Modified portmanteau tests for ARMA models with dependent errors: a self-normalisation approach

Yacouba Boubacar Maïnassara and **Bruno Saussereau**

Laboratoire de mathématiques de Besançon,
UMR CNRS 6623,
16 route de Gray,
25030 Besançon, France.

e-mail: yacouba.boubacar_mainassara@univ-fcomte.fr; bruno.saussereau@univ-fcomte.fr

Abstract: In this paper, we derive the asymptotic distribution of residual empirical autocovariances and autocorrelations under weak assumptions on the noise. We propose new portmanteau statistics for autoregressive moving-average (ARMA) models with uncorrelated but non-independent innovations by using a self-normalisation approach. We establish the asymptotic distribution of the proposed statistics. This asymptotic distribution is quite different from the usual chi-squared approximation used under the independent and identically distributed assumptions on the noise, or the weighted sum of independent chi-squared random variables obtained under nonindependent innovations. A set of Monte Carlo experiments, and an application to the daily returns of the CAC40 is presented.

AMS 2000 subject classifications: Primary 62M10, 62F03, 62F05; secondary 91B84, 62P05.
Keywords and phrases: Goodness-of-fit test, least-square estimation, Box-Pierce and Ljung-Box portmanteau tests, residual autocorrelation, self-normalization, weak ARMA models.

1. Introduction

In 1970, Box and Pierce have proposed (see [4]) a goodness-of-fit test, the so-called portmanteau test, for univariate strong ARMA models, namely models with an independent and identically distributed noise sequence. A modification of their test has been proposed by Ljung and Box (see [13]). It is nowadays one of the most popular diagnostic checking tools in ARMA modeling of time series. Both of theses tests are based on the residual empirical autocorrelations $\hat{\rho}(h)$. The Box-Pierce statistic is defined by

$$Q_m^{\mathrm{BP}} = n \sum_{h=1}^{m} \hat{\rho}^2(h), \tag{1}$$

where n is the length of the series and m is a fixed integer. If the noise sequence is independent and identically distributed, the standard test procedure consists in rejecting the null hypothesis of an ARMA(p,q) models if the statistic Q_m^{BP} is greater than a certain quantile of a chi-squared distribution. We mention that the Ljung-Box statistic defined by

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

has the same asymptotic distribution as Q_m^{BP} and has the reputation of doing better for small or medium sized sample (see [13]). Nowadays, the time series literature has shown a growing interest in non-linear models. Roughly speaking, these models are those for which the assumption that the noise is a sequence of independent and identically distributed random variables is relaxed. A detailed and concise presentation is proposed in the next section.

Henceforth, we deal with some models with uncorrelated but dependent noise process. For such models, the asymptotic distribution of the Box-Pierce statistic is no more a chi-square distribution but a mixture of m chi-squared distributions, weighted by eigenvalues of the asymptotic covariance matrix of the vector of autocorrelations (see [8]).

In this work, we propose an alternative method that do not estimate an asymptotic covariance matrix. It is based on a new self-normalization based approach to construct a new test-statistic that is asymptotically distribution-free under the null hypothesis. The idea comes from Lobato (see [14]) and has been already used in [16]. Indeed, under some technical condition, Shao has constructed some confidence region for the parameter vector for fractional autoregressive integrated moving average process. Our new test-statistic is defined by

$$Q_m^{\mathrm{SN}} = n\,\hat{\sigma}^4\,\hat{\rho}_m'\hat{C}_m^{-1}\hat{\rho}_m, \tag{3}$$

where $\hat{C}_m^{-1}$ is a normalization matrix that is observable, $\hat{\sigma}^2$ is the estimator of the common variance of the noise process and $\hat{\rho}_m$ is the vector of the first m sample autocorrelations. We prove in Theorem 2, that the asymptotic distribution of Q_m^{SN} is the distribution of a random variable $\mathcal{U}_m$ that depends only on m and is free of all the parameters of the model. It as an explicit expression by means of Brownian bridges but its law is not explicitly known. Nevertheless, it can be easily tabulated by Monte-Carlo experiments (see Table 1 in [14] or our own table given by Table 1). We emphasize the fact that in [8], the authors have proposed some modified versions of the Box-Pierce and Ljung-Box statistics that are more difficult to implement because their critical values have to be computed from the data, whereas those of our new portmanteau statistics are not computed from the data since they are tabulated. In some sense, our method is finally closer to the standard versions which are simply given in a χ^2-table.

In Monte Carlo experiments, we illustrate that the proposed test statistics Q_m^{SN} have reasonable finite sample performance, at least for the models considered in our study. Under nonindependent errors, it appears that the standard test statistics are generally unreliable, overrejecting severely, while the proposed test statistics offers satisfactory levels in most cases. Even for independent errors, the proposed versions may be preferable to the standard ones, when the number m of autocorrelations is small. Moreover, the error of first kind is well controlled. For all these reasons, we think that the modified versions that we propose in this paper are preferable to the standard ones for diagnosing ARMA models under nonindependent errors.

The paper is organised as follows. In the next section, we introduce the model and our methodology. The main results (Theorem 1 and Theorem 2) are given in Section 3. Their proofs are postponed to Section 6. Numerical illustrations are presented in Section 4. An application to the daily returns of the CAC40 is presented in Section 5. The associated numerical tables are gathered in Section 7, after the bibliography.

We complete this introduction by some basic notations that will be used throughout this paper. We denote by A' the transpose of a matrix A. We denote by $\xrightarrow{\mathrm{d}}$ and $\xrightarrow{\mathbb{P}}$ the convergence in distribution and in probability, respectively. The symbol $\mathrm{o}_{\mathbb{P}}(1)$ is used for a sequence of random variables that converges to zero in probability. We recall that the Skorokhod space $\mathbb{D}^d[0,1]$ is the set of $\mathbb{R}^d-$valued functions defined on [0,1] which are right continuous and have left limits. It is endowed with the Skorokhod topology and the weak convergence on $\mathbb{D}^d[0,1]$ is mentioned by $\xrightarrow{\mathbb{D}^d}$. We finally denote by $\lfloor a \rfloor$ the integer part of the real a. The identity matrix of order d is denoted by I_d.

2. Model, assumptions and methodology

First, we introduce the notions of weak and strong ARMA representations, which differ by the assumptions on the error terms. Then we recall some results concerning the estimation of weak ARMA models and we present the methodology of our test, which is that based on the self-normalization procedure.

2.1. Strong and weak ARMA representations

In order to deal with quite general linear models, it is usually assumed that the linear innovations, which are uncorrelated by construction, are not independent, nor martingale differences. Indeed, the Wold decomposition (see Section 5.7 in [5]) stipulates that any zero-mean purely non deterministic stationary process can be expressed as

$$X_t = \sum_{\ell=0}^{\infty} \varphi_\ell \, \epsilon_{t-\ell}, \tag{4}$$

where $\varphi_0 = 1$, $\sum_{l=0}^{\infty} \varphi_\ell^2 < \infty$, and the linear innovation process $\epsilon := (\epsilon_t)_{t\in\mathbb{Z}}$ is assumed to be a stationary sequence of centered and uncorrelated random variables with common variance σ^2. Under the above assumptions, the process ϵ is called a weak white noise. In practice the sequence $(\varphi_\ell)_{\ell\geq 0}$ is often parameterized by assuming that the process $X = (X_t)_{t\in\mathbb{Z}}$ admits an ARMA(p,q) representation, *i.e.* that there exists integers p and q and constants $a_1, \dots, a_p, b_1, \dots, b_q$, such that for any $t \in \mathbb{Z}$

$$X_t - \sum_{i=1}^{p} a_i X_{t-i} = \epsilon_t + \sum_{j=1}^{q} b_j \epsilon_{t-j}. \tag{5}$$

This representation is said to be a weak ARMA(p,q) representation under the assumption that ϵ is a weak white noise. For the statistical inference of ARMA models, the weak white noise assumption is often replaced by the strong white noise assumption, *i.e.* the assumption that ϵ is an independent and identically distributed (i.i.d. for short) sequence of random variables with mean 0 and common variance σ^2. An intermediate assumption for the noise when ϵ is that is a stationary martingale-difference sequence. In such a case, it holds that ϵ is a stationary sequence satisfying $\mathbb{E}\left(\epsilon_t \mid \epsilon_u, u < t\right) = 0$ and $\mathrm{Var}(\epsilon_t) = \sigma^2$. An interesting example of such a noise is the generalized autoregressive conditional heteroscedastic (GARCH) model. At last, a process $(X_t)_{t\in\mathbb{Z}}$ is said to be linear when the sequence ϵ is i.i.d. in the representation (4), and is said to be non-linear in the opposite case. With this definition, GARCH-type processes are considered as non-linear. Linear and non-linear processes may also have exact weak ARMA representations, and many important classes of nonlinear processes admit weak ARMA representations (see [8] and the references therein). Obviously the strong white noise assumption is more restrictive than the weak white noise assumption, because independence entails uncorrelatedness. Consequently weak ARMA representations are more general than the strong ones. We end this general presentation by recalling that any process satisfying (4) is the limit, in L^2 as n tends to ∞, of a sequence of processes satisfying weak ARMA(p_n, q_n) representations (see [11] page 244). In this sense, the subclass of the processes admitting weak ARMA(p_n, q_n) representations is dense in the set of the purely non deterministic stationary processes.

2.2. Estimating weak ARMA representations

Now, we present the least square estimation procedure of the parameter of an ARMA model as well as the asymptotic behavior of the least squares estimator (LSE in short). The LSE method is the standard estimation procedure for ARMA models and it coincides with the maximum-likelihood estimator in the Gaussian case. As usual, it will be convenient to write (5) as

$$a(L)X_t = b(L)\epsilon_t,$$

where L is the backshift operator, $a(z) = 1 - \sum_{i=1}^{p} a_i z^i$ is the AR polynomial (AR stands for auto regressive) and $b(z) = 1 + \sum_{j=1}^{q} b_j z^j$ is the MA polynomial (MA stands for moving average) . The unknown parameter $\theta_0 = (a_1, \dots, a_p, b_1, \dots, b_q)'$ is supposed to belong to the interior of a compact subspace Θ^* of the following parameter space

$$\Theta := \Big\{ \theta = (\theta_1, \dots, \theta_p, \theta_{p+1}, ..., \theta_{p+q})' \in \mathbb{R}^{k_0},\ k_0 := p + q : \\ a(z) = 1 - \sum_{i=1}^{p} a_i z^i \text{ and } b(z) = 1 + \sum_{i=1}^{q} b_i z^i \\ \text{have all their zeros outside the unit disk} \Big\}.$$

For a $\theta \in \Theta$, the polynomial functions a and b have all their zeros outside the unit disk and we also assume that a and b have no zero in common. As for the usual strong ARMA models, it is assumed that $p + q > 0$ and $a_p^2 + b_q^2 \neq 0$ (by convention $a_0 = b_0 = 1$).

For all $\theta \in \Theta$, let

$$\epsilon_t(\theta) = b^{-1}(L)a(L)X_t = X_t + \sum_{i=1}^{\infty} c_i(\theta) X_{t-i}.$$

Note that for any $t \in \mathbb{Z}$, $\epsilon_t(\theta_0) = \epsilon_t$ almost-surely. Given a realization $X_1, X_2, \dots, X_n$ of length n, $\epsilon_t(\theta)$ can be approximated, for $0 < t \leq n$, by $e_t(\theta)$ which is defined recursively by

$$e_t(\theta) = X_t - \sum_{i=1}^{p} \theta_i X_{t-i} - \sum_{i=1}^{q} \theta_{p+i} e_{t-i}(\theta) \tag{6}$$

where the unknown starting values are set to zero:

$$e_0(\theta) = e_{-1}(\theta) = \dots = e_{-q+1}(\theta) = X_0 = X_{-1} = \dots = X_{-p+1} = 0 \ .$$

The random variable $\hat{\theta}_n$ is called LSE if it satisfies, almost surely,

$$Q_n(\hat{\theta}_n) = \min_{\theta \in \Theta^*} Q_n(\theta), \text{ with}$$

$$Q_n(\theta) = \frac{1}{2n} \sum_{t=1}^{n} e_t^2(\theta). \tag{7}$$

In order to state our results, we shall need the functions

$$O_n(\theta) = \frac{1}{2n}\sum_{t=1}^{n} \epsilon_t^2(\theta). \tag{8}$$

As in [9], we shall work with a noise that will have the strong mixing property. We recall that the strong mixing coefficients of a stationary process $Z = (Z_t)_{t\in\mathbb{T}}$ (here $\mathbb{T}$ denotes the set of indices of the process Z that may be $\mathbb{R}$, $\mathbb{Z}$ or any subset of these sets) are defined by

$$\alpha_Z(h) = \sup_{A\in\sigma(Z_u,u\leq t),B\in\sigma(X_u,u\geq t+h)} \left|\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)\right|.$$

For a given h, $\alpha_Z(h)$ is measuring the temporal dependence of the process Z at time lag h. The asymptotic behavior of the LSE is well known in the strong ARMA case. This assumption being very restrictive, a weak ARMA representation of stationary processes satisfying $\mathbb{E}(|X_t|^{4+2\nu}) < \infty$ and $\sum_{k=0}^{\infty}\{\alpha_X(k)\}^{\frac{\nu}{2+\nu}} < \infty$ for some $\nu > 0$, has been considered in [9]. Hereafter, it has been noted in [11] that the strong mixing and the moments conditions on the process X can be replaced by the following ones on the noise process ϵ:

$$\textbf{(A1)}:\ \mathbb{E}(|\epsilon_t|^{4+2\nu}) < \infty \text{ and } \sum_{k=0}^{\infty}\{\alpha_\epsilon(k)\}^{\frac{\nu}{2+\nu}} < \infty \text{ for some } \nu > 0\ ; \tag{9}$$

$$\textbf{(A2)}:\ \epsilon \text{ is ergodic and stationnary.} \tag{10}$$

Note that assumption **(A1)** does not require independence of the noise, nor the fact that it is a martingale difference. The above hypotheses will be assumed in all our work.

In [9], it is proved that if $(X_t)_{t\in\mathbb{Z}}$ is a strictly stationary and ergodic process satisfying the ARMA model (5) with a weak white noise ϵ, then under the Assumption **A1**, it holds that $\hat{\theta}_n$ is a consistent estimator of θ_0 and

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow[n\to\infty]{\mathrm{d}} \mathcal{N}(0, \mathrm{V}(\theta_0)), \tag{11}$$

where $\mathrm{V}(\theta_0) = J(\theta_0)^{-1}I(\theta_0)J(\theta_0)^{-1}$, with

$$I(\theta) = \lim_{n\to\infty} \mathrm{Var}\left(\sqrt{n}\frac{\partial}{\partial\theta}Q_n(\theta)\right) \quad \text{and}$$
$$J(\theta) = \lim_{n\to\infty} \frac{\partial^2 Q_n(\theta)}{\partial\theta\partial\theta'} \text{ a.s..} \tag{12}$$

The matrix J can easily be estimated by its empirical counterpart (see Theorem 2 in [10])

$$\hat{J}_n = \frac{1}{n}\sum_{t=1}^{n}\frac{\partial e_t(\hat{\theta}_n)}{\partial\theta}\frac{\partial e_t(\hat{\theta}_n)}{\partial\theta'} \tag{13}$$

because we have (see the proof of Lemma A4 in [10])

$$J(\theta_0) = \mathbb{E}\left[\frac{\partial\epsilon_t(\theta_0)}{\partial\theta}\frac{\partial\epsilon_t(\theta_0)}{\partial\theta'}\right] \tag{14}$$

Some details on this estimation will be given hereafter. The matrices $J(\theta)$ and $I(\theta)$ are usually called information matrices. They are involved in the inference steps, such as in portmanteau adequacy tests (see [8]).

2.3. Methodology of test

As mentioned in the introduction, the goodness-of-fit test are based on the residual autocorrelations. We introduce further notations to explain our strategy.

For a time $t \geq 0$, let $\hat{e}_t = e_t(\hat{\theta}_n)$ be the least-square residuals when $p > 0$ or $q > 0$, and let $\hat{e}_t = e_t = X_t$ when $p = q = 0$. When $p + q \neq 0$, we use (6) to notice that we have $\hat{e}_t = 0$ for $t \leq 0$ and $t > n$. By (5), it holds that

$$\hat{e}_t = X_t - \sum_{i=1}^{p} \hat{\theta}_i \, \hat{X}_{t-i} + \sum_{i=1}^{q} \hat{\theta}_{p+i} \, \hat{e}_{t-i}, \tag{15}$$

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

We denote

$$\gamma(h) = \frac{1}{n} \sum_{t=h+1}^{n} \epsilon_t \, \epsilon_{t-h}$$

the white noise "empirical" autocovariances. It should be noted that $\gamma(h)$ is not a statistic (unless if $p = q = 0$) because it depends on the unobserved innovations ϵ. The residual autocovariances are defined by

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=h+1}^{n} \hat{e}_t \, \hat{e}_{t-h}.$$

It is worth to notice that the residual autocovariances are observable. For a fixed integer $m \geq 1$, we denote

$$\gamma_m = (\gamma(1), \dots, \gamma(m))' \text{ and } \hat{\gamma}_m = (\hat{\gamma}(1), \dots, \hat{\gamma}(m))'.$$

The theoretical and sample autocorrelations at lag ℓ are respectively defined by $\rho(\ell) = \gamma(\ell)/\gamma(0)$ and $\hat{\rho}(\ell) = \hat{\gamma}(\ell)/\hat{\gamma}(0)$, with $\gamma(0) := \sigma^2$. In the sequel, we will also need the vector of the first m sample autocorrelations

$$\hat{\rho}_m = (\hat{\rho}(1), \dots, \hat{\rho}(m))'.$$

Based on the residual empirical autocorrelations $\hat{\rho}(h)$, the Box-Pierce statistic is defined in (1) and is used to test the null hypothesis

(H0) : $(X_t)_{t\in\mathbb{Z}}$ satisfies an ARMA(p, q) representation;

against the alternative

(H1) : $(X_t)_{t\in\mathbb{Z}}$ does not admit an ARMA(p, q) representation or $(X_t)_{t\in\mathbb{Z}}$ satisfies an ARMA(p', q') representation with $p' > p$ or $q' > q$.

These tests are very useful tools for checking the overall significance of the residual autocorrelations. Under the assumption that the data generating process follows a strong ARMA(p, q) model, the asymptotic distribution of the statistic Q_m^{BP} is generally approximated by the $\chi^2_{m-k_0}$ distribution with $m > k_0$.

The main objective is to investigate the asymptotic behaviour of the law of Q_m^{SN}, or a modification of it. It is obvious that this task will be a consequence of the study of the limit of the vector $\hat{\gamma}_m$.

2.4. Self-normalization

When the noise process in (5) is not independent, it has been shown in [8] that the asymptotic covariance matrix of the sample autocorrelations depends on the data generation process and the unknown parameters. To derive the asymptotic distribution of the portmanteau statistics under weak assumptions on the noise, one needs a consistent estimator of the asymptotic covariance matrix $\Sigma_{\hat{\rho}_m}$ of a vector of autocorrelations for residuals. In the econometric literature the nonparametric kernel estimator, also called heteroscedastic autocorrelation consistent estimator (see [1, 15]), is widely used to estimate covariance matrices. However, this causes serious difficulties as regard to the choice of the sequence of weights. An alternative method consists in using a parametric autoregressive estimate of the spectral density of a stationary process. This approach, which has been studied in [2, 7], is also confronted to the problem of choosing the truncation parameter. Indeed, this method is based on an infinite autoregressive representation of the stationary process. So the choice of the order of truncation is crucial and difficult. The methodology employed in [8] is, to our knowledge, the only work that deals with goodness-of-fit tests for weak ARMA models (an extension is proposed in [3] for multivariate ARMA models). Nevertheless, their method presents both difficulties that we just discussed: it supposes to weight appropriately some empirical fourth-order moments by means of a window and a truncation point. To circumvent the problem, we propose to adopt a self-normalization approach as in [14] and [16].

The self-normalization procedure is very well described in Section 2 of the paper of Lobato (see [14]). We also refer to [16] (Section 2.1). It has the outstanding advantage to be free of nuisance parameters such as the choice of the bandwidth in the truncation or the order of the truncation. We present here the main difficulties and differences that appear in our framework. More precisely, it will arise, after some quite technical considerations, that if we denote Γ the matrix in $\mathbb{R}^{m\times(p+q+m)}$ defined in block formed by

$$\Gamma = \Big(-\Phi_m J^{-1} | I_m \Big), \tag{16}$$

where J is defined in (12) and Φ_m is defined by

$$\Phi_m = \mathbb{E}\left\{ \begin{pmatrix} \epsilon_{t-1} \\ \vdots \\ \epsilon_{t-m} \end{pmatrix} \frac{\partial \epsilon_t(\theta_0)}{\partial \theta'} \right\}, \tag{17}$$

then one may write

$$\sqrt{n}\ \hat{\gamma}_m = \frac{1}{\sqrt{n}}\ \Gamma U_t + \mathrm{o}_{\mathbb{P}}(1),$$

with

$$U_t = \left(-\epsilon_t(\theta_0) \frac{\partial \epsilon_t(\theta_0)}{\partial \theta}, \epsilon_t\epsilon_{t-1}, ..., \epsilon_t\epsilon_{t-m} \right)' . \tag{18}$$

The above expression comes from (44) in Section 6. At this stage, we do not rely on the classical method that would consist in estimating the asymptotic covariance matrix of ΓU_t. We rather try to apply Lemma 1 in [14]. So we need to check that a functional central limit theorem holds for the process $U := (U_t)_{t\geq 1}$.

If there were no first entry of U_t, we would be in the context of Lobato and the functional central limit theorem would be clear thanks to the mixing condition on the noise process ϵ. Unfortunately, it is more difficult to deal with the process U itself. In order to prove that $\frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor nr\rfloor}\Gamma U_j$ converges on the Skorokhod space to a Brownian motion, we will employ a cutoff argument on the representation of $\partial\epsilon_t(\theta_0)/\partial\theta$ as an infinite sum of the past of the noise (see (38) for precisions). This difficulty is overcome in the proof of Theorem 1 (see Subsection 6.2.2).

We are now able to state the following theorems, which are the main results of our work.

3. Main results

We begin by introducing further notations. Let $(B_m(r))_{r\geq 0}$ be a m-dimensional Brownian motion starting from 0. For $m\geq 1$, we denote $\mathcal{U}_m$ the random variable defined by

$$\mathcal{U}_m = B_m'(1)V_m^{-1}B_m(1) \tag{19}$$

where

$$V_m = \int_0^1 \left(B_m(r)-rB_m(1)\right)\left(B_m(r)-rB_m(1)\right)' dr. \tag{20}$$

The critical values for $\mathcal{U}_m$ have been tabulated by Lobato in [14]. We have produce our own table (see Table 1 in Section 7).

Finally, we denote

$$S_t = \frac{1}{n}\sum_{j=1}^{t}\Gamma\left(U_j - \frac{1}{n}\sum_{i=1}^{n}U_i\right) \tag{21}$$

where the process U is defined by (18).

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

Theorem 1. *Assume that $p>0$ or $q>0$. Under assumptions* (**A1**) *and* (**A2**), *we have*

$$n\,\hat{\gamma}_m' C_m^{-1}\hat{\gamma}_m \xrightarrow[n\to\infty]{\mathrm{d}} \mathcal{U}_m \tag{22}$$

where the normalization matrix $C_m\in\mathbb{R}^{m\times m}$ is defined by

$$C_m = \frac{1}{n^2}\sum_{t=1}^{n} S_tS_t'. \tag{23}$$

The sample autocorrelations satisfy

$$n\,\sigma^4\,\hat{\rho}_m' C_m^{-1}\hat{\rho}_m \xrightarrow[n\to\infty]{\mathrm{d}} \mathcal{U}_m \tag{24}$$

The proof of this result is postponed to Section 6.

Of course, the above theorem is useless for practical purpose, because it does not involve any observable quantities. This gap will be fixed hereafter. Right now, we make several important comments that are necessary to compare Theorem 1 with the existing results.

Remark 1. *When $p = q = 0$, we do no more need to estimate the unknown parameter θ_0. Thus a careful reading of the proofs shows that the vector U_t is replaced by*

$$\tilde{U}_t = (\epsilon_t\epsilon_{t-1}, ..., \epsilon_t\epsilon_{t-m})'$$

and Γ is replaced by the identity matrix. Then we obtain the result of Lobato (see Lemma 1 in [14]) and we thus generalized his result to the ARMA model.

Remark 2. *In [16], the author has proposed a self-normalization approach to the construction of confidence regions for the parameter vector θ. The unknown parameter is estimated by the Whittle estimator, but this difference is not significant. The striking difference consists in the lists of assumptions that the author has used. Even if they are natural, they are quite strong (especially the 8th moment condition) whereas our work only supposes that assumption* (**A1**)*, on the moments, is satisfied. Moreover, it is not clear to us that, in the framework of [16], a goodness-of-fit test can be obtained.*

As mentioned before, the above theorem has to be completed. To make it useful in practice, one has to replace the matrix C_m and the variance of the noise σ^2 by their empirical or observable counterparts. For C_m, the idea is to use $e_t(\hat{\theta}_n)$ instead of the unobservable noise $\epsilon_t(\theta_0)$. The matrix Φ_m can be easily estimated by its empirical counterpart

$$\hat{\Phi}_m = \frac{1}{n}\sum_{t=1}^{n}\left\{(\hat{e}_{t-1}, \ldots, \hat{e}_{t-m})'\frac{\partial e_t(\theta)}{\partial\theta'}\right\}_{\theta=\hat{\theta}_n}. \tag{25}$$

The matrix J is estimated by $\hat{J}_n$ defined in (13). Thus we define

$$\hat{\Gamma} = \Big(-2\hat{\Phi}_m\hat{J}_n^{-1}|I_m\Big). \tag{26}$$

Finally we denote

$$\hat{U}_t = \left(-\hat{e}_t\frac{\partial e_t(\theta)}{\partial\theta}, \hat{e}_t\hat{e}_{t-1}, ..., \hat{e}_t\hat{e}_{t-m}\right)'\Bigg|_{\theta=\hat{\theta}_n} \quad \text{and} \tag{27}$$

$$\hat{S}_t = \frac{1}{n}\sum_{j=1}^{t}\hat{\Gamma}\left(\hat{U}_j - \frac{1}{n}\sum_{i=1}^{n}\hat{U}_i\right). \tag{28}$$

The above quantities are all observable and we are able state our second theorem which is the applicable counterpart of Theorem 1.

Theorem 2. *Assume that $p > 0$ or $q > 0$. Under the assumptions* (**A1**) *and* (**A2**)*, we have*

$$n\,\hat{\gamma}_m'\hat{C}_m^{-1}\hat{\gamma}_m \xrightarrow[n\to\infty]{\mathrm{d}} \mathcal{U}_m \tag{29}$$

where the normalization matrix $\hat{C}_m \in \mathbb{R}^{m\times m}$ is defined by

$$\hat{C}_m = \frac{1}{n^2}\sum_{t=1}^{n}\hat{S}_t\hat{S}_t'. \tag{30}$$

The sample autocorrelations satisfy

$$Q_m^{\mathrm{SN}} = n\,\hat{\sigma}^4\,\hat{\rho}_m'\hat{C}_m^{-1}\hat{\rho}_m \xrightarrow[n\to\infty]{\mathrm{d}} \mathcal{U}_m, \tag{31}$$

where $\hat{\sigma}^2$ is a consistent estimator of the common variance of the noise process ϵ.

The proof of this result is postponed to Section 6.

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

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

where the matrix $D_{n,m}$ is diagonal with $(1/(n-1), ..., 1/(n-m))$ as diagonal terms.

Remark 3. *In [8], the authors proposed some modified versions of the Box-Pierce and Ljung-Box statistics that are more difficult to implement because their critical values have to be computed from the data, whereas those of our new portmanteau statistics are tabulated once for all. In some sense, our methods is finally closer to the standard versions who are simply given in a χ^2-table.*

4. Numerical illustrations

In this section, by means of Monte Carlo experiments, we investigate the finite sample properties of the modified version of the portmanteau tests introduced in this paper. The numerical illustrations of this section are made with the free statistical software R (see http://cran.r-project.org/). The tables are gathered in Section 7.

We indicate the conventions that we adopt in the discussion and in the tables. One refers to

- $\mathrm{BP_S}$ for the standard Box-Pierce test using the statistic Q_m^{BP};
- $\mathrm{LB_S}$ for the standard Ljung-Box test using the statistic Q_m^{LB};
- $\mathrm{BP_{SN}^{Lo}}$ for the modified test using the statistic Q_m^{SN} with the values of the quantiles of $\mathcal{U}_m$ simulated in Table 1 of [14];
- $\mathrm{BP_{SN}}$ for the modified test using the statistic Q_m^{SN} with the values of the quantiles of $\mathcal{U}_m$ that we have simulated in Table 1;
- $\mathrm{LB_{SN}}$ for the modified test using the statistic $\tilde{Q}_m^{\text{SN}}$ with the values of the quantiles of $\mathcal{U}_m$ that we have simulated in Table 1;
- $\mathrm{BP_{FRZ}}$ for the standard Ljung-Box test using the statistic Q_m^{BP} and the method presented in [8];
- $\mathrm{LB_{FRZ}}$ for the standard Box-Pierce test using the statistic Q_m^{BP} and the method presented in [8].

As an example, the $p-$value of the test $\mathrm{BP_{SN}}$ is denoted by $\mathrm{p}_{\text{SN}}^{\text{BP}}$. We adopt similar notations for the other tests.

We will see in the tables that the numerical results using the Ljung-Box tests are very close to those of the Box-Pierce tests. Nevertheless, they are still presented here, for the sake of completeness.

4.1. Simulated models

To generate the strong and the weak ARMA models, we consider the following ARMA$(1,1)$-ARCH(1) model

$$X_t = aX_{t-1} + \epsilon_t + b\epsilon_{t-1}, \tag{32}$$

with $\epsilon_t = \eta_t(1+\alpha_1\epsilon_{t-1}^2)^{1/2}$ and $(\eta_t)_{t\geq 1}$ a sequence of independent and identically distributed standard Gaussian random variable.

4.2. Empirical size

We simulated $N = 1,000$ independent trajectories of different size of model (32) with different values of the parameters $\theta_0 = (a, b)'$ and α_1. For each of these N replications, we estimated the coefficient θ_0 and we applied portmanteau tests to the residuals for different values of m, where m is the number of autocorrelations used in the portmanteau test statistic. The nominal asymptotic level of the tests is $\alpha = 5\%$.

4.2.1. Strong ARMA model case

To generate the strong ARMA model, we assume that $\alpha_1 = 0$ in (32). We first simulated N independent trajectories of size $n = 100$, $n = 300$ and $n = 1,000$ with the parameter $\theta_0 = (0, 0)'$ (see Table 2). Secondly, we simulated N independent trajectories of size $n = 500$ and $n = 1,000$ with the parameter $\theta_0 = (0.95, -0.6)'$ (see Table 3). For the standard Box-Pierce test, the model is therefore rejected when the statistic Q_m^{BP} is greater than $\chi^2_{(m-2)}(0.95)$. We know that the asymptotic level of this test is indeed $\alpha = 5\%$ when $\theta_0 = (0, 0)'$. Note however that, even when the noise is strong, the asymptotic level is not exactly $\alpha = 5\%$ when $\theta_0 \neq (0, 0)$.

For the modified Box-Pierce test (BP_{SN}), the model is rejected when the statistic Q_m^{SN} is greater than $\mathcal{U}_m(0.95)$, where the critical values $\mathcal{U}_m(0.95)$ are tabuled in Lobato (see Table 1 in [14] or our own table given by Table 1). When the roots of $(1-az)(1-bz) = 0$ are near the unit disk, the asymptotic distribution of Q_m^{SN} is likely to be far from its $\chi^2_{(m-2)}$ approximation. Table 2 displays the relative rejection frequencies of the null hypothesis H_0 that the data generating process follows a ARMA$(0, 0)$, over the N independent replications. As expected the observed relative rejection frequencies of the standard and modified Box-Pierce tests are close from the nominal $\alpha = 5\%$. Table 3 displays the relative rejection frequencies of the null hypothesis H_0 that the data generating process follows a ARMA$(1, 1)$, with the parameter $\theta_0 = (0.95, -0.6)'$, over the N independent replications. As expected the observed relative rejection frequency of the standard Box-Pierce test is very far from the nominal $\alpha = 5\%$ when the number m of autocorrelations used in the statistic is small. This is in accordance with the results in the literature on the standard ARMA models. The theory that the $\chi^2_{(m-2)}$ approximation is better for larger m is confirmed. In contrast, our modified Box-Pierce test well controls the error of first kind, even when m is small. Note that, for large n, when m is large the results are very similar for the standard and modified BP tests. Contrary to the standard test, the test based on Q_m^{SN} can be used safely for m small. Note that for $m \leq 2$, the empirical size is not available (n.a.) for the standard Box-Pierce test because this test is not applicable to $m \leq p + q$. From this example we draw the conclusion that, even in the strong ARMA case (with coefficients far from zero), the modified version is preferable to the standard one, when the number m of autocorrelations used is small.

4.2.2. Weak ARMA model case

We now repeat the same experiment on model (32) by assuming that $\alpha_1 = 0.4$ (*i.e.* a weak ARMA$(1, 1)$ model). We first simulated N independent trajectories, of this model, of size $n = 300$ and $n = 1,000$ with the parameter $\theta_0 = (0, 0)'$. Secondly, we simulated N independent trajectories of size $n = 1,000$ and $n = 2,000$ with the parameter $\theta_0 = (0.95, -0.6)'$.

As expected, Tables 4 and 5 show that the standard Box-Pierce test poorly performs to assess the adequacy of this weak ARMA models. In view of the observed relative rejection frequency, the standard test rejects very often the true ARMA$(0,0)$ or ARMA$(1,1)$ and all the relative rejection frequencies are very far from the nominal $\alpha = 5\%$. By contrast, the error of first kind is well controlled by the modified version of the BP test. We draw the conclusion that, for this particular weak ARMA model, the modified version is clearly preferable to the standard one.

4.3. Empirical power

In this part, we simulated $N = 1,000$ independent trajectories of different size of a ARMA$(2,1)$ defined by

$$X_t = X_{t-1} - 0.2X_{t-2} + \epsilon_t + 0.8\epsilon_{t-1}, \tag{33}$$

$$\tag{34}$$

where $\epsilon_t = \eta_t(1+\alpha_1\epsilon_{t-1}^2)^{1/2}$ and $(\eta_t)_{t\geq 1}$ is again a sequence of iid standard Gaussian random variables.

For each of these $N = 1,000$ replications we fitted a ARMA$(1,1)$ model and perform standard and modified Box-Pierce test based on $m = 1,2,3,4$, 6 and 10 residual autocorrelations. The adequacy of the ARMA$(1,1)$ model is rejected, in the strong and the weak ARMA cases, when the standard statistic Q_m^{BP} is greater than $\chi^2_{(m-2)}(0.95)$ and when the proposed statistic Q_m^{SN} is greater than $\mathcal{U}_m(0.95)$.

Tables 6 and 7 display the relative rejection frequencies of over the $N = 1,000$ independent replications. Table 6 shows that the powers of all the standard tests are very greater than the modified tests in the strong ARMA model. This is not surprising, because we have seen that, even in the strong ARMA case, the actual level of the standard version is generally much greater than the 5% nominal level, when the number m of autocorrelations used is small (see Table 3). By contrast, Table 7 shows that the powers of the modified tests are very similar in the weak ARMA case. The empirical powers of the standard tests are hardly interpretable for this weak ARMA model, because we have already seen in Table 5 that the standard versions of the tests do not well control the error of first kind in this particular weak ARMA framework.

5. Illustrative example

We now consider an application to the daily returns of the CAC40. The observations cover the period from the starting date of CAC40 to July 26, 2010. The length of the series is $n = 5154$. In Financial Econometrics, the returns are often assumed to be martingale increments (though they are not generally independent sequences), and the squares of the returns have often second-order moments close to those of an ARMA(1,1) (which is compatible with a GARCH(1,1) model for the returns). We will test these hypotheses by fitting weak ARMA models on the returns and on their squares.

First, we apply portmanteau tests for checking the hypothesis that the CAC40 returns constitute a white noise. Table 8 displays the statistics of the standard and modified BP tests.

Since the p-values of the standard test are very small, the white noise hypothesis is rejected at the nominal level $\alpha = 5\%$. This is not surprising because the standard tests required the iid assumption and, in particular in view of the so-called volatility clustering, it is well known that the strong white noise model is not adequate for these series. By contrast, the white noise hypothesis is not rejected by the modified tests, since for the modified tests, the statistic is not greater than the critical values (see Table 1). To summarize, the outputs of Table 8 are in accordance with the common belief that these series are not strong white noises, but could be weak white noises.

Next, turning to the dynamics of the squared returns, we fit an ARMA$(1,1)$ model to the squares of the CAC40 returns. Denoting by (X_t) the mean corrected series of the squared returns, we obtain the model

$$X_t = 0.98262 X_{t-1} + \epsilon_t - 0.89978 \epsilon_{t-1}, \text{ where } \mathrm{Var}(\epsilon_t) = 26.66141 \times 10^{-8}.$$

Table 9 displays the statistics of the standard and modified BP tests. From Table 9, we draw the same conclusion, on the squares of the previous daily returns, that the strong ARMA$(1,1)$ model is rejected, but a weak ARMA$(1,1)$ model is not rejected. Note that the first and second-order structures we found for the CAC40 returns, namely a weak white noise for the returns and a weak ARMA$(1,1)$ model for the squares of the returns, are compatible with a GARCH$(1,1)$ model.

6. Proofs

First, we shall need some technical results which are essentially contained in [8, 9, 10]. They are essential to understand the proof, but were not necessary to give the main ideas of the self-normalization approach. This is the reason why these facts are presented here.

6.1. Reminder on technical issues on quasi likelihood method for ARMA models

We recall that, given a realization $X_1, ..., X_n$ of length n, the noise $\epsilon_t(\theta)$ is approximated by $e_t(\theta)$ which is defined in (6).

The starting point in the asymptotic analysis, is the property that $\epsilon_t(\theta) - e_t(\theta)$ converges uniformly to 0 (almost-surely) as t goes to infinity. Similar properties also holds for the derivatives with respect to θ of $\epsilon_t(\theta) - e_t(\theta)$. We sum up the fact that we shall need in the sequel. We refer to the appendix of [10] (see also [9]) for a more detailed treatment.

For any $\theta \in \Theta$ and any $(l,m) \in \{1, ..., p+q\}^2$, there exists absolutely summable and deterministic sequences $(c_i(\theta))_{i\geq 0}$, $(c_{i,l}(\theta))_{i\geq 1}$ and $(c_{i,l,m}(\theta))_{i\geq 1}$ such that, almost surely,

$$\epsilon_t(\theta) = \sum_{i=0}^{\infty} c_i(\theta) X_{t-i} \ , \ \frac{\partial \epsilon_t(\theta)}{\partial \theta_l} = \sum_{i=1}^{\infty} c_{i,l}(\theta) X_{t-i} \text{ and } \frac{\partial^2 \epsilon_t(\theta)}{\partial \theta_l \partial \theta_m} = \sum_{i=2}^{\infty} c_{i,l,m}(\theta) X_{t-i} \tag{35}$$

$$e_t(\theta) = \sum_{i=0}^{t-1} c_i(\theta) X_{t-i} \ , \ \frac{\partial e_t(\theta)}{\partial \theta_l} = \sum_{i=1}^{t-1} c_{i,l}(\theta) X_{t-i} \text{ and } \frac{\partial^2 e_t(\theta)}{\partial \theta_l \partial \theta_m} = \sum_{i=2}^{t-1} c_{i,l,m}(\theta) X_{t-i}. \tag{36}$$

We strength the fact that, in the above identities, $c_0(\theta) = 1$.

A useful property of the sequences c, is that they are asymptotically exponentially small. Indeed, there exists $\rho \in [0,1[$ such that, for all $i \geq 1$,

$$\sup_{\theta \in \Theta^*} \Big(|c_i(\theta)| + |c_{i,l}(\theta)| + |c_{i,l,m}(\theta)| \Big) \leq K \rho^i \ , \tag{37}$$

where K is a positive constant.

From (5), this implies that there exists some other absolutely summable and deterministic sequences $(d_i(\theta))_{i\geq 0}$, $(d_{i,l}(\theta))_{i\geq 1}$ and $(d_{i,l,m}(\theta))_{i\geq 1}$ such that, almost surely,

$$\epsilon_t(\theta) = \sum_{i=0}^{\infty} d_i(\theta)\epsilon_{t-i} \ , \ \frac{\partial \epsilon_t(\theta)}{\partial \theta_l} = \sum_{i=1}^{\infty} d_{i,l}(\theta)\epsilon_{t-i} \text{ and } \frac{\partial^2 \epsilon_t(\theta)}{\partial \theta_l \partial \theta_m} = \sum_{i=2}^{\infty} d_{i,l,m}(\theta)\epsilon_{t-i} \tag{38}$$

$$e_t(\theta) = \sum_{i=0}^{t-1} d_i(\theta)e_{t-i} \ , \ \frac{\partial e_t(\theta)}{\partial \theta_l} = \sum_{i=1}^{t-1} d_{i,l}(\theta)e_{t-i} \text{ and } \frac{\partial^2 e_t(\theta)}{\partial \theta_l \partial \theta_m} = \sum_{i=2}^{t-1} d_{i,l,m}(\theta)e_{t-i}. \tag{39}$$

Again, we have $d_0(\theta) = 1$ and the three above sequences also satisfy

$$\sup_{\theta\in\Theta^*} \Big(|d_i(\theta)| + |d_{i,l}(\theta)| + |d_{i,l,m}(\theta)| \Big) \leq K\,\rho^i \ . \tag{40}$$

Finally, from the above estimates, we are able to deduce that for any $(l,m) \in \{1,...,p+q\}^2$, it holds

$$\sup_{\theta\in\Theta^*} |\epsilon_t(\theta) - e_t(\theta)| \xrightarrow[t\to\infty]{\text{a.s.}} 0 \ , \tag{41}$$

$$\rho^t \sup_{\theta\in\Theta^*} |\epsilon_t(\theta)| \xrightarrow[t\to\infty]{\text{a.s.}} 0 \ . \tag{42}$$

Analogous estimates to (41) and (42) are satisfied for first and second order derivatives of ϵ_t and e_t.

This implies that the sequences $\sqrt{n}\frac{\partial}{\partial\theta}Q_n(\theta_0)$ and $\sqrt{n}\frac{\partial}{\partial\theta}O_n(\theta_0)$ have the same asymptotic distribution. More precisely, we have

$$\sqrt{n}\left(\frac{\partial}{\partial\theta}Q_n(\theta_0) - \frac{\partial}{\partial\theta}O_n(\theta_0)\right) = \mathrm{o}_{\mathbb{P}}(1) \ . \tag{43}$$

6.2. *Proof of Theorem 1*

The proof is divided in several steps.

6.2.1. *Taylor's expansion of $\hat{\gamma}_m$*

The aim of this step is to prove that

$$\sqrt{n}\ \hat{\gamma}_m = \frac{1}{\sqrt{n}}\ \Gamma\ U_t + \mathrm{o}_{\mathbb{P}}(1). \tag{44}$$

Let $h \in \{1,...,m\}$. We apply the mean value theorem to the function

$$\theta \mapsto (1/n) \sum_{t=h+1}^{n} \epsilon_t(\theta)\,\epsilon_{t-h}(\theta)$$

between the points θ_0 and $\hat{\theta}_n$. Thus there exists θ_n^* between θ_0 and $\hat{\theta}_n$ such that

$$\frac{1}{n}\sum_{t=h+1}^{n} \epsilon_t(\hat{\theta}_n)\,\epsilon_{t-h}(\hat{\theta}_n) = \gamma(h) + \frac{1}{n}\sum_{t=h+1}^{n} D_t(\theta_n^*)\ (\hat{\theta}_n - \theta_0)$$

where the function D_t is defined on Θ by

$$D_t(\theta) = \frac{\partial \epsilon_t(\theta)}{\partial \theta} \epsilon_{t-h}(\theta) + \epsilon_t(\theta) \frac{\partial \epsilon_{t-h}(\theta)}{\partial \theta'}.$$

Therefore,

$$\begin{aligned} \hat{\gamma}(h) &= \frac{1}{n} \sum_{t=h+1}^{n} e_t(\hat{\theta}_n) \, e_{t-h}(\hat{\theta}_n) \\ &= \gamma(h) + \mathbb{E}(D_t(\theta_0)) \, (\hat{\theta}_n - \theta_0) + R_n \ , \end{aligned} \tag{45}$$

where $R_n = R_n^1 + R_n^2 + R_n^3$ with

$$\begin{aligned} R_n^1 &= \frac{1}{n} \sum_{t=h+1}^{n} \Big\{ e_t(\hat{\theta}_n) \, e_{t-h}(\hat{\theta}_n) - \epsilon_t(\hat{\theta}_n) \, \epsilon_{t-h}(\hat{\theta}_n) \Big\} \\ R_n^2 &= \left(\frac{1}{n} \sum_{t=h+1}^{n} \left(D_t(\theta_n^*) - D_t(\theta_0) \right) \right) (\hat{\theta}_n - \theta_0) \\ R_n^3 &= \left(\frac{1}{n} \sum_{t=h+1}^{n} D_t(\theta_0) - \mathbb{E}(D_t(\theta_0)) \right) \ (\hat{\theta}_n - \theta_0) \ . \end{aligned}$$

We recall that $\lim_{n\to\infty} \theta_n = \theta_0$. Then, using (35) to (42), combined with the arguments of Lemma A3 in [10], one has $\lim_{n\to\infty} R_n = 0$ almost-surely. By (35), one may find an absolutely summable and deterministic sequence $(\tilde{d}_i(\theta))_{i\geq 1}$ such that $\partial \epsilon_{t-h}(\theta_0)/\partial \theta' = \sum_{i=1}^{\infty} \tilde{d}_i(\theta) \epsilon_{t-h-i}(\theta)$. Hence $\partial \epsilon_{t-h}(\theta_0)/\partial \theta'$ is not correlated with ϵ_t when $h \geq 0$. Consequently,

$$\mathbb{E}(D_t(\theta_0)) = \mathbb{E}\left(\frac{\partial \epsilon_t(\theta_0)}{\partial \theta'} \epsilon_{t-h}(\theta_0) \right) := \phi^h \in \mathbb{R}^{p+q}$$

and (45) becomes

$$\hat{\gamma}(h) = \gamma(h) + \phi^h \, (\hat{\theta}_n - \theta_0) + \mathrm{o}_{\mathbb{P}}(1). \tag{46}$$

Now we use the Taylor expansion of the derivative of Q_n which is defined in (7). This classical argument uses the fact that $\partial Q_n(\hat{\theta}_n)/\partial \theta = 0$, because $\hat{\theta}_n$ minimizes the function $\theta \mapsto Q_n(\theta)$ (see the proof of Theorem 2 in [9]). Therefore there exists $\theta^\sharp$ is between θ_0 and $\hat{\theta}_n$ such that

$$\begin{aligned} 0 &= \sqrt{n} \, \frac{\partial Q_n(\hat{\theta}_n)}{\partial \theta} \\ &= \sqrt{n} \, \frac{\partial Q_n(\theta_0)}{\partial \theta} + \frac{\partial^2 Q_n(\theta^\sharp)}{\partial \theta \partial \theta'} \sqrt{n} \left(\hat{\theta}_n - \theta_0 \right) \\ &= \sqrt{n} \, \frac{\partial O_n(\theta_0)}{\partial \theta} + \frac{\partial^2 O_n(\theta_0)}{\partial \theta \partial \theta'} \sqrt{n} \left(\hat{\theta}_n - \theta_0 \right) + R_n^4 \end{aligned} \tag{47}$$

with

$$R_n^4 = \sqrt{n} \left(\frac{\partial Q_n(\theta_0)}{\partial \theta} - \frac{\partial O_n(\theta_0)}{\partial \theta} \right) + \left(\frac{\partial^2 Q_n(\theta^\sharp)}{\partial \theta \partial \theta'} - \frac{\partial^2 O_n(\theta_0)}{\partial \theta \partial \theta'} \right) \sqrt{n} \left(\hat{\theta}_n - \theta_0 \right) \ .$$

The first therm of R_n^4 converges in probability to 0 by (43). The last term of R_n^4 tends to 0 almost-surely by the same arguments used in the proof of Theorem 2 in [9]. Consequently, $R_n^4 = \mathrm{o}_{\mathbb{P}}(1)$.

By the definition of the matrix J (see (12)), we get from (47) that

$$\hat{\theta}_n - \theta_0 = -J^{-1}\frac{\partial O_n(\theta_0)}{\partial\theta} + \mathrm{o}_{\mathbb{P}}(1) = -J^{-1}\frac{1}{n}\sum_{t=1}^{n} Y_t + \mathrm{o}_{\mathbb{P}}(1)$$

where

$$Y_t = -\epsilon_t(\theta_0)\frac{\partial \epsilon_t(\theta_0)}{\partial\theta}. \tag{48}$$

Thus we may rewrite (46) as

$$\begin{aligned}\hat{\gamma}(h) &= \gamma(h) - \frac{1}{n}\phi^h\, J^{-1}\sum_{t=1}^{n} Y_t + \mathrm{o}_{\mathbb{P}}(1)\\ &= \frac{1}{n}\sum_{t=h+1}^{n}\epsilon_t\,\epsilon_{t-h} - (\phi^h\, J^{-1})\frac{1}{n}\sum_{t=1}^{n} Y_t + \mathrm{o}_{\mathbb{P}}(1).\end{aligned} \tag{49}$$

Now, we come back to the vector $\hat{\gamma}_m = (\hat{\gamma}(1), ..., \hat{\gamma}(m))'$. We remark that the matrix Φ_m defined by (17) is the matrix in $\mathbb{R}^{m\times(p+q)}$ whose line h is ϕ^h. By (48) and (49), we obtain

$$\sqrt{n}\ \hat{\gamma}_m = \frac{1}{\sqrt{n}}\,\Gamma\left(\sum_{t=1}^{n} Y_t, \sum_{t=1+1}^{n}\epsilon_t\epsilon_{t-1}, ..., \sum_{t=m+1}^{n}\epsilon_t\epsilon_{t-m}\right)' + \mathrm{o}_{\mathbb{P}}(1). \tag{50}$$

Therefore, the above Taylor's expansion (44) of $\hat{\gamma}_m$ is proved. This ends our first step.

Now, it is clear that the asymptotic behaviour of $\hat{\gamma}_m$ is related to the limit distribution of $U_t = (Y_t, \epsilon_t\epsilon_{t-1}, ..., \epsilon_t\epsilon_{t-m})'$. The next step deals with the asymptotic distribution of ΓU_t.

6.2.2. Functional central limit theorem for $(\Gamma U_t)_{t\in\mathbb{Z}}$

Our purpose is to prove that there exists a lower triangular matrix Ψ, with nonnegative diagonal entries, such that

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor nr\rfloor}\Gamma U_j \xrightarrow[n\to\infty]{\mathbb{D}^m} \Psi B_m(r) \tag{51}$$

where $(B_m(r))_{r\geq 0}$ is a m-dimensional standard Brownian motion.

By (38), one rewrites U_t as

$$U_t = \left(\sum_{t=1}^{n}\sum_{i=1}^{\infty} d_i(\theta)\epsilon_t\epsilon_{t-i}, \sum_{t=1+1}^{n}\epsilon_t\epsilon_{t-1}, ..., \sum_{t=m+1}^{n}\epsilon_t\epsilon_{t-m}\right)'$$

and thus it has zero expectation with values in $\mathbb{R}^{k_0+m}$. In order to apply the function central limit theorem for strongly mixing process, one has to introduce, for any integer k, the random variables

$$U_t^k = \left(\sum_{t=1}^{n}\sum_{i=1}^{k} d_i(\theta)\epsilon_t\epsilon_{t-i}, \sum_{t=1+1}^{n}\epsilon_t\epsilon_{t-1}, ..., \sum_{t=m+1}^{n}\epsilon_t\epsilon_{t-m}\right)'. \tag{52}$$

Since U^k depends on a finite number of values of the noise-process ϵ, it satisfies a mixing property of the form (9). Using this truncation procedure, one may prove as in [9] that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_t \xrightarrow[n\to\infty]{\mathrm{d}} \mathcal{N}(0,\Xi)$$

where

$$\Xi := 2\pi f_U(0) = \sum_{h=-\infty}^{+\infty} \mathrm{Cov}(U_t, U_{t-h}) = \sum_{h=-\infty}^{+\infty} \mathbb{E}(U_t U'_{t-h}), \tag{53}$$

where $f_U(0)$ is the spectral density of the stationary process $(U_t)_{t\in\mathbb{Z}}$ evaluated at frequency 0 (see for example [5]). The main issue is the existence of the sum of the right-hand side of (53). It is ensured by the Davydov inequality (see [6]) and the arguments developed in the Lemma A.1 in [8] (se also [9]). Since the matrix Ξ is positive definite, it can be factorized as $\Xi = \Upsilon\Upsilon'$ where the $(k_0+m)\times(k_0+m)$ lower triangular matrix Υ has nonnegative diagonal entries. Therefore, we have

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} \Gamma U_t \xrightarrow[n\to\infty]{\mathrm{d}} \mathcal{N}(0,\Gamma\Xi\Gamma'),$$

and the new variance matrix can also been factorized as $\Gamma\Xi\Gamma' = (\Gamma\Upsilon)(\Gamma\Upsilon)' := \Psi\Psi'$. Thus, $n^{-1/2}\sum_{t=1}^{n}\Psi^{-1}\Gamma U_t \xrightarrow[n\to\infty]{\mathrm{d}} \mathcal{N}(0, I_m)$ where I_m is the identity matrix of order m. The above arguments also apply to the sequence U^k. There exists a sequence of matrix $(\Xi_k)_{k\geq 1}$, such that $\lim_{k\to\infty}\Xi_k = \Xi$. Some matrix Ψ_k are defined analogously as Ψ. Consequently,

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} \Gamma U_t^k \xrightarrow[n\to\infty]{\mathrm{d}} \mathcal{N}(0,\Gamma\Xi_k\Gamma').$$

Using [12], the functional central limit theorem also holds: for any $r \in (0,1)$,

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor nr\rfloor} \Psi_k^{-1}\Gamma U_j^k \xrightarrow[n\to\infty]{\mathbb{D}^m} B_m(r).$$

We write

$$\Psi^{-1}\Gamma U_j^k = \left(\Psi^{-1} - \Psi_k^{-1}\right)\Gamma U_j^k + \Gamma\Psi_k^{-1}\Gamma U_j^k$$

and we obtain that

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor nr\rfloor} \Psi^{-1}\Gamma U_j^k \xrightarrow[n\to\infty]{\mathbb{D}^m} B_m(r). \tag{54}$$

In order to conclude that (51) is true, it remains to observe that, uniformly with respect to n,

$$Z_n^k(r) := \frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor nr\rfloor} \Psi^{-1}\Gamma V_j^k \xrightarrow[k\to\infty]{\mathbb{D}^m} 0, \tag{55}$$

where

$$V_t^k = \left(\sum_{t=1}^{n}\sum_{i=k+1}^{\infty} d_i(\theta)\epsilon_t\epsilon_{t-i}, \sum_{t=1+1}^{n}\epsilon_t\epsilon_{t-1}, \ldots, \sum_{t=m+1}^{n}\epsilon_t\epsilon_{t-m}\right)'.$$

By Lemma 4 in [9],

$$\sup_n \mathrm{VAR}\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n} V_j^k\right) \xrightarrow[k\to\infty]{} 0$$

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

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

Thus (55) is true. We additionally use (54) to conclude that (51) is true.

6.2.3. Limit theorem

We prove Theorem 1. We follow the arguments developed in the Sections 2 and 3 in [14]. The main difference is that we shall work with the sequence $(\Gamma U_t)_{t\geq 1}$ instead of the sequence $\left((\epsilon_t\epsilon_{t-1},...,\epsilon_t\epsilon_{t-m})'\right)_{t\geq 1}$. The previous step ensures us that Assumption 1 in [14] is satisfied for the sequence $(\Gamma U_t)_{t\geq 1}$. Since $C_m = (1/n^2)\sum_{t=1}^n S_t S_t'$, the continuous mapping theorem on the Skorokhod space implies that

$$C_m \xrightarrow[n\to\infty]{\mathrm{d}} \Psi V_m \Psi', \tag{56}$$

where the random variable V_m is defined in (20). Since by (50), $\sqrt{n}\hat{\gamma}_n = n^{-1/2}\sum_{t=1}^n \Gamma U_t + \mathrm{o}_{\mathbb{P}}(1)$, we use (51) and (56) in order to obtain

$$\begin{aligned} n\hat{\gamma}_m' C_m^{-1}\hat{\gamma}_m &= \frac{1}{n}\sum_{t=1}^{n}\left((\Gamma U_t)' C_m^{-1}(\Gamma U_t)\right) \\ &\xrightarrow[n\to\infty]{\mathrm{d}} (\Psi B_m(1))'\left(\Psi V_m \Psi'\right)^{-1}(\Psi B_m(1)) = B_m'(1) V_m^{-1} B_m(1), \end{aligned}$$

and we recognize the random variable $\mathcal{U}_m$ defined in (19). Consequently we have proved (22). The property (24) is straightforward since $\hat{\rho}(h) = \hat{\gamma}(h)/\sigma^2$ for $h = 1,...,m$. The proof of Theorem 1 is then complete.

6.3. Proof of Theorem 2

We write $\hat{C}_m = C_m + \Upsilon_n$ where

$$\Upsilon_n = \frac{1}{n^2}\sum_{t=1}^{n}(S_t S_t' - \hat{S}_t\hat{S}_t').$$

There are three kinds of entries in the matrix Υ_n. The first one is a sum composed of

$$\upsilon_t^{k,k'} = \epsilon_t^2(\theta_0)\epsilon_{t-k}(\theta_0)\epsilon_{t-k'}(\theta_0) - e_t^2(\hat{\theta}_n)e_{t-k}(\hat{\theta}_n)e_{t-k'}(\hat{\theta}_n)$$

for $(k,k') \in \{1,...,m\}^2$. By (40) and the consistency of $\hat{\theta}_n$, we have $\upsilon_t^{k,k'} = \mathrm{o}(1)$ almost-surely. The two last kinds of entries of Υ_n come from the following quantities

$$\tilde{\upsilon}_t^{k,k'} = \epsilon_t^2(\theta_0)\epsilon_{t-k'}(\theta_0)\frac{\partial \epsilon_{t-k}(\theta_0)}{\partial\theta} - e_t^2(\hat{\theta}_n)e_{t-k'}(\hat{\theta}_n)\left.\frac{\partial e_{t-k}(\theta)}{\partial\theta}\right|_{\theta=\hat{\theta}_n}$$

$$\bar{\upsilon}_t^{k,k'} = \epsilon_{t-k}(\theta_0)\epsilon_{t-k'}(\theta_0)\frac{\partial \epsilon_{t-k}(\theta_0)}{\partial\theta}\frac{\partial \epsilon_{t-k'}(\theta_0)}{\partial\theta} - e_{t-k}(\hat{\theta}_n)e_{t-k'}(\hat{\theta}_n)\left.\frac{\partial e_{t-k}(\theta)}{\partial\theta}\right|_{\theta=\hat{\theta}_n}\left.\frac{\partial e_{t-k'}(\theta)}{\partial\theta}\right|_{\theta=\hat{\theta}_n}$$

and they also satisfy $\upsilon_t^{k,k'} + \bar{\upsilon}_t^{k,k'} = \mathrm{o}(1)$ almost-surely. Consequently, $\Upsilon_n = \mathrm{o}(1)$ almost-surely as n goes to infinity. Thus one may find a matrix Υ_n^*, that tends to the null matrix almost-surely, such that

$$\begin{aligned} n\,\hat{\gamma}_m' \hat{C}_m^{-1} \hat{\gamma}_m &= n\,\hat{\gamma}_m' (C_m + \Upsilon_n)^{-1} \hat{\gamma}_m \\ &= n\,\hat{\gamma}_m' C_m^{-1} \hat{\gamma}_m + n\,\hat{\gamma}_m' \Upsilon_n^* \hat{\gamma}_m \ . \end{aligned}$$

Thanks to the arguments developed in the proof of Theorem 1, $n\hat{\gamma}_m'\hat{\gamma}_m$ converges in distribution. So $n\hat{\gamma}_m'\Upsilon_n^*\hat{\gamma}_m$ tends to zero in distribution, hence in probability. Then $n\hat{\gamma}_m'\hat{C}_m^{-1}\hat{\gamma}_m$ and $n\hat{\gamma}_m' C_m^{-1}\hat{\gamma}_m$ have the same limit in distribution and the result is proved.

Acknowledgements

The research of Y. Boubacar Maïnassara was supported by a BQR (Bonus Qualité Recherche) of the Université de Franche-Comté.

References

[1] Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59**, 3, 817–858.

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

[3] Boubacar Mainassara, Y. (2011). Multivariate portmanteau test for structural VARMA models with uncorrelated but non-independent error terms. *J. Statist. Plann. Inference* **141**, 8, 2961–2975.

[4] 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.

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

[6] Davydov, J. A. (1968). The convergence of distributions which are generated by stationary random processes. *Teor. Verojatnost. i Primenen. 13*, 730–737.

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

[8] Francq, C., Roy, R., and Zakoïan, J.-M. (2005). Diagnostic checking in ARMA models with uncorrelated errors. *J. Amer. Statist. Assoc.* **100**, 470, 532–544.

[9] Francq, C. and Zakoïan, J.-M. (1998). Estimating linear representations of nonlinear processes. *J. Statist. Plann. Inference* **68**, 1, 145–165.

[10] Francq, C. and Zakoïan, J.-M. (2000). Covariance matrix estimation for estimators of mixing weak ARMA models. *J. Statist. Plann. Inference* **83**, 2, 369–394.

[11] 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., Vol. **1**. Springer, New York, 241–265.

[12] Herrndorf, N. (1984). A functional central limit theorem for weakly dependent sequences of random variables. *Ann. Probab.* **12**, 1, 141–153.

[13] Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika* **65**, 2.

[14] Lobato, I. N. (2001). Testing that a dependent process is uncorrelated. *J. Amer. Statist. Assoc.* **96**, 455, 1066–1076.
[15] Newey, W. K. and West, K. D. (1987). A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* **55**, 3, 703–708.
[16] Shao, X. (2012). Parametric inference in stationary time series models with dependent errors. *Scand. J. Stat.* **39**, 4, 772–783.

7. Tables

Table 1
Upper Critical Values of the Distribution of $\mathcal{U}_m$.

$m\backslash\alpha$	90	95	97.5	99	99.5
1	28.43	45.73	66.57	100.02	129.26
2	70.68	102.94	138.85	194.00	239.47
3	126.58	174.62	227.21	304.16	370.03
4	194.23	258.68	330.13	429.21	510.05
5	274.32	357.02	444.63	566.13	665.13
6	365.09	466.60	571.74	717.55	836.68
7	467.53	587.12	710.42	882.78	1018.78
8	581.20	721.75	869.05	1065.60	1223.84
9	703.28	865.77	1034.87	1260.32	1441.43
10	838.06	1023.06	1205.85	1462.63	1654.02
11	984.14	1190.11	1400.45	1674.11	1884.53
12	1137.44	1368.23	1596.98	1898.21	2126.25
13	1304.32	1551.80	1806.80	2139.87	2409.02
14	1477.48	1747.29	2017.97	2383.78	2654.10
15	1660.83	1965.79	2261.07	2638.53	2956.67
16	1855.49	2184.97	2498.27	2923.30	3229.06
17	2062.69	2413.09	2748.80	3205.62	3541.93
18	2283.94	2657.10	3028.39	3500.37	3856.29
19	2505.54	2907.39	3298.23	3814.26	4200.32
20	2734.79	3161.32	3572.52	4122.45	4541.10
21	2981.93	3432.20	3869.38	4461.20	4897.68
22	3232.17	3707.30	4176.87	4782.44	5243.65
23	3499.66	4011.67	4516.08	5157.49	5646.50
24	3772.57	4313.92	4835.70	5491.94	5990.39
36	7789.22	8710.34	9587.04	10701.47	11566.64
48	13126.85	14515.20	15754.06	17328.84	18466.33

The critical values for $\mathcal{U}_m$ have also been tabulated by Lobato in [14]. We remark that the critical values that we have computed are close to the one obtained by Lobato. Moreover, the difference between these two tables do not affect the results of the testing procedure (see the comments in Section 4).

TABLE 2

Empirical size (in %) of the standard and modified versions of the BP test in the case of the strong ARMA : model (32), with the parameter $\theta_0 = (0,0)'$ and $\alpha_1 = 0$. The nominal asymptotic level of the tests is $\alpha = 5\%$.

	$m=2$			$m=3$			$m=4$		
Length n	100	300	1,000	100	300	1,000	100	300	1,000
LB_{SN}	4.7	5.6	3.7	4.2	6.1	3.2	4.3	6.3	4.0
$\text{BP}_{\text{SN}}^{\text{Lo}}$	4.4	5.5	3.7	3.9	6.1	3.2	3.8	6.3	3.9
BP_{SN}	4.4	5.6	3.7	3.9	6.1	3.2	3.8	6.3	3.9
LB_{FRZ}	3.7	4.5	4.7	3.8	4.2	4.3	2.4	4.6	4.8
BP_{FRZ}	3.2	4.2	4.7	2.5	4.0	4.3	2.3	4.3	4.8
LB_{S}	4.5	4.5	4.7	3.3	5.7	4.4	3.8	5.7	5.0
BP_{S}	4.2	4.5	4.7	3.0	5.2	4.2	3.2	5.3	5.0
	$m=6$			$m=10$			$m=18$		
Length n	100	300	1,000	100	300	1,000	100	300	1,000
LB_{SN}	4.4	6.0	5.3	3.5	5.4	6.1	2.4	4.0	6.5
$\text{BP}_{\text{SN}}^{\text{Lo}}$	3.5	5.6	5.3	2.3	5.4	6.1	1.3	3.0	6.0
BP_{SN}	3.5	5.6	5.2	2.3	5.4	6.1	1.3	3.0	6.1
LB_{FRZ}	2.1	4.6	4.8	2.5	4.6	4.4	2.9	4.0	5.0
BP_{FRZ}	1.7	4.3	4.8	1.7	4.1	4.3	1.6	3.5	4.8
LB_{S}	4.4	6.0	4.8	4.9	6.2	4.9	6.1	6.2	5.0
BP_{S}	3.4	5.7	4.8	3.8	5.5	4.7	3.3	4.9	5.0

Table 3
Empirical size (in %) of the standard and modified versions of the BP test in the case of the strong ARMA$(1,1)$: model (32), with the parameter $\theta_0 = (0.95, -0.6)'$ and $\alpha_1 = 0$. The nominal asymptotic level of the tests is $\alpha = 5\%$.

	$m=1$		$m=2$		$m=3$	
Length n	500	1,000	500	1,000	500	1,000
$\mathrm{LB_{SN}}$	5.1	5.4	5.7	5.4	4.2	3.9
$\mathrm{BP_{SN}^{Lo}}$	5.1	5.3	5.6	5.4	4.2	3.9
$\mathrm{BP_{SN}}$	5.0	5.2	5.7	5.4	4.2	3.9
$\mathrm{LB_{FRZ}}$	5.6	4.9	5.5	4.9	3.5	3.3
$\mathrm{BP_{FRZ}}$	5.3	4.8	5.5	4.9	3.3	3.3
$\mathrm{LB_S}$	n.a.	n.a.	n.a.	n.a.	15.3	13.8
$\mathrm{BP_S}$	n.a.	n.a.	n.a.	n.a.	15.0	13.5
	$m=6$		$m=12$		$m=18$	
Length n	500	1,000	500	1,000	500	1,000
$\mathrm{LB_{SN}}$	7.2	5.8	6.4	6.0	5.4	7.1
$\mathrm{BP_{SN}^{Lo}}$	7.0	5.7	6.2	5.9	4.7	6.4
$\mathrm{BP_{SN}}$	7.0	5.7	6.2	5.8	4.7	6.5
$\mathrm{LB_{FRZ}}$	2.6	1.8	2.2	1.5	3.3	1.8
$\mathrm{BP_{FRZ}}$	2.2	1.7	1.9	1.3	2.4	1.7
$\mathrm{LB_S}$	8.3	7.6	7.2	5.9	7.6	7.7
$\mathrm{BP_S}$	7.8	7.5	6.9	5.6	6.4	6.5

Table 4
Empirical size (in %) of the standard and modified versions of the BP test in the case of the weak ARMA$(0,0)$: model (32), with the parameter $\theta_0 = (0,0)'$ and $\alpha_1 = 0.4$. The nominal asymptotic level of the tests is $\alpha = 5\%$.

	$m=2$		$m=3$		$m=4$	
Length n	300	1,000	300	1,000	300	1,000
$\mathrm{LB_{SN}}$	5.1	3.8	3.4	4.5	4.2	4.4
$\mathrm{BP_{SN}^{Lo}}$	4.9	3.8	3.4	4.5	4.0	4.3
$\mathrm{BP_{SN}}$	5.1	3.9	3.4	4.5	4.0	4.4
$\mathrm{LB_{FRZ}}$	3.2	4.0	3.0	3.9	3.1	4.0
$\mathrm{BP_{FRZ}}$	3.1	4.0	2.5	3.9	3.1	4.0
$\mathrm{LB_S}$	19.1	19.4	18.3	20.1	16.1	18.9
$\mathrm{BP_S}$	18.8	19.4	17.5	20.0	15.6	18.7
	$m=6$		$m=12$		$m=18$	
Length n	300	1,000	300	1,000	300	1,000
$\mathrm{LB_{SN}}$	4.8	4.6	1.9	4.7	1.3	4.9
$\mathrm{BP_{SN}^{Lo}}$	4.4	4.4	1.9	4.7	1.1	4.7
$\mathrm{BP_{SN}}$	4.3	4.4	1.9	4.7	1.1	4.7
$\mathrm{LB_{FRZ}}$	2.7	3.8	2.6	2.1	1.9	3.4
$\mathrm{BP_{FRZ}}$	2.3	3.8	2.3	2.0	1.4	3.2
$\mathrm{LB_S}$	14.3	16.2	11.3	13.7	11.7	13.1
$\mathrm{BP_S}$	13.8	16.0	10.8	13.4	9.7	12.3

TABLE 5
Empirical size (in %) of the standard and modified versions of the BP test in the case of the weak ARMA(1,1) : model (32), with the parameter $\theta_0 = (0.95, -0.6)'$ and $\alpha_1 = 0.4$. The nominal asymptotic level of the tests is $\alpha = 5\%$.

	$m = 1$		$m = 2$		$m = 3$	
Length n	1,000	2,000	1,000	2,000	1,000	2,000
$\mathrm{LB}_{\mathrm{SN}}$	4.6	4.8	5.0	4.8	3.8	2.9
$\mathrm{BP}_{\mathrm{SN}}^{\mathrm{Lo}}$	4.5	4.8	5.0	4.8	3.8	2.9
$\mathrm{BP}_{\mathrm{SN}}$	4.5	4.7	5.0	4.8	3.8	2.9
$\mathrm{LB}_{\mathrm{FRZ}}$	4.3	4.8	3.9	4.7	2.1	2.7
$\mathrm{BP}_{\mathrm{FRZ}}$	4.3	4.8	3.8	4.7	2.1	2.7
LB_{S}	n.a.	n.a.	n.a.	n.a.	25.4	26.8
BP_{S}	n.a.	n.a.	n.a.	n.a.	25.4	26.7
	$m = 6$		$m = 12$		$m = 18$	
Length n	1,000	2,000	1,000	2,000	1,000	2,000
$\mathrm{LB}_{\mathrm{SN}}$	6.2	5.7	5.7	4.1	3.6	4.7
$\mathrm{BP}_{\mathrm{SN}}^{\mathrm{Lo}}$	6.2	5.6	5.3	3.9	3.5	4.4
$\mathrm{BP}_{\mathrm{SN}}$	6.2	5.6	5.3	3.8	3.5	4.6
$\mathrm{LB}_{\mathrm{FRZ}}$	1.2	1.0	0.7	0.6	0.8	0.6
$\mathrm{BP}_{\mathrm{FRZ}}$	1.2	1.0	0.7	0.6	0.6	0.6
LB_{S}	13.4	15.2	9.6	10.7	7.7	9.6
BP_{S}	13.3	15.0	9.0	10.6	7.2	8.9

TABLE 6
Empirical power (in %) of the standard and modified versions of the BP test in the case of the weak ARMA(2,1) model (33) and $\alpha_1 = 0$.

	$m = 1$		$m = 2$		$m = 3$	
Length n	500	1,000	500	1,000	500	1,000
$\mathrm{LB}_{\mathrm{SN}}$	77.4	95.1	65.8	88.4	46.2	73.9
$\mathrm{BP}_{\mathrm{SN}}^{\mathrm{Lo}}$	77.2	95.1	65.8	88.1	46.1	73.8
$\mathrm{BP}_{\mathrm{SN}}$	77.0	95.0	65.8	88.5	46.2	73.8
$\mathrm{LB}_{\mathrm{FRZ}}$	95.3	99.6	92.4	99.0	66.5	76.7
$\mathrm{BP}_{\mathrm{FRZ}}$	95.3	99.6	92.2	99.0	65.9	76.4
LB_{S}	n.a.	n.a.	n.a.	n.a.	95.3	99.9
BP_{S}	n.a.	n.a.	n.a.	n.a.	95.3	99.9
	$m = 4$		$m = 6$		$m = 10$	
Length n	500	1,000	500	1,000	500	1,000
$\mathrm{LB}_{\mathrm{SN}}$	53.9	82.0	48.5	77.4	40.2	73.5
$\mathrm{BP}_{\mathrm{SN}}^{\mathrm{Lo}}$	53.4	81.9	48.3	77.3	39.3	73.4
$\mathrm{BP}_{\mathrm{SN}}$	53.5	81.8	48.3	77.3	38.9	73.5
$\mathrm{LB}_{\mathrm{FRZ}}$	72.2	88.1	68.9	91.4	57.4	87.5
$\mathrm{BP}_{\mathrm{FRZ}}$	71.6	88.0	68.0	91.3	55.8	87.1
LB_{S}	90.3	99.7	88.3	99.4	82.3	98.6
BP_{S}	90.1	99.7	87.6	99.4	81.6	98.6

Table 7
Empirical power (in %) of the standard and modified versions of the BP test in the case of the weak ARMA$(2,1)$ model (33) and $\alpha_1 = 0.4$.

	$m=1$		$m=2$		$m=3$	
Length n	500	1,000	500	1,000	500	1,000
LB_{SN}	60.7	84.6	49.1	72.9	30.3	56.1
$\text{BP}_{\text{SN}}^{\text{Lo}}$	60.6	84.6	48.7	72.7	29.9	56.1
BP_{SN}	60.5	84.4	49.1	72.9	29.9	56.1
LB_{FRZ}	81.7	97.6	75.1	95.1	29.6	59.4
BP_{FRZ}	81.8	97.5	74.8	94.9	29.4	59.3
LB_{S}	n.a.	n.a.	n.a.	n.a.	91.1	99.5
BP_{S}	n.a.	n.a.	n.a.	n.a.	91.0	99.5
	$m=4$		$m=6$		$m=10$	
Length n	500	1,000	500	1,000	500	1,000
LB_{SN}	35.5	63.7	27.6	54.6	23.2	50.0
$\text{BP}_{\text{SN}}^{\text{Lo}}$	35.2	63.7	27.8	54.5	22.8	49.3
BP_{SN}	35.4	63.7	27.5	54.5	22.4	49.1
LB_{FRZ}	43.2	77.5	45.9	81.9	36.0	75.2
BP_{FRZ}	42.9	77.4	45.1	81.9	35.3	74.8
LB_{S}	84.7	98.5	81.8	97.5	75.0	95.8
BP_{S}	84.2	98.4	81.5	97.5	74.2	95.5

Table 8
Standard and modified versions of portmanteau tests to check the null hypothesis that the CAC40 returns is a white noise.

Lag m	2	3	4	5	10	18	24
$\hat{\rho}(m)$	-0.02829	-0.05308	0.04064	-0.05296	0.00860	-0.02182	0.00466
LB_{SN}	36.8397	65.3110	141.899	183.391	435.224	669.439	880.159
BP_{SN}	36.8010	65.2472	141.727	183.144	434.646	668.402	878.556
LB_{FRZ}	4.89097	19.4172	27.9413	42.4158	52.7845	61.2431	67.2210
BP_{FRZ}	4.88733	19.3994	27.9137	42.3685	52.7171	61.1467	67.0991
$\text{p}_{\text{SN}}^{\text{LB}}$	0.24154	0.27548	0.18218	0.22384	0.43726	0.90622	0.98502
$\text{p}_{\text{SN}}^{\text{BP}}$	0.24178	0.27574	0.18250	0.22440	0.43814	0.90674	0.98519
$\text{p}_{\text{FRZ}}^{\text{LB}}$	0.29699	0.03480	0.03837	0.00911	0.02085	0.17452	0.24341
$\text{p}_{\text{FRZ}}^{\text{BP}}$	0.29725	0.03491	0.03810	0.00916	0.02100	0.17544	0.24482
$\text{p}_{\text{S}}^{\text{LB}}$	0.08668	0.00022	0.00000	0.00000	0.00000	0.00000	0.00000
$\text{p}_{\text{S}}^{\text{BP}}$	0.08684	0.00023	0.00000	0.00000	0.00000	0.00000	0.00000

Table 9

Standard and modified versions of portmanteau tests to check the null hypothesis that the CAC40 squared returns follow an ARMA(1, 1) model.

Lag m	1	2	3	4	5	6	7
$\hat{\rho}(m)$	-0.04724	0.010467	0.01275	-0.00252	0.08579	-0.03532	-0.02772
LB_{SN}	8.96411	17.2907	21.0192	20.9689	21.0344	21.8014	25.0933
BP_{SN}	8.95890	17.2786	21.0026	20.9526	21.0172	21.7755	25.0477
LB_{FRZ}	11.5095	12.0746	12.9140	12.9468	50.9362	57.3777	61.3462
BP_{FRZ}	11.5028	12.0674	12.9061	12.9388	50.8767	57.3081	61.2697
p_{SN}^{LB}	0.30050	0.45977	0.66164	0.84375	0.93811	0.97700	0.98987
p_{SN}^{BP}	0.30061	0.45998	0.66183	0.84391	0.93822	0.97709	0.98991
p_{FRZ}^{LB}	0.11777	0.18293	0.34192	0.48658	0.29325	0.36110	0.38848
p_{FRZ}^{BP}	0.11789	0.18310	0.34218	0.48687	0.29368	0.36159	0.38899
p_{S}^{LB}	n.a.	n.a.	0.00033	0.00154	0.00000	0.00000	0.00000
p_{S}^{BP}	n.a.	n.a.	0.00033	0.00155	0.00000	0.00000	0.00000
Lag m	8	9	10	12	18	20	24
$\hat{\rho}(m)$	-0.04099	0.02048	0.04568	0.03584	0.03929	-0.04526	-0.04526
LB_{SN}	27.7828	27.5084	67.8011	93.7494	167.937	155.600	364.860
BP_{SN}	27.7576	27.4791	67.5936	93.1298	169.383	154.868	363.935
LB_{FRZ}	70.0230	72.1898	82.9700	89.6424	117.933	128.546	143.244
BP_{FRZ}	69.9297	72.0919	82.8470	89.5013	117.697	128.265	142.890
p_{SN}^{LB}	0.99599	0.99896	0.99275	0.99673	0.99988	0.99999	0.99995
p_{SN}^{BP}	0.99600	0.99897	0.99278	0.99679	0.99987	0.99999	0.99995
p_{FRZ}^{LB}	0.33766	0.40190	0.33819	0.39101	0.48717	0.45510	0.45367
p_{FRZ}^{BP}	0.33821	0.40250	0.33892	0.39191	0.48829	0.45635	0.45503
p_{S}^{LB}	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
p^{BP}	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000