omega = 0.04$*,* $\alpha_1 = 0.13$ *and* $\beta_1 = 0.88$*. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong FARIMA assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak FARIMA assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 2. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 7.*

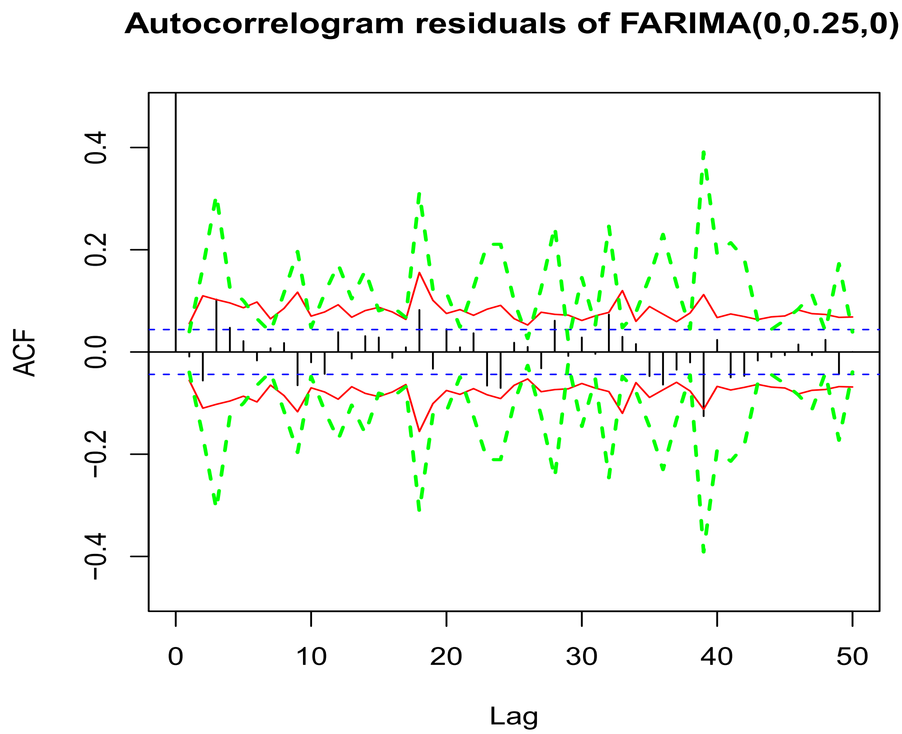

Fig 9. *Autocorrelation of a realization of size* $n = 2,000$ *for a weak FARIMA*$(0, 0.25, 0)$ *model* (19)–(20) *with* $\theta_0 = (0, 0, 0.25)$ *and where* $\omega = 0.04$, $\alpha_1 = 0.13$ *and* $\beta_1 = 0.88$. *The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong FARIMA assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak FARIMA assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 2. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 7.*

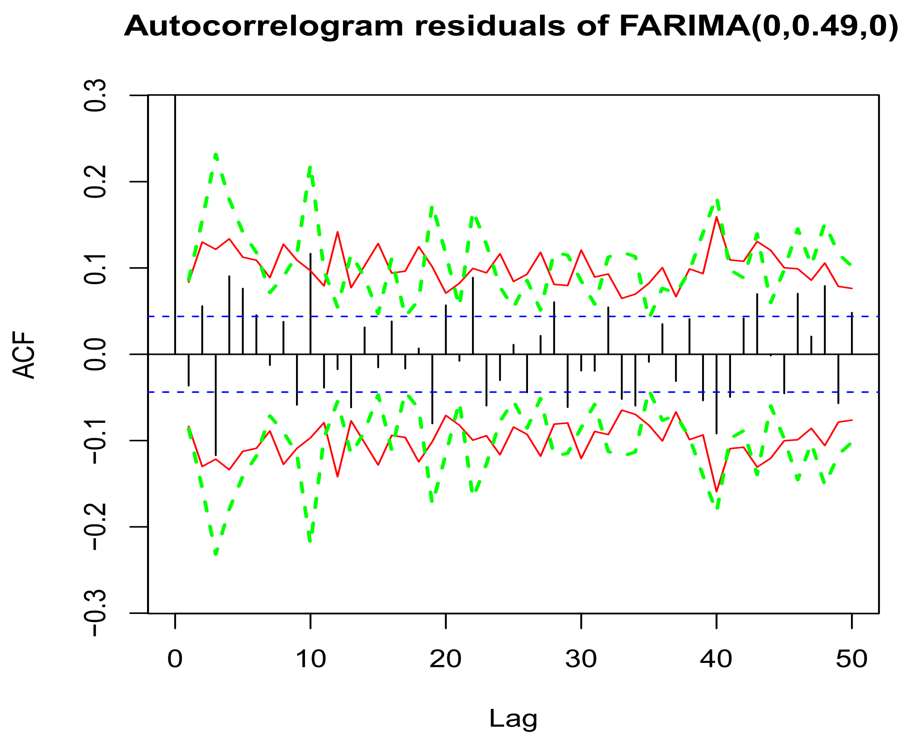

Fig 10. *Autocorrelation of a realization of size* $n = 2,000$ *for a weak FARIMA*$(0, 0.49, 0)$ *model* (19)–(20) *with* $\theta_0 = (0, 0, 0.49)$ *and where* $\omega = 0.04$, $\alpha_1 = 0.13$ *and* $\beta_1 = 0.88$. *The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong FARIMA assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak FARIMA assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 2. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 7.*

References

[1] Akutowicz, E. J. (1957). On an explicit formula in linear least squares prediction. *Math. Scand.* **5** 261–266. MR97854

[2] Baillie, R. T., Chung, C.-F. and Tieslau, M. A. (1996). Analysing inflation by the fractionally integrated ARFIMA-GARCH model. *Journal of applied econometrics* **11** 23–40.

[3] Beran, J., Feng, Y., Ghosh, S. and Kulik, R. (2013). *Long-memory processes*. Springer, Heidelberg Probabilistic properties and statistical methods. MR3075595

[4] Berk, K. N. (1974). Consistent autoregressive spectral estimates. *Ann. Statist.* **2** 489–502. Collection of articles dedicated to Jerzy Neyman on his 80th birthday. MR0342345

[5] Boubacar Mainassara, Y., Carbon, M. and Francq, C. (2012). Computing and estimating information matrices of weak ARMA models. *Comput. Statist. Data Anal.* **56** 345–361. MR2842343

[6] Boubacar Maïnassara, Y., Esstafa, Y. and Saussereau, B. (2021). Estimating FARIMA models with uncorrelated but non-independent error terms. *Stat. Inference Stoch. Process.* **24** 549–608. MR4321851

[7] Boubacar Maïnassara, Y. and Saussereau, B. (2018). Diagnostic Checking in Multivariate ARMA Models With Dependent Errors Using Normalized Residual Autocorrelations. *J. Amer. Statist. Assoc.* **113** 1813–1827. MR3902248

[8] Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Amer. Statist. Assoc.* **65** 1509–1526. MR0273762

[9] Brockwell, P. J. and Davis, R. A. (1991). *Time series: theory and methods*, second ed. *Springer Series in Statistics*. Springer-Verlag, New York. MR1093459

[10] Cavaliere, G., Nielsen, M. Ø. and Taylor, A. M. R. (2017). Quasi-maximum likelihood estimation and bootstrap inference in fractional time series models with heteroskedasticity of unknown form. *Journal of Econometrics* **198** 165–188. MR3628104

[11] Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. *Ann. Statist.* **17** 1749–1766. MR1026311

[12] Davidson, J. (1994). *Stochastic limit theory. Advanced Texts in Econometrics*. The Clarendon Press, Oxford University Press, New York An introduction for econometricians. MR1430804

[13] Delgado, M. A. and Velasco, C. (2011). An asymptotically pivotal transform of the residuals sample autocorrelations with application to model checking. *J. Amer. Statist. Assoc.* **106** 946–958. MR2894755

[14] den Haan, W. J. and Levin, A. T. (1997). A practitioner's guide to robust covariance matrix estimation. In *Robust inference. Handbook of Statist.* **15** 299–342. North-Holland, Amsterdam. MR1492717

[15] Ding, Z., Granger, C. W. and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of empirical*

finance **1** 83–106.

[16] Doukhan, P. and León, J. (1989). Cumulants for stationary mixing random sequences and applications to empirical spectral density. *Probab. Math. Stat* **10** 11–26. MR0990396

[17] Fan, J. and Yao, Q. (2008). *Nonlinear time series: nonparametric and parametric methods*. Springer Science & Business Media. MR1964455

[18] Fox, R. and Taqqu, M. S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.* **14** 517–532. MR0840512

[19] Francq, C., Roy, R. and Zakoïan, J.-M. (2005). Diagnostic Checking in ARMA Models With Uncorrelated Errors. *J. Amer. Statist. Assoc.* **100** 532–544. MR2160557

[20] Francq, C. and Zakoïan, J.-M. (1998). Estimating linear representations of nonlinear processes. *J. Statist. Plann. Inference* **68** 145–165. MR1627612

[21] Francq, C. and Zakoïan, J.-M. (2005). Recent results for linear time series models with non independent innovations. In *Statistical modeling and analysis for complex data problems. GERAD 25th Anniv. Ser.* **1** 241–265. Springer, New York. MR2189539

[22] Francq, C. and Zakoïan, J.-M. (2007). HAC estimation and strong linearity testing in weak ARMA models. *J. Multivariate Anal.* **98** 114–144. MR2292920

[23] Francq, C. and Zakoïan, J.-M. (2009). Bartlett's formula for a general class of nonlinear processes. *J. Time Series Anal.* **30** 449–465. MR2536062

[24] Francq, C. and Zakoïan, J.-M. (2010). *GARCH Models: Structure, Statistical Inference and Financial Applications*. Wiley.

[25] Granger, C. W. J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *J. Time Ser. Anal.* **1** 15–29. MR0605572

[26] Hallin, M., Taniguchi, M., Serroukh, A. and Choy, K. (1999). Local asymptotic normality for regression models with long-memory disturbance. *Ann. Statist.* **27** 2054–2080. MR1765628

[27] Herrndorf, N. (1984). A functional central limit theorem for weakly dependent sequences of random variables. *Ann. Probab.* **12** 141–153. MR0723735

[28] Hong, Y. (1996). Consistent testing for serial correlation of unknown form. *Econometrica* **64** 837–864. MR1399220

[29] Hosking, J. R. M. (1981). Fractional differencing. *Biometrika* **68** 165–176. MR0614953

[30] Hualde, J. and Robinson, P. M. (2011). Gaussian pseudo-maximum likelihood estimation of fractional time series models. *Ann. Statist.* **39** 3152–3181. MR3012404

[31] Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika* **48** 419–426. MR0137199

[32] Kuan, C.-M. and Lee, W.-M. (2006). Robust M tests without consistent estimation of the asymptotic covariance matrix. *J. Amer. Statist. Assoc.*

101 1264–1275. MR2328312

[33] Li, W. K. and McLeod, A. I. (1986). Fractional time series modelling. *Biometrika* **73** 217–221. MR0836451

[34] Ling, S. (2003). Adaptive estimators and tests of stationary and non-stationary short- and long-memory ARFIMA-GARCH models. *J. Amer. Statist. Assoc.* **98** 955–967. MR2041484

[35] Ling, S. and Li, W. K. (1997). On fractionally integrated autoregressive moving-average time series models with conditional heteroscedasticity. *J. Amer. Statist. Assoc.* **92** 1184–1194. MR1482150

[36] Ljung, G. M. and Box, G. E. P. (1978). On a Measure of Lack of Fit in Time Series Models. *Biometrika* **65** pp. 297–303.

[37] Lobato, I. N. (2001). Testing that a dependent process is uncorrelated. *J. Amer. Statist. Assoc.* **96** 1066–1076. MR1947254

[38] Lobato, I. N., Nankervis, J. C. and Savin, N. E. (2001). Testing for Autocorrelation Using a Modified Box- Pierce Q Test. *Inter. Econ. Review* **42** 187–205. MR1816014

[39] Lobato, I. N., Nankervis, J. C. and Savin, N. E. (2002). Testing for Zero Autocorrelation in the Presence of Statistical Dependence. *Econ. Theory* **18** 730–743. MR1906333

[40] Magnus, J. R. and Neudecker, H. (1999). *Matrix differential calculus with applications in statistics and econometrics. Wiley Series in Probability and Statistics*. John Wiley & Sons, Ltd., Chichester Revised reprint of the 1988 original. MR1698873

[41] Nielsen, M. Ø. (2015). Asymptotics for the Conditional-Sum-of-Squares Estimator in Multivariate Fractional Time-Series Models. *Journal of Time Series Analysis* **36** 154–188. MR3316465

[42] Palma, W. (2007). *Long-memory time series. Wiley Series in Probability and Statistics*. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ Theory and methods. MR2297359

[43] Romano, J. P. and Thombs, L. A. (1996). Inference for autocorrelations under weak assumptions. *J. Amer. Statist. Assoc.* **91** 590–600. MR1395728

[44] Shao, X. (2010). A self-normalized approach to confidence interval construction in time series. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **72** 343–366. MR2758116

[45] Shao, X. (2010). Corrigendum: A self-normalized approach to confidence interval construction in time series. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **72** 695–696. MR2758242

[46] Shao, X. (2011). Testing for white noise under unknown dependence and its applications to diagnostic checking for time series models. *Econometric Theory* **27** 312–343. MR2782041

[47] Shao, X. (2012). Parametric inference in stationary time series models with dependent errors. *Scand. J. Stat.* **39** 772–783. MR3000848

[48] Shao, X. (2015). Self-normalization for time series: a review of recent developments. *J. Amer. Statist. Assoc.* **110** 1797–1817. MR3449074

[49] Tong, H. (1990). *Non-linear time series: a dynamical system approach.* Oxford University Press. MR1079320

[50] van der Vaart, A. W. (1998). *Asymptotic statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge University Press, Cambridge. MR1652247

[51] Velasco, C. and Wang, X. (2015). A joint portmanteau test for conditional mean and variance time-series models. *J. Time Series Anal.* **36** 39–60. MR3300204

[52] Wang, X. and Sun, Y. (2020). A Simple Asymptotically F-Distributed Portmanteau Test for Diagnostic Checking of Time Series Models with Uncorrelated Innovations. *Journal of Business & Economic Statistics* **0** 1–40. MR4410878

[53] Zhu, K. and Li, W. K. (2015). A bootstrapped spectral test for adequacy in weak ARMA models. *J. Econometrics* **187** 113–130. MR3347298