Robust tests for independence of two time series

Pierre DUCHESNE[*,†] Roch ROY[‡,§]

CRM-2751

August 2001

[*]Service d'enseignement des méthodes quantitatives de gestion, École des Hautes Études Commerciales de Montréal
[†]This work was supported by grants from the Natural Science and Engineering Research Council of Canada and the Institut de Finance Mathématique de Montréal IFM[2].
[‡]Département de mathématiques et statistique, Université de Montréal
[§]This work was supported by grants from the Natural Science and Engineering Research Council of Canada, the Fonds FCAR (Government of Québec) and the Network of Centres of Excellence on the Mathematics of Information Technology and Complex Systems.

Abstract

This paper aims at developing a robust and omnibus procedure for checking the independence of two time series. Li and Hui (1994) proposed a robustified version of Haugh's (1976) classic portmanteau statistic which is based on a fixed number of lagged residual cross-correlations. In order to obtain a consistent test for independence against an alternative of serial cross-correlation of an arbitrary form between the two series, Hong's (1996a) introduced a class of statistics that take into account all possible lags. The test statistic is a weighted sum of residual cross-correlations and the weighting is determined by a kernel function. With the truncated uniform kernel, we retrieve a normalized version of Haugh's statistic. However, several kernels lead to a greater power. Here, we introduce a robustified version of Hong's statistic. We suppose that for each series, the true ARMA model is estimated by a $n^{1/2}$-consistent robust method and the robust cross-correlation is so obtained. Under the null hypothesis of independence, we show that the robust statistic asymptotically follows a $N(0,1)$ distribution. Using a result of Li and Hui, we also propose a robust procedure for checking independence at individual lags and a descriptive causality analysis in the Granger's sense is discussed. The level and power of the robust version of Hong's statistic are studied by simulation in finite samples. Finally, the proposed robust procedures are applied to a set of financial data.

Key words and phrases: ARMA model; robust serial correlation; robust estimation; coherency; independence; causality.

Mathematics subject classification codes (2000): primary 62M10; secondary 62F35.

Résumé

L'objectif premier de cet article est de développer une approche robuste afin de vérifier l'indépendance de deux séries chronologiques. Li et Hui (1994) ont proposé une version robuste de la statistique portmanteau de Haugh (1976) qui est basée sur les corrélations croisées résiduelles à un nombre fixé de délais. Afin d'obtenir un test convergent pour l'hypothèse d'indépendance par rapport à une alternative de corrélation sérielle croisée de forme quelconque, Hong (1996a) a introduit une classe de statistiques prenant en compte tous les délais. La statistique de test est une somme pondérée de toutes les corrélations croisées résiduelles, la pondération étant déterminée par une fonction de noyau. Avec le noyau uniforme tronqué, nous retrouvons une version normalisée de la statistique de Haugh. Cependant, plusieurs noyaux permettent d'obtenir une plus grande puissance. Dans ce travail, nous introduisons une version robuste de la statistique de Hong. Nous supposons que pour chaque série, le vrai modèle ARMA est estimé par une méthode robuste qui est $n^{1/2}$-convergente et les corrélations croisées résiduelles sont ainsi obtenues en utilisant une version robuste des corrélations croisées sérielles. Sous l'hypothèse nulle d'indépendance, nous montrons que la statistique robustifiée de Hong suit asymptotiquement une loi $N(0,1)$. De plus, en utilisant un résultat de Li et Hui, nous proposons une procédure robuste pour vérifier l'indépendance à chaque délai et une analyse descriptive de causalité au sens de Granger y est aussi décrite. Le niveau et la puissance de la statistique robustifiée de Hong sont étudiés par simulation en échantillons finis. Finalement, l'application de la méthodologie robuste proposée est illustrée à l'aide d'un jeu de données financières.

1. INTRODUCTION

Several physical and economic phenomena can be described by time series; see among others Akaike and Kitagawa (1999) for physical applications and Judge, Hill, Griffiths, Lütkepohl and Lee (1985) for economic applications. When we have to deal with two time series, the question often arises of describing the interrelationships existing between them. In economics, for example, elucidation of causality relationships between time series may be very important in a prediction context. Before applying sophisticated methods for describing relationships between two time series, it is important to check whether they are independent (or serially uncorrelated in the non-Gaussian case) or not. If sufficiently powerful methods which are simple both to apply and to interpret were available for checking the independence of two time series, more sophisticated multivariate analyses as causality analysis might become redundant. Haugh's (1976) paper is at our knowledge the first attempt at developing a procedure based on residual cross-correlations for checking the independence of two stationary ARMA series. It is a two-stage procedure. First, it involves fitting univariate models to each of the series and then, cross-correlating the two resulting residual series. Let $r_{\hat{u}\hat{v}}(j)$ be the residual cross-correlation at lag j:

$$r_{\hat{u}\hat{v}}(j) = \frac{\sum_{t=j+1}^{n} \hat{u}_t \hat{v}_{t-j}}{(\sum_{t=1}^{n} \hat{u}_t^2 \sum_{t=1}^{n} \hat{v}_t^2)^{1/2}},$$

for $|j| \leq n-1$, where $\hat{u}_t$, $\hat{v}_t$, $t = 1, \ldots, n$, are the two residual series and n is the length of each time series. The asymptotic distribution of a fixed number of residual cross-correlations is given by Haugh (1976). He considered a portmanteau statistic which is based on the first M residual cross-correlations, where $M \leq n-1$ is a fixed integer. More precisely, he introduced the statistic

$$S_M = n \sum_{j=-M}^{M} r_{\hat{u}\hat{v}}^2(j). \tag{1}$$

Its asymptotic distribution is chi-square under the null hypothesis of independence, and the hypothesis is rejected for large values of the test statistic.

Haugh's procedure was extended in various directions. Koch and Yang (1986) introduced a modification of S_M that allows for a potential pattern in the residual cross-correlation function. El Himdi and Roy (1997) proposed a version of S_M for two stationary vector ARMA (VARMA) that was recently extended to partially non stationary (cointegrated) VARMA series by Pham, Roy and Cédras (2000). Hallin and Saidi (2001) have recently proposed a generalization of Koch and Yang procedure for VARMA series.

In the application of Haugh's procedure, a value of M must be chosen a priori. It cannot be too large in order that the asymptotic chi-square approximation be satisfactory. Since S_M does not take into account the cross-correlations at lag j for $j > M$, it will not detect serial correlation at high lags and leads therefore to an inconsistent test procedure.

The previous observation led Hong (1996a) to propose a generalization of Haugh's statistic that takes into account all possible lags. Its test statistic is a weighted sum of residual cross-correlations of the form

$$Q_n = \frac{n \sum_{j=1-n}^{n-1} k^2(j/m) r_{\hat{u}\hat{v}}^2(j) - M_n(k)}{[2V_n(k)]^{1/2}}, \tag{2}$$

where $M_n(k) = \sum_{j=1-n}^{n-1}(1-|j|/n)k^2(j/m)$ and $V_n(k) = \sum_{j=2-n}^{n-2}(1-|j|/n)(1-(|j|+1)/n)k^4(j/m)$. The weighting depends on a kernel function k. With the truncated uniform kernel, Q_n corresponds to a normalized version of Haugh's statistic. However several kernels lead to a higher power than the truncated uniform kernel. Indeed, he shows that the so-called Daniell kernel maximizes the asymptotic power of his test in a certain class of smooth kernels and under some local alteratives. Also, Hong (1996a) avoids the ARMA modeling by fitting autoregressions of appropriate high orders. Hong's techniques of proofs are different from those of Haugh, since he directly establishes the asymptotic distribution of Q_n, without making use of the asymptotic distribution of a fixed length vector of residual cross-correlations. Under the null hypothesis, the test statistic Q_n is asymptotically $N(0,1)$. The test is unilateral and rejects for the large values of Q_n.

In practice, outliers in time series can create serious problems. They can occur for various reasons, namely because of measurement errors or equipment failure, etc. Haugh and Hong tests for independence are based on estimation methods that are sensitive to outliers. Furthermore, the usual cross-correlation function can be considerable affected by outliers. An interesting review of robust methods in time series is presented in Martin and Yohai (1985); see also Hampel, Ronchetti, Rousseeuw and Stahel (1986), Rousseeuw and Leroy (1987). Robust estimation methods in ARMA models are studied in Bustos and Yohai (1986), and a robustified autocovariance function is introduced. Robust estimation of VARMA models is considered in Ben, Martinez and Yohai (1999). Li and Hui (1994) proposed a robustified cross-correlation function between two time series and employed it to develop a robust version of Haugh's (1976) and McLeod's (1979) tests for checking independence. They established the asymptotic normality of a fixed length vector of robust cross-correlations. Their robust portmanteau statistics also follow asymptotically a chi-square distribution. Another approach of the non-parametric type for checking the independence of two autoregressive time series based on autoregressive rank scores is studied by Hallin, Jurevcková, Picek and Zahaf (1999).

The main objective of this paper is to develop a robust version of Hong's statistic for checking the independence of two univariate stationary and invertible time series. The test statistic is based on the robust cross-correlation function introduced by Li and Hui (1994). We suppose that for each series, the true ARMA model is estimated by a $n^{1/2}$-consistent robust method and the residual robust cross-correlation function is so obtained. In the empirical part of the paper, we used the residual autocovariances estimators (RA estimators) introduced by Bustos and Yohai (1986). A robustified version of Q_n is obtained by replacing in (2) the usual residual cross-correlation by the robust ones. As in El Himdi and Roy (1997), we also describe a robust procedure for checking independence at individual lags. It relies on a result of Li and Hui (1994). As emphasized by Box, Jenkins and Reinsel (1994) and others, it is important not only to look at the value of a global statistic but also to examine the values of the cross-correlations at individual lags. It may well happened that the procedure based on individual lags lead us to reject the null hypothesis whilst the global test does not reject. Also, if in a first step we reject with the global test, the tests at individual lags will identify the lags where there is a significant correlation.

The organization of the paper is as follows. In Section 2, we introduce the notations, concepts and results employed thereafter. Invoking a general result in Li and Hui (1994), we describe in Section 3 a robust procedure for checking independence at individual lags and we present a robustified version of Haugh's (1976) statistic. A descriptive approach for causality analysis in the Granger's (1969) sense is also discussed. The main result of the paper is presented in Section 4. We establish under the null hypothesis the asymptotic normality of a robustified version of Hong's statistic. The proof relies on a central limit theorem for martingale difference. It is long and involved and is deferred to an Appendix. In Section 5, we describe the results of a small Monte Carlo experiment conducted in order to analyze the level and power of the new statistics and to compare them to their non robust versions. The procedures introduced in the paper are applied to a set of financial data in Section 6 and we end with some concluding remarks.

2. PRELIMINARIES

2.1 Assumptions on the process

Let $\{(X_t, Y_t), t \in \mathbb{Z}\}$ be a bivariate second-order stationary process. Without loss of generality, we can assume that $\{(X_t, Y_t)\}$ is centered at zero. We further suppose that $\{X_t\}$ and $\{Y_t\}$ are of the ARMA type, that is

$$\begin{aligned} \phi_1(B)X_t &= \theta_1(B)u_t, \\ \phi_2(B)Y_t &= \theta_2(B)v_t, \end{aligned}$$

where $\phi_h(B) = \sum_{i=0}^{p_h} \phi_{hi} B^i$, $\theta_h(B) = \sum_{i=0}^{q_h} \theta_{hi} B^i$, $h = 1, 2$, with $\theta_{10} = \theta_{20} = 1$ and B is the backward shift operator. We assume that the two processes are invertible. >From the stationarity and inversibility assumptions, it follows that the power series of the complex variable z given by $\{\phi_h(z)\}^{-1}\theta_h(z) = \psi_h(z) = \sum_{r=0}^{\infty} \psi_{h,r} z^r$ and $\theta_h(z)^{-1} = \pi_h(z) = \sum_{r=0}^{\infty} \pi_{h,r} z^r$ converge for $|z| \leq 1$ and that the coefficients $\{\psi_{h,r}\}$ and $\{\pi_{h,r}\}$ tends to zero exponentially.

The innovation processes $\{u_t\}$ and $\{v_t\}$ are strong white noise, that is the u_t's and the v_t's are two sequence of independent and identically distributed random variables with mean zero and variance σ_u^2 and σ_v^2, respectively. It is also supposed that the fourth-order moment of u_t and v_t exist and that the probability distributions of u_t and v_t are symmetric with respect to zero. The symmetry assumption is frequent in the robustness context, see for example Denby and Martin (1979), Bustos, Fraiman and Yohai (1984) and Bustos and Yohai (1986).

2.2 Outliers in time series

In time series, Fox (1972) introduced two types of outliers: the innovation outliers and the additive outliers. That terminology is frequently used in robustness studies, see for exemple Martin and Yohai (1985), Rousseeuw and Leroy (1987), Hampel, Ronchetti, Rousseeuw and Stahel (1986). Here, we briefly described these two types of outliers using a probability models as it is often done in distributional studies.

a) Innovation outliers

Suppose that $\{X_t\}$ is an ARMA(p,q) process, as described in Section 2.1 and that the innovation process $\{u_t\}$ has an heavy tail distribution F with is not far from the Gaussian distribution, as for example a contaminated normal:

$$F = (1-\epsilon)N(0,\sigma^2) + \epsilon N(0,\tau^2),$$

where $\epsilon > 0$ is small, and $\tau \geq \sigma$. In this situation, the innovations u_t are from a $N(0,\sigma^2)$ with probability $1-\epsilon$, or from a $N(0,\tau^2)$ with a larger variance with probability ϵ. The latter innovations are considered as outliers.

The important point with innovation outliers is that the ARMA(p, q) model is still the exact model for the observations. However, if an outlier occurs at t_0, then u_{t_0} will affect not only X_{t_0}, but many future observations X_{t_0+1}, X_{t_0+2}, and so on. After a while, the effect disappears. Bustos and Yohai (1986) give several results showing that innovation outliers do not affect too seriously the least square estimators of autoregressive and moving avevage parameters of an ARMA process.

b) Additive outliers

Suppose that the observations are obtained from the model

$$X_t = \tilde{X}_t + V_t,$$

where $\{\tilde{X}_t\}$ is ARMA, that is $\phi_1(B)\tilde{X}_t = \theta_1(B)u_t$, the innovations u_t are Gaussian with a common variance σ^2, the V_t are iid random variables, independent of $\tilde{X}_t$ whose distribution is

$$H = (1-\epsilon)\delta_0 + \epsilon G,$$

where δ_0 is the degenerated distribution at zero, and G is an arbitrary distribution.

In that situation, the ARMA process itself is observed without error with probability $1-\epsilon$ and there is a probability ϵ that the process ARMA plus an error of distribution G be observed. Least-square estimators and even M-estimators are quite sensible to additive outliers. Bustos and Yohai (1986) proposed alternative estimation methods with a good behavior in presence of such outliers.

Since the ARMA(p, q) process $\{X_t\}$: $\phi(B)X_t = \theta(B)u_t$ is stationary and invertible, we can write $X_t = \phi^{-1}(B)\theta(B)u_t$ and $u_t = \theta(B)^{-1}\phi(B)X_t$. To robustly estimate the ARMA parameters of the process $\{X_t\}$, Bustos and Yohai (1986) introduced the RA estimators (for *Residual Autocovariances*). These estimators are obtained by solving the equation system (3), which is similar to the one leading to the least squares estimators except that the usual sample innovation autocorrelation function $\gamma_u(\cdot)$ is replaced by a robustified version $\gamma_u(\cdot;\eta)$. Let $\phi^{-1}(z) = \sum_{j\geq 0} s_j z^j$ and $\theta^{-1}(z) = \sum_{j\geq 0} \pi_j z^j$. The RA estimators are obtained through the resolution of the following system:

$$\begin{aligned}
\sum_{h=0}^{n-j-p-1} s_j \gamma_u(h+j;\eta) &= 0,\ j=1,\ldots,p; \\
\sum_{h=0}^{n-j-p-1} \pi_j \gamma_u(h+j;\eta) &= 0,\ j=1,\ldots,q; \\
\sum_{t=p+1}^{n} \psi(u_t/\sigma_u) &= 0.
\end{aligned} \tag{3}$$

The robust autocorrelation $\gamma_u(i;\eta)$ at lag i is defined by

$$\gamma_u(i;\eta) = n^{-1} \sum_{t=p+1+i}^{n} \eta(u_t/\sigma_u, u_{t-i}/\sigma_u). \tag{4}$$

We defined similarly $\gamma_{\hat{u}}(i;\eta)$ computed with the residual series. The scale parameter σ_u is estimated simultaneously from the observations using for example

$$\hat{\sigma}_u = \text{med}(|\hat{u}_{p+1}|,\ldots,|\hat{u}_n|)/.6745. \tag{5}$$

Several choices for the function η are possible, as for example Mallows type or Hampel type functions:

$$\eta_M(u,v) = \psi(u)\psi(v), \quad \text{Mallows}, \tag{6}$$

$$\eta_H(u,v) = \psi(uv), \quad \text{Hampel}, \tag{7}$$

where the function ψ is continuous and odd. Examples for the function ψ is the Huber family

$$\psi_H(u;c) = \text{sign}(u)\min(|u|,c); \tag{8}$$

and the bisquare family proposed by Beaton and Tukey (1974) which is defined by

$$\psi_B(u;c) = \begin{cases} u(1-u^2/c^2)^2, & \text{if } 0 \le |u| \le c, \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

Note that the function $\eta(u,v) = uv$ leads to the usual autocorrelation function and the least-squares estimators are obtained by solving (3) with $\eta(u,v) = uv$ and $\psi(u) = u$.

Bustos and Yohai (1986) also introduced a class of truncated RA estimators (TRA) that behave well when additive outliers are present. An heuristic argument shows that these estimators are qualitatively robust when a moving average component is present in the ARMA model. The results presented in Section 3 are based on RA estimators although the main result of this paper described in Section 4 only required that the AR and MA estimators be $n^{1/2}$-consistent and robust. Other robust and consistent estimators could be used in conjunction with the robustified autocorrelation function $\gamma_u(i;\eta)$. For example, GM estimators as well as RA and TRA estimators are $n^{1/2}$-consistent and robust. RA estimators with Mallows type function are particularly convenient to compute since an iterative scheme is possible. For rigor and completeness, we state explicitly the consistency condition in Assumption A.

Assumption A: Let $\{X_t\}$ and $\{Y_t\}$ be ARMA processes as described in Section 2.1. We suppose that the estimators of the parameters satisfy the following conditions:

$$\begin{aligned}
\phi_{hj} - \hat{\phi}_{hj} &= O_p(n^{-1/2}), j=1,\ldots,p_h; h=1,2; \\
\theta_{hj} - \hat{\theta}_{hj} &= O_p(n^{-1/2}), j=1,\ldots,q_h; h=1,2; \\
\sigma_u - \hat{\sigma}_u &= O_p(n^{-1/2});\ \sigma_v - \hat{\sigma}_v = O_p(n^{-1/2}).
\end{aligned}$$

With two time series, the cross-correlation function is often used to appreciate the existing dependency between the two processes. However, as the autocorrelation function, it is sensible to outliers. Li and Hui (1994) proposed the following robust cross-correlation function:

$$\gamma_{uv}(j;\eta) = \begin{cases} n^{-1}\sum_{t=j+1}^{n}\eta(u_t/\sigma_u, v_{t-j}/\sigma_v), \ j \geq 0, \\ n^{-1}\sum_{t=-j+1}^{n}\eta(u_{t+j}/\sigma_u, v_t/\sigma_v), \ j < 0, \end{cases} \tag{10}$$

where $\{u_t\}$ and $\{v_t\}$ are the two innovation processes, with variances σ_u^2 and σ_v^2 respectively. When $\eta(u,v) = uv$, we retrieve the usual sample cross-correlation function, and we write $\gamma_{uv}(j;\eta) = r_{uv}(j)$. Similarly, a robust residual cross-correlation function is obtained by replacing $\{u_t\}$ and $\{v_t\}$ by the residual series $\{\hat{u}_t\}$ and $\{\hat{v}_t\}$. Li and Hui (1994) studied the asymptotic properties of a fixed length vector of robust residual cross-correlations.

3. TESTS BASED ON A SUBSET OF LAGS

By making use of a general result of Li and Hui (1994) on the asymptotic distribution of a vector of robustified residual cross-correlations, we present in this section a robust procedure for testing the hypothesis of non correlation based on the robustified cross-correlations at a particular lag or at a finite number of lags. We also described a robustified version of the procedure for testing causality in the Granger (1969) sense discussed in El Himdi and Roy (1997).

3.1 Asymptotic distribution of robust cross-correlations

The theorem of Li and Hui (1994) is based on RA-estimation of the ARMA parameters and is a direct generalization of a similar result in McLeod (1979) for least squares estimators. It also extends Haugh's (1976) theorem. However, when going through the Taylor series expansions involved in Haugh's proof (see also El Himdi and Roy, 1997), it is seen that Haugh's approach remains valid for any $n^{1/2}$-consistent ARMA estimators. That remark leads us to Theorem 1 below which is slightly more general than the result of Li and Hui (1994) under the null hypothesis of non correlation between the two processes $\{X_t\}$ and $\{Y_t\}$.

Let $j_1, \ldots, j_m$ be a sequence of m distinct integers independent of n such that $|j_i| < n$ and let us consider the vector

$$\boldsymbol{\gamma}_{uv}^{\eta} = (\gamma_{uv}(j_1;\eta), \ldots, \ldots, \gamma_{uv}(j_m;\eta))'.$$

We have the following result.

Theorem 1 *Let $\{X_t\}$ and $\{Y_t\}$ be two second-order stationary and invertible ARMA processes satisfying the assumptions of Section 2.1. Let $\{\hat{u}_t\}$ and $\{\hat{v}_t\}$ be the residual series resulting from $n^{1/2}$-consistent and robust estimation of the parameters. If the two processes are independent, then $\sqrt{n}\boldsymbol{\gamma}_{\hat{u}\hat{v}}^{\eta}$ and $\sqrt{n}\boldsymbol{\gamma}_{uv}^{\eta}$ have the same asymptotic distribution $N_m(0, a\mathrm{I}_m)$, where $a = E[\eta^2(u_1/\sigma_u, v_1/\sigma_v)]$.*

A similar result is given in Haugh (1976) for univariate time series and in El Himdi and Roy (1997) for multivariate time series, in the the particular case $\boldsymbol{\gamma}_{\hat{u}\hat{v}}^{\eta} = \mathrm{r}_{\hat{u}\hat{v}} = (r_{\hat{u}\hat{v}}(j_1), \ldots, r_{\hat{u}\hat{v}}(j_m))'$ and when the parameters are estimated by conditional least-squares. In the next two sections we apply Theorem 1.

3.2 Test based on the robust cross-correlation at a particular lag

Suppose that we want to test in a robust manner that the two stationary and invertible ARMA processes $\{X_t\}$ and $\{Y_t\}$ are uncorrelated, that is:

$$H_0 : \rho_{uv}(j) = 0, \ j \in \mathbb{Z},$$

where $\rho_{uv}(j) = \mathrm{corr}(u_t, v_{t-j})$ is the population cross-correlation of lag j, between $\{u_t\}$ and $\{v_t\}$. Theorem 1 allows us to construct a robust version of the tests given in El Himdi and Roy (1997).

Under H_0, from Theorem 1, we have that $\sqrt{n}\gamma_{\hat{u}\hat{v}}(j_i;\eta)/\sqrt{a}$, $i = 1, \ldots, m$, are asymptotically independent and identically distributed $N(0,1)$. For the alternative hypothesis,

$$H_{1j} : \rho_{uv}(j) \neq 0,$$

it is natural to consider the statistic

$$S_R(j) = n\gamma^2_{\hat{u}\hat{v}}(j;\eta)/\hat{a},$$

and under H_0, $S_R(j)$ follows a χ^2_1 distribution, where $\hat{a}$ is a consistent estimator of a. For example, when η is of the Mallows type, a consistent estimator for a is $\hat{a} = [n^{-1}\sum_{t=1}^{n}\psi^2(\hat{u}_t/\hat{\sigma}_u)] \times [n^{-1}\sum_{t=1}^{n}\psi^2(\hat{v}_t/\hat{\sigma}_v)]$. Thus, at the signification level α, H_0 is rejected if $S_R(j) > \chi^2_{1,1-\alpha}$, where $\chi^2_{m,p}$ is the pth quantile of the χ^2_m distribution. Haugh (1976) used the statistics

$$S(j) = nr^2_{\hat{u}\hat{v}}(j)$$

and noticed from a simulation study that the exact level can be much smaller than the asymptotic level, specially for high lags. He proposed the modified statistic

$$S^*(j) = \frac{n}{n-|j|}S(j) \tag{11}$$

which is much better approximated by the asymptotic distribution. For that reason, we also consider the robust modified statistic

$$S^*_R(j) = \frac{n}{n-|j|}S_R(j) \tag{12}$$

which is asymptotically equivalent to $S_R(j)$.

In practice, we are often interested in considering simultaneously several lags, for example all lags such that $|j| \le M$, where $M \le n-1$. Thus, the alternative hypothesis becomes in that situation

$$H_1^{(M)}: \text{ There exists at least one } j, |j| \le M, \text{ such that } \rho_{uv}(j) \neq 0.$$

A global test for H_0 based on the statistics $S_R(j)$, $|j| \le M$, consists in rejecting H_0 if for at least one lags j, $S_R(j) > \chi^2_{1,1-\alpha_0}$. To obtain a global level α, since the $S_R(j)$ are asymptotically independent, the marginal level α_0 for each test must be $\alpha_0 = 1-(1-\alpha)^{1/(2M+1)}$.

That method has the advantage of individual examination at each lag in a robust way. Portmanteau tests do not share that property. As in El Himdi and Roy (1997), that procedure can be represented graphically with the lag j on the horizontal axis and the value of $S_R(j)$ on the vertical axis. Such a graphical representation allows us to appreciate the dependency between the two residual series at several lags, and can be useful to explaining the reasons of the rejection of the null hypothesis. Such an analysis can also be useful in causality studies.

3.3 Links with causality

In many economic studies, it is important to know the causality directions between two time series when they are correlated at possibly many lags. Following the lines of El Himdi and Roy (1997), the robust cross-correlation analysis described in the previous section can be pursued further for that purpose. Granger causality is now a basic concept in econometrics and is discussed in most textbooks in the field. See for example Hamilton (1994, chap.11) and Lütkepohl (1993, chap.2).

Suppose that $\{(X_t, Y_t), t \in Z\}$ is stationary. In a general framework, the process $\{X_t\}$ does not cause $\{Y_t\}$ in the Granger (1969) sense if the minimum mean square error linear predictor of Y_t based on the information set $\{(X_s, Y_s), s \le t-1\}$ is the same as the one based on the information set $\{Y_s, s \le t-1\}$. In other words, $\{X_t\}$ does not cause $\{Y_t\}$ is Y_t cannot be predicted more efficiently if the information contained in $X_s, s \le t-1$ is taken into account in addition to the information in $Y_s, s \le t-1$. A more formal definition of causality and characterizations of non causality in vector ARMA models in terms of AR and MA parameters are given in Boudjellaba, Dufour and Roy (1992, 1994).

Suppose that $\{X_t\}$ and $\{Y_t\}$ are individually described by univariate stationary and invertible ARMA models. Pierce and Haugh (1977) obtained a characterization of the absence of causality between $\{X_t\}$ and $\{Y_t\}$ in function of the cross-correlation between the two innovation processes $\{u_t\}$ and $\{v_t\}$. They showed that $\{X_t\}$ does not cause $\{Y_t\}$ if and if $\rho_{uv}(j) = 0$, $\forall j < 0$ and $\{Y_t\}$ does not cause $\{X_t\}$ if and only if $\rho_{uv}(j) = 0$, $\forall j > 0$. Therefore, the residual cross-correlation function can be useful for detecting causality directions between $\{X_t\}$ and $\{Y_t\}$.

Suppose that we want to test the null hypothesis of non correlation H_0 against the alternative

$$H_{1M}^{+} : \rho_{uv}(j) \neq 0, \text{ for at least one } j \text{ such that } 1 \leq j \leq M,$$

where $M > 0$ is a fixed integer. It seems natural to consider residual cross-correlations at positive lags to define the test statistic and if H_0 is rejected in favor of H_{1M}^{+}, we will conclude that $\{Y_t\}$ causes $\{X_t\}$. As in the previous section, we can do simultaneous tests at lags $1, 2, \ldots, M$ or employ the portmanteau type statistic

$$S_{RM}^{+*} = \sum_{j=1}^{M} S_R^*(j). \tag{13}$$

which asymptotically follows a χ_M^2 distribution under H_0 and the null hypothesis is rejected for large values of S_{RM}^{+}. Similarly, for the alternative hypothesis

$$H_{1M}^{-} : \rho_{uv}(j) \neq 0, \text{ for at least one } j \text{ such that } -M \leq j \leq 1,$$

the test statistic will be

$$S_{RM}^{-*} = \sum_{j=-M}^{-1} S_R^*(j), \tag{14}$$

which is also asymptotically χ_M^2 under H_0. If H_0 is rejected in favor of H_{1M}^{-}, we will conclude that $\{X_t\}$ causes $\{Y_t\}$.

The procedure described here for causality analysis does not constitute a statistical test in the usual sense, since the hypothesis of non causality which is the hypothesis of interest stands for the alternative hypothesis rather than the null. It is however a robustified version of a similar approach used, among others, by Pierce (1977) and Pierce and Haugh (1977) to identity causality directions between univariate time series. >From a statistical point of view, we would like to directly consider the null hypothesis

$$H_0^{+} : \rho_{uv}(j) \neq 0, j > 0, \text{ or } H_0^{-} : \rho_{uv}(j) \neq 0, j < 0$$

versus the simple negation of H_0^{+} or H_0^{-}. Li and Hui (1994) give the asymptotic distribution of the vector of residual robust cross-correlations $\hat{\boldsymbol{\gamma}}_{\hat{u}\hat{v}}^{\eta}$ in the general case of two correlated univariate ARMA series. The asymptotic covariance matrix is quite involved since it depends on the true models describing the two series, and except for the special case of only instantaneous causality between the two processes $\{X_t\}$ and $\{Y_t\}$, it seems difficult to exploit it in practice.

4. TEST PROCEDURE BASED ON ALL LAGS

For the null hypothesis of non-correlation, Li and Hui (1994) proposed the following portmanteau test statistic

$$S_{RM} = \frac{n}{\hat{a}} \sum_{j=-M}^{M} \gamma_{\hat{u}\hat{v}}^2(j;\eta), \tag{15}$$

and its modified version

$$S_{RM}^* = \frac{n}{\hat{a}} \sum_{j=-M}^{M} \frac{n}{n-|j|} \gamma_{\hat{u}\hat{v}}^2(j;\eta), \tag{16}$$

which are robust versions of Haugh's (1976) statistics. Hong (1996a) introduces another generalization of Haugh's statistic which leads to a class of consistent tests for serial cross-correlation of unknown form between the two time series. Here, we propose a robust version of Hong's approach. It is based on a kernel-based statistic. The kernel $k(\cdot)$ satisfies the following assumption.

Assumption B: The kernel $k : \mathbb{R} \to [-1, 1]$ is a symmetric function, continuous at 0, having at most a finite number of discontinuity points and such that $k(0) = 1$, $\int_{-\infty}^{\infty} k^2(z)dz < \infty$.

The test statistic, denoted Q_{Rn} is the following.

$$Q_{Rn} = \frac{n\hat{a}^{-1}\sum_{j=1-n}^{n-1} k^2(j/m)\gamma_{\hat{u}\hat{v}}^2(j;\eta) - M_n(k)}{[2V_n(k)]^{1/2}}, \tag{17}$$

where $M_n(k) = \sum_{j=1-n}^{n-1}(1-|j|/n)k^2(j/m)$ and $V_n(k) = \sum_{j=2-n}^{n-2}(1-|j|/n)(1-(|j|+1)/n)k^4(j/m)$. The smoothing parameter m is such that $m = m(n) \to \infty$, but $m(n)/n \to 0$ as $n \to \infty$. Using the truncated uniform kernel, we retrieve a standardized version of Li and Hui's (1994) statistic. Since $\lim_{n\to\infty} m^{-1}M_n(k) = M(k) = \int_{-\infty}^{\infty} k^2(z)dz < \infty$, $\lim_{n\to\infty} m^{-1}V_n(k) = V(k) = \int_{-\infty}^{\infty} k^4(z)dz < \infty$, $M_n(k)$ and $V_n(k)$ can be replaced by $mM(k)$ and $mV(k)$ respectively. The resulting statistic is

$$Q_{Rn}^* = \frac{na^{-1}\sum_{j=1-n}^{n-1} k^2(j/m)\gamma_{\hat{u}\hat{v}}^2(j;\eta) - mM(k)}{(2mV(k))^{1/2}}, \tag{18}$$

which is asymptotically equivalent in distribution to Q_{Rn}. These substitutions may lead to better finite sample approximations.

Although Q_{Rn} and Q_{Rn}^* are defined in terms of cross-correlations, there are also coherency-based statistics (in the frequency domain). For a function $g : [-\pi, \pi] \to \mathbb{C}$, the normalized L^2 norm is defined by

$$||g||^2 = (2\pi)^{-1}\int_{-\pi}^{\pi} |g(\omega)|^2 d\omega,$$

where $|\cdot|$ denotes the modulus of a complex number. The cross-spectral density (based on correlations) of two stationary processes $\{u_t\}$ and $\{v_t\}$ is given by

$$f_{uv}(\omega) = \sum_{j=\infty}^{\infty} \rho_{uv}(j)e^{-ij\omega}.$$

To simplify the discussion, assume that η is of the Mallows type. Then, $a^{-1/2}\gamma_{uv}(j;\eta)$ can be interpreted as the sample cross-correlation at lag j of the winsorized series $\{\psi(u_t/\sigma_u)\}$ and $\{\psi(v_t/\sigma_v)\}$. A kernel-based robust estimator is given by

$$\hat{f}_{uv}(\omega) = a^{-1/2}\sum_{j=-n+1}^{n-1} k(j/m)\gamma_{uv}(j;\eta)e^{-ij\omega},$$

whose L^2 norm is

$$||\hat{f}_{uv}||^2 = a^{-1}\sum_{j=-n+1}^{n-1} k^2(j/m)\gamma_{uv}^2(j;\eta).$$

In our situation, $\{u_t\}$ and $\{v_t\}$ are two white noises and the complex coherency (at frequency ω) is given by

$$C_{uv}(\omega) = \frac{f_{uv}(\omega)}{\{f_{uu}(\omega)f_{vv}(\omega)\}^{1/2}} = \frac{2\pi f_{uv}(\omega)}{\sigma_u\sigma_v}.$$

Therefore, Q_{Rn} and Q_{Rn}^* are essentially standardized versions of the estimated coherency $||\hat{C}_{uv}||$.

Under the hypothesis that $\{u_t\}$ and $\{v_t\}$ are two mutually independent processes, the asymptotic distribution of Q_{Rn} is the following. The symbol $\to_L$ stands for convergence in law.

Theorem 2 *Suppose that the assumptions of Section 2.1, assumptions A and B are satisfied, and that* $m \to \infty$, $m/n \to 0$. *If the innovation processes* $\{u_t\}$ *and* $\{v_t\}$, *are independent, then* $Q_{Rn} \to_L N(0,1)$.

The proof which is an adaptation of the one of Hong (1996b) is long and technical and is presented in the Appendix. It is written in two parts. First, we establish the asymptotic normality of the pseudo-statistic

$$\tilde{Q}_{Rn} = \frac{na^{-1}\sum_{j=1-n}^{n-1} k^2(j/m)\gamma_{uv}^2(j;\eta) - M_n(k)}{(2V_n(k))^{1/2}}$$

which is based on the two innovation processes $\{u_t\}$ and $\{v_t\}$. To derive the asymptotic normality, we make use of Brown's (1971) central limit theorem for martingale differences that can be found in many textbooks, namely in Taniguchi and Kakizawa (2000, Chap 1). A more general version for a triangular array of random variables is presented in Fuller (1996, Chap.6). The ARMA models describing the two series do not intervene in the first part since $\tilde{Q}_{Rn}$ is completely determined by the two innovation series $\{u_t\}$ and $\{v_t\}$. The observed data and the estimated models are taken into account in the second part in which it is shown that $\tilde{Q}_{Rn} - Q_{Rn} = o_p(1)$ and Theorem 1 follows.

5. SIMULATION RESULTS

In the previous sections, we have introduced a robustified versions of Hong's (1996a) statistics for checking lagged cross correlation between two ARMA time series which have interesting asymptotic properties. They also extend the robust portmanteau test proposed by Li and Hui (1994) which is a robustified version of Haugh's (1976) classic statistic. >From a practical point of view, it is natural to inquire for the finite sample properties of these statistics, in particular their exact level and power whether or not there is outliers in the series under study. To partially answer that question, a Monte Carlo simulation was conducted. The following modified versions of Haugh's (1976) and Hong's (1996a) statistics S_M^* and Q_n^* were also included in the experiment.

$$S_M^* = n \sum_{j=-M}^{M} \frac{n}{(n-|j|)} r_{\hat{u}\hat{v}}^2(j), \tag{19}$$

and

$$Q_n^* = \frac{n \sum_{j=1-n}^{n-1} k^2(j/m) r_{\hat{u}\hat{v}}^2(j) - mM(k)}{[2mV(k)]^{1/2}}. \tag{20}$$

5.1 Description of the experiment

To compare the omnibus robust statistics Q_{Rn}, Q_{Rn}^* to their non-robust counterparts Q_n, Q_n^* and also to the portmanteau statistics S_M^*, S_{RM} and S_{RM}^*, independent realizations were generated from the following bivariate Gaussian VAR(1) process:

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} 0.5 & 0.0 \\ 0.0 & 0.5 \end{pmatrix} \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} + \begin{pmatrix} u_t \\ v_t \end{pmatrix}, t \in \mathbb{Z}. \tag{21}$$

The covariance matrix of $(u_t, v_t)'$ is

$$\Sigma = \begin{pmatrix} 1.0 & \rho \\ \rho & 1.0 \end{pmatrix}. \tag{22}$$

We considered $\rho = 0.0$ and $\rho = 0.2$. The first value corresponds to the situation where two independent AR(1) series are generated, and that situation allows us to study the level of the tests. In the second case, the two series are serially correlated and it allows us to compare the powers.

For each realization, the portmanteau statistics were computed for $M = 6, 12, 18$. Each omnibus statistic was calculated for five different kernels that are described in Table 1. For each kernel, three different rates m given respectively by $\lfloor \log(n) \rfloor$, $\lfloor 3.5n^{0.2} \rfloor$ and $\lfloor 3n^{0.3} \rfloor$, were employed ($\lfloor a \rfloor$ denotes the integer part of a). These rates are discussed in Hong (1996c, p. 849). They lead respectively to the values $m = 5, 8, 12$ for the series length $n = 100$ and to $m = 5, 9, 15$ with $n = 200$.

For more details on kernels, see Priestley (1981).

To investigate the effect of outliers on the robust and non robust statistics, 10 000 realizations were generated from the Gaussian VAR(1) model (21) for each series length ($n = 100, 200$). For each realization, the following three scenarios were considered:

Scenario 1 No contamination.

Scenario 2 For $n = 100$, we subtracted systematically the value 10 to observations X_{26} and Y_{76}. Furthermore, for $n = 200$, we added the value 10 to X_{101} and Y_{151}, and we subtracted 10 to X_{176}.

Table 1: Kernels used in the calculation of the omnibus statistics Q_n and Q_{Rn}

Truncated uniform kernel (TR): $k(z) = \begin{cases} 1 & \text{if } |u| \leq 1, \\ 0 & \text{elsewhere.} \end{cases}$

Bartlett (BAR): $k(z) = \begin{cases} 1 - |z| & \text{if } |z| \leq 1, \\ 0 & \text{elsewhere,} \end{cases}$

Daniell (DAN): $k(z) = \dfrac{\sin(\pi z)}{\pi z}, \; z \in \mathbb{R}.$

Parzen (PAR): $k(z) = \begin{cases} 1 - 6(\pi z/6)^2 + 6|\pi z/6|^3 & \text{if } |z| \leq 3/\pi, \\ 2(1 - |\pi z/6|)^3 & \text{if } 3/\pi \leq |z| \leq 6/\pi, \\ 0 & \text{elsewhere.} \end{cases}$

Bartlett-Priestley (BP): $k(z) = \dfrac{9}{5\pi^2 z^2}\{\dfrac{\sin(\pi\sqrt{5/3}z)}{\pi\sqrt{5/3}z} - \cos(\pi\sqrt{5/3}z)\}, z \in \mathbb{R}.$

Scenario 3 In addition to scenario 2, for $n = 100$, we added to X_{51} and Y_{51} the value 10 and for $n = 200$, we subtracted to X_{126} and Y_{126} the value 10.

Similar scenarios were employed in Li (1988) and Li and Hui (1994). The first scenario allows us to evaluate the performance of the statistics in the context of Gaussian VAR(1) series. The second scenario permits to evaluate the performance of the procedures when outliers occur at different points in time. Finally, in the third scenario, outliers occur in both series at the same points in time.

For the robust statistics, we chose the Mallows type functions η defined by (6) and two ψ functions were employed: Huber (8) and bisquare (9) families. The tuning constant c in Huber family is 1.65 and in the bisquare family is 5.58 (Bustos and Yohai, 1986; Li and Hui, 1994).

In the level study, 10 000 independent realizations were generated from model (21) and (22) with $\rho = 0$ for each value of n and the computations were made in the following way. All the computer programs were written in Fortran 77.

(1) The Gaussian white noise $(u_t, v_t)'$, $t = 1, \ldots, n$, was generated using the subroutine G05EZF from the NAG library.

(2) The initial values $(X_0, Y_0)'$ was generated from the exact bivariate Gaussian distribution using Ansley's (1980) algorithm. The values $(X_t, Y_t)', t = 1, \ldots, n$, were obtained by solving the difference equation (21).

(3) For each realization, univariate AR(1) models were separately estimated for each of the two series and the residual series $\{\hat{u}_t, t = 1, \ldots, n\}$ and $\{\hat{v}_t, t = 1, \ldots, n\}$ were obtained. The autoregressive parameter was estimated by least squares for the non robust statistics and the RA estimators was obtained from the iterative algorithm described in Bustos and Yohai (1986, Section 2) for the robust statistics. The Mallows type function η combined with Huber and bisquare functions ψ were used. The three scenarios for outliers were applied to each realization, which leads to 9 distinct estimates and therefore to 9 residual series $\{(\hat{u}_t, \hat{v}_t)', t = 1, 2, \ldots, n\}$.

(4) For each realization, the test statistic S_M^*, Q_n and Q_n^* were computed from the least-squares residuals. The value of S_{RM}, S_{RM}^*, Q_{Rn} and Q_{Rn}^* were obtained from the RA residuals.

(5) Finally, for each series of length n and for each of the three nominal level (1, 5 and 10%), we obtained for each statistic, the number of rejections of the null hypothesis of non correlation between the two series, based on 10 000 realizations. The detailed results can be found in Duchesne (2000, Chap. 4) and a summary is presented in Table 2. The standard error of the number of rejections is 9.9 for the nominal level 1%, 21.8 for 5% and 30 for 10%. For example, at the nominal level 5%, the observed number of rejections is not significantly different from the expected number of rejections if it lies

in the interval [458, 542] at the significance level $\alpha = 0.05$ and in the interval [444, 556] at the significance level $\alpha = 0.01$.

The power study was conducted in a similar way except that the correlation between u_t and v_t is $\rho = 0.2$.

5.2 Discussion of the level study

For briefness, we just present the results for the nominal level 5% in Table 2. Further, since the results for Q_n and Q_{Rn} are similar to those for Q_n^* and Q_{Rn}^* respectively, only the results for the modified statistics are reported. Finally, since the numbers of rejections for the three kernels DAN, PAR and BP are similar, we just give the results for Daniell kernel which is optimal (greatest asymptotic power) kernel in a certain class of smooth kernels for Q_n and Q_n^* (Hong, 1996a, Section 4).

Since the nominal level is 5%, all the numbers in Table 2 must be compared to 500. Under scenario 1 (no outlier), the levels of the portmanteau statistics S_M^* and S_{RM}^* are perfectly controled. Hong's statistic Q_n^* tends to overreject. When $m = 5$, its empirical level varies arround 7% with the BAR and DAN kernels, around 8% with the TR kernel and becomes closer to 5% as m increases. A similar trend was observed by Hong (1996a). The behavior of the robustified version Q_{Rn}^* is similar to the one of Q_n^* irrespective of the ψ function.

Under scenarios 2 and 3, it is immediately seen that the unrobustified statistics S_M^* and Q_n^* are very sensitive to outliers. In general, they dramatically underreject under scenario 2 and overreject under scenario 3. By opposition, the empirical level of the robustified statistics S_M^* under scenarios 2 and 3, those of Q_{Rn}^* with either Huber or the bisquare function ψ under scenario 2 and those of Q_{Rn}^* with the bisquare function are similar to the ones observed under scenario 1 and are therefore reasonably close to 5 %. However, Q_{Rn}^* with Huber function overrejects considerably under scenario 3, specially for small values of m with the three kernels and for all values of m with the kernel BAR. For that reason, the bisquare function seems preferable in practice.

Globally as a function of m, the sizes of the robust statistics are improving as m increases. With respect to the series length, there is no obvious improvement when n goes from 100 to 200.

5.3 Discussion of the power study

The number of rejections under the alternative hypothesis based on 10 000 realizations at the 5 % nominal level are reported in Table 3. The values on lines 1, 3 and 4 are based on the asymptotic critical values for scenarios 1, 2 and 3 respectively, whilst those in parentheses on line 2 correspond to scenario 1 and are based on the empirical critical values (ECV) obtained from the level study. In general, the powers obtained with the ECV are slightly smaller that those found with the ACV. It is not surprising since generally, the omnibus statistics Q_n^* and Q_{Rn}^* overreject. The power of S_M^* and Q_n^* under scenarios 2 and 3 are of limited interest since these statistics are severally affected by outliers. The only scenario for which all the tests are comparable is scenario 1. First, it is very interesting to notice that the price to pay for using a robustified statistic rather than its unrobustified version is rather small. Indeed, the power of S_{RM}^* is smaller than the one for S_M^* but the difference is rather small and the same can be said of Q_{Rn}^* in comparison with Q_n^*. Since the behavior of Q_n^* was already discussed in Hong (1996a), we focus on its robustified version Q_{Rn}^*.

Among the three kernels, the BAR kernel is in general more powerful than the DAN kernel and the latter is considerably more powerful than the TR kernel. This observation does not contradict the optimality result of Hong (1996a) which says that the DAN kernel is optimal in a certain class of smooth kernels since the BAR kernel is not in that class.

The two ψ functions lead to similar powers in scenarios 1 and 2. In some cases with scenario 3, Huber functions leads to larger powers than the bisquare function. It is very likely related to the fact that the empirical levels are much larger than the nominal level.

As a function of m, the power of Q_{Rn}^* decreases considerably as m increases except with Huber function in scenario 3. As for Hong's statistic Q_n^*, the power of Q_{Rn}^* increases very slightly when n goes from 100 to 200. Finally, among the three kernels, the TR kernel is the lest powerful but it is still more powerful than Li and Hui's (1994) robust portmanteau statistics S_M^*. The empirical results presented here strongly support the use of the new robust statistic Q_{Rn}^* when additive outliers are suspected. Indeed, the loss of

Table 2: Number of rejections under the null hypothesis based on 10 000 realizations at the 5% nominal level. For each value of M or m, the three lines correspond to the three scenarios.

								Q^*_{Rn}					
					Q^*_n			Huber			bisquare		
M	S^*_M	$S^*_{RM}(H)$	$S^*_{RM}(b)$	m	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR
$n = 100$													
6	460	469	478	5	753	657	835	776	657	842	769	664	828
	36	463	462		175	119	121	737	625	828	748	649	850
	2266	565	457		9480	8632	4754	1101	894	988	740	651	822
12	488	487	496	8	673	555	645	670	542	651	674	544	647
	19	510	471		94	70	42	650	552	636	670	556	658
	309	544	484		8581	6722	1876	960	715	750	661	545	654
18	469	478	484	12	575	432	473	566	425	470	580	438	477
	10	534	476		57	31	17	586	427	500	590	437	461
	58	576	471		6902	4073	407	767	539	535	565	436	474
$n = 200$													
6	509	505	515	5	781	669	899	764	666	898	779	687	901
	21	480	509		156	95	73	714	627	865	762	677	900
	7211	711	495		9997	9977	9101	1673	1365	1307	786	678	933
12	532	499	504	9	678	590	737	677	585	709	676	582	732
	9	496	500		62	45	28	636	564	696	671	575	715
	2186	619	499		9936	9740	5581	1348	1088	927	701	609	717
18	467	478	472	15	614	513	546	602	501	525	607	489	531
	8	494	481		28	31	10	578	482	557	589	491	554
	514	622	496		9498	9105	1437	1088	827	687	595	478	556

Table 3: Number of rejections under the alternative hypothesis based on 10 000 realizations at the 5 % nominal level. For each value of M or m, the three lines correspond to the three scenarios. The values in parentheses represent the number of rejections based on the empirical critical values for scenario 1.

								Q^*_{Rn}					
					Q^*_n			Huber			bisquare		
M	S^*_M	$S^*_{RM}(H)$	$S^*_{RM}(b)$	m	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR
$n = 100$													
6	1750	1656	1656	5	4781	3802	2876	4571	3658	2724	4562	3655	2705
	(1845)	(1744)	(1720)		(4033)	(3400)	(2153)	(3890)	(3242)	(2001)	(3861)	(3229)	(1968)
	98	1432	1576		924	546	265	3962	3068	2337	4248	3388	2533
	5254	2240	1584		9948	9787	7745	5970	4930	3543	4446	3528	2576
12	1349	1283	1291	8	3964	3020	2022	3814	2866	1947	3821	2863	1936
	(1382)	(1315)	(1310)		(3520)	(2871)	(1668)	(3421)	(2746)	(1642)	(3402)	(2733)	(1621)
	40	1168	1228		545	276	109	3213	2370	1693	3561	2621	1869
	1246	1708	1257		9782	9153	4559	5092	3976	2530	3693	2708	1891
18	1138	1075	1088	12	3288	2301	1411	3129	2129	1350	3124	2116	1347
	(1198)	(1117)	(1112)		(3067)	(2488)	(1460)	(2918)	(2325)	(1417)	(2926)	(2311)	(1398)
	23	1023	1024		298	112	39	2615	1773	1211	2901	2000	1275
	308	1459	1037		9241	7323	1623	4301	3021	1808	3003	2019	1318
$n = 200$													
6	3780	3585	3555	5	7716	6858	5304	7489	6639	5082	7492	6623	5077
	(3748)	(3568)	(3510)		(7145)	(6394)	(4103)	(6919)	(6157)	(3945)	(6902)	(6137)	(3933)
	128	2817	3258		1422	890	377	6556	5627	4137	7095	6185	4663
	9612	4883	3351		10000	10000	9943	8787	8138	6469	7271	6339	4761
12	2687	2528	2545	9	6769	5691	3872	6540	5459	3671	6534	5441	3676
	(2583)	(2536)	(2538)		(6292)	(5376)	(3218)	(6074)	(5156)	(3055)	(6055)	(5147)	(3083)
	42	2021	2286		746	430	120	5538	4426	3005	6113	4960	3382
	6304	3558	2419		10000	9996	9072	8036	6999	4873	6281	5134	3514
18	2152	2058	2051	15	5749	4410	2662	5539	4192	2516	5499	4177	2510
	(2238)	(2122)	(2121)		(5429)	(4377)	(2522)	(5169)	(4192)	(2428)	(5178)	(4202)	(2422)
	16	1650	1848		323	241	30	4500	3363	2034	5071	3811	2296
	2546	2798	1921		9985	9938	5016	7027	5747	3418	5261	3939	2390

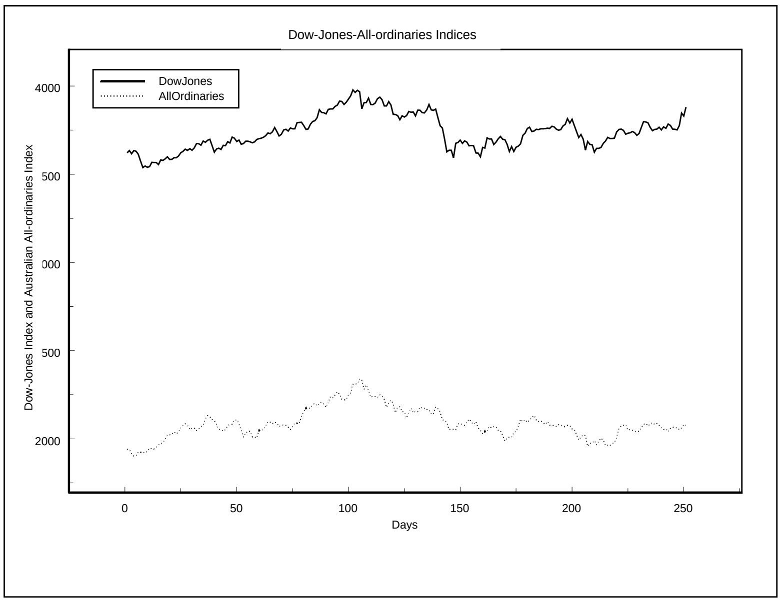
power in using Q^*_{Rn} rather than its unrobustified version Q^*_n is not important and we have also seen that Q^*_n is very sensitive to outliers. If the outliers are at the same point in time, the bisquare function seems preferable to Huber function.

6. APPLICATION

6.1 The data

We illustrate here the new robust tests for non correlation with a real data set coming from the financial literature. It was analyzed by Brockwell and Davis (1996, Chap. 7) who aimed at developing a stationary bivariate time series model. However, in practice, before deciding to build a multivariate model, a first important step is to check for cross-correlation between the series. If we conclude that they are indeed dependent, it is then appropriate to develop a multivariate model.

Figure 1: Closing values of the Dow-Jones Index and of the Australian All-ordinaries index, for the period September 13, 1993 to August 26, 1994.



The variables are the closing values of the Dow-Jones Index of stocks (D_t) on the New York Stock Exchange and the closing values of the Australian All-ordinaries Index (A_t) of Share Prices during 251 successive trading days from September 13, 1993 to August 26th, 1994. The two series are displayed in Figure 1. According to the efficient market hypothesis, these two series should behave individually as random walks with uncorrelated increments. In order to have stationary series, the data are transformed as percentage relative price changes:

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} 100(D_t - D_{t-1})/D_{t-1} \\ 100(A_t - A_{t-1})/A_{t-1} \end{pmatrix},$$

$t = 1, \ldots, 250$. The sample autocorrelations of each series indicate that both of them can be considered as white noise since all their autocorrelations are within or very close to the significant limits $\pm 1.96 n^{-1/2}$ (Brockwell and Davis, 1996, Example 7.1.1). The sample cross-correlations are also insignificant except at lag -1 where $r_{XY}(-1) = 0.46$, which indicates that X_{t-1} and Y_t are dependent.

6.2 Correlation analysis

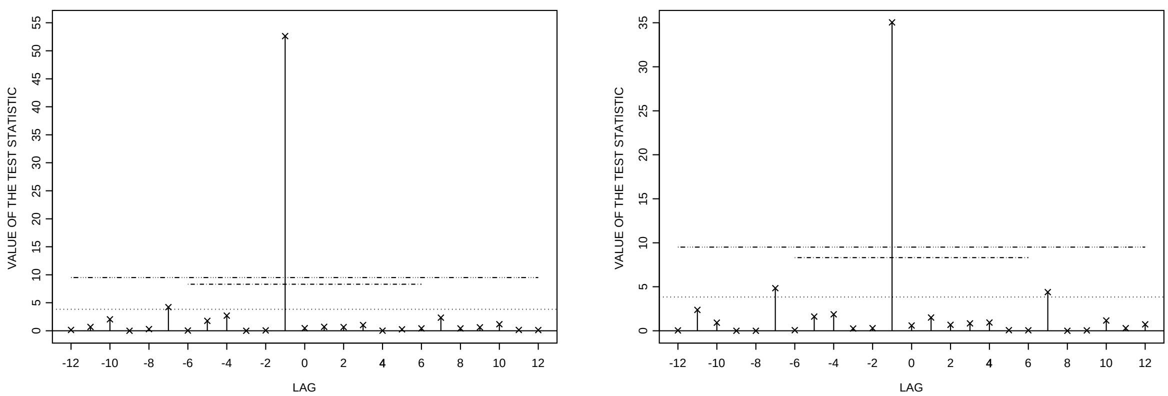
The test procedures described in the previous sections can be directly applied to the two series since there is no significant serial autocorrelation. The values of the statistics $S^*(j)$ defined by (11), $|j| \leq 12$, are presented in Figure 2 (left part). With tests at individually lags, we strongly reject the hypothesis of non-correlation at lag $j = -1$. We also reject with simultaneous tests at lags $-6, \ldots, +6$ and even at lags $-12, \ldots, +12$ at the global level $\alpha = 0.05$. The values of the robust statistics $S^*_R(j)$, $|j| \leq 12$, are also

Table 4: Values of the statistics Q_n^* and Q_{Rn}^* for the real data and for the data contaminated by two outliers.

	Real data		Contaminated data	
M	Q_n^*	Q_{Rn}^*	Q_n^*	Q_{Rn}^*
5	16.75	11.25	0.45	10.18
9	13.58	9.05	0.37	8.18
16	10.43	7.03	-0.13	6.27

shown in the right part of Figure 2. Again, the non-correlation hypothesis is rejected by the test at the individual lag $j = -1$ and by the simultaneous tests at the global level $\alpha = 0.05$.

Figure 2: Values of the statistics $S^*(j)$ (left part) and of $S_R^*(j)$ (right part) for different lags j with the real data. The horizontal dotted line represents the marginal critical value at the level $\alpha = 0.05$. The dash lines give the critical values at the global level $\alpha = 0.05$, for simultaneous tests at lags $j = -6, \ldots, 6$ and $j = -12, \ldots, 12$.



In order to take into account all possible lags, Hong's test Q_n^* and its robustified version Q_{Rn}^* where carried out. The statistic Q_{Rn}^* was calculated with the Daniell kernel and for the three values $5, 9, 16$ for m that correspond to the three rates employed in Section 5. The scale parameters σ_X and σ_Y in (4) were estimated with the median estimator defined by (5). The values of the statistics are reported in Table 4. At the 1% significance level, these values must be compared to the 99th quantile of the $N(0,1)$ distribution and again, the hypothesis of non correlation is rejected.

Since there is no apparent outliers in the two series under study, we have replaced X_{219} by -5.5 and X_{205} by +5.5 which gives rise to two additive outliers in the Dow-Jones series. The contaminated series $\{X_t\}$ and the series $\{Y_t\}$ are shown in Figure 3. Their behavior is comparable to the one of many financial time series. The new series $\{X_t\}$ can still be considered as a white noise and the cross-correlation at lag $j = -1$ is now 0.14 rather that 0.46. For these new series, we have redone the correlation analysis. With the non robust procedure, cross-correlation is detected by the tests at individual lags as shown in Figure 4 (left part). However, no significant correlation is detected by the simultaneous tests at lags $j = -12, \ldots, 12$ and $j = -6, \ldots, 6$ as well as with Hong's statistic whose values are reported in Table 4. With the robust procedure as illustrated in Figure 2 (right part) for the statistic $S_R^*(j)$ and in Table 4 for the statistics Q_{Rn}^*, the hypothesis of non correlation is strongly rejected once again.

6.3 A descriptive causality analysis

Once the hypothesis H_0 of non correlation between the two series is rejected, it is natural to inquire into the directions of causality between the two variables, specially in economic studies. In Section 3.3, we have seen that $\{X_t\}$ does not cause $\{Y_t\}$ if and only if $\rho_{uv}(j) = 0$, $\forall j < 0$, and $\{Y_t\}$ does not cause $\{X_t\}$ is

Figure 3: The two variables X_t and Y_t with two outliers.

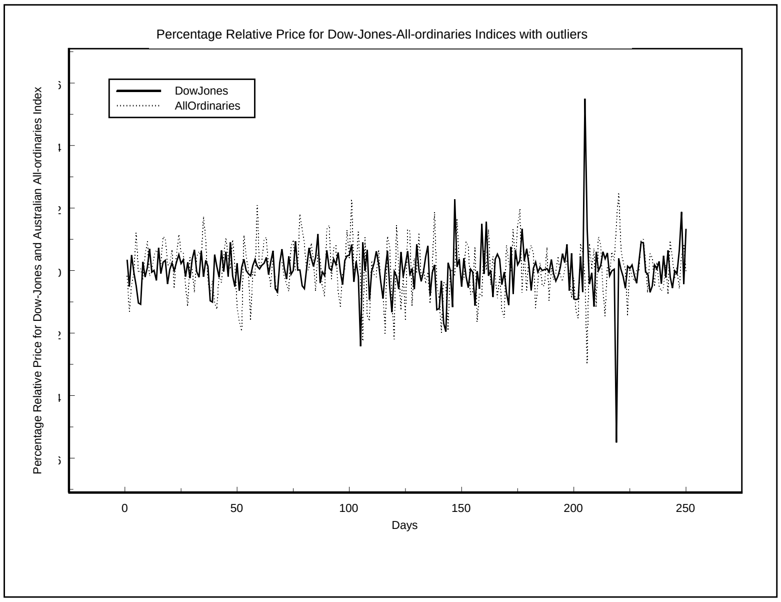


Table 5: p-values of the non-robust tests S_M^{+*}, S_M^{-*} and of the robust tests S_{RM}^{+*} and S_{RM}^{-*} for the real data and for the contaminated data.

	Real data				Contaminated data			
M	S_M^+	S_M^-	S_{RM}^+	S_{RM}^-	S_M^+	S_M^-	S_{RM}^+	S_{RM}^-
1	0.40	0^+	0.22	0^+	0.39	0.03	0.20	0^+
2	0.51	0^+	0.34	0^+	0.47	0.09	0.30	0^+
3	0.50	0^+	0.39	0^+	0.68	0.19	0.39	0^+
6	0.80	0^+	0.67	0^+	0.76	0.16	0.72	0^+
12	0.79	0^+	0.55	0^+	0.89	0.23	0.66	0^+

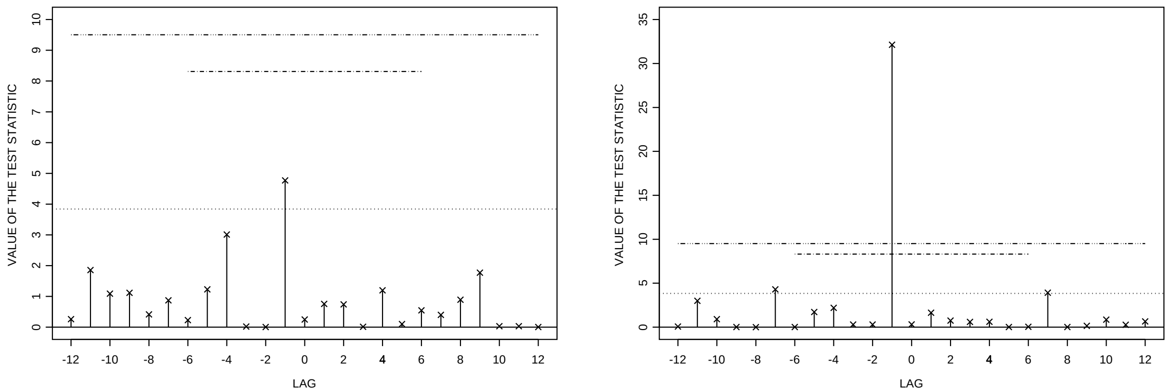
Note: 0^+ denotes a positive number smaller than 10^{-5}.

and only if $\rho_{uv}(j) = 0$, $\forall j > 0$, where $\{u_t\}$ and $\{v_t\}$ are the two corresponding innovation processes. In the first part of the correlation analysis of the previous section, we have tested H_0 with statistics based on individual lags and it is immediately seen from Figure 2 that $\rho_{uv}(-1)$ is significantly different from zero and all the other lags, the values of the test statistics are either smaller than the 5% critical value or very close of it. Figure 2 leads us to think that the Dow-Jones Index influenced the Australian Index and that the Australian Index does not influence the Dow-Jones. Given the relative size of the two economies, it is not a surprising conclusion.

In order to simultaneously consider many lags, we can employ the statistic $S_M^{+*} = \sum_{j=1}^{M} S^*(j)$ or $S_M^{-*} = \sum_{j=-M}^{-1} S^*(j)$ which are asymptotically χ^2_M under H_0. For H_0 against H_{1M}^+, the p-values of S_M^{+*} for various values of M are presented in Table 5. For H_0 against H_{1M}^-, similar results are given for S_M^{-*}. Thus, for all the values of M considered, H_0 is rejected against H_{1M}^- and not rejected against H_{1M}^+. These tests confirms the conclusion deduced from the graphical analysis. The robust statistics S_{RM}^{+*} and S_{RM}^{-*} defined by (13) and (14) respectively were also employed for these data and the conclusion are the same.

The causality analysis was redone with the contaminated data and the resulting p-values are shown in Table 5. With the non robust statistics, H_0 is not rejected against H_{1M}^+ as before. Against H_{1M}^-, H_0 is barely rejected at the 5% level at $M = 1$ and is not rejected at the other values of M. It is therefore difficult to conclude in that case. However, with the robust statistics, we retrieve the conclusions obtained with the true data.

Figure 4: Values of the statistics $S^*(j)$ (left part) and of $S_R^*(j)$ (right part) for different lags j with the contaminated data. The horizontal dotted line represents the marginal critical value at the level $\alpha = 0.05$. The dash lines give the critical values at the global level $\alpha = 0.05$, for simultaneous tests at lags $j = -6, \ldots, 6$ and $j = -12, \ldots, 12$.



CONCLUSION

We proposed in this paper a robust version of Hong's statistic for checking the independence of two univariate stationary and invertible ARMA time series. The robustified test statistic is obtained by replacing the usual residual cross-correlations by the robust ones, as introduced by Li and Hui (1994). We established under the null hypothesis the asymptotic normality of the new robustified test statistic. We described a robust procedure for checking independence at individual lags, following an approach similar to El Himdi and Roy (1997). A descriptive approach for causality analysis in the Granger's (1969) sense is also discussed. In a Monte Carlo experiment, we analyzed the level and power of the new statistics and we compare them to their non robust versions. We used the residual robust autocovariance estimators introduced by Bustos and Yohai (1986) to perform the robust estimation of the ARMA parameters. However, any $n^{1/2}$-consistent robust method in conjunction with the use of the residual robust cross-correlation function are sufficient for the asymptotic normality. The empirical results presented show that Hong's test is very sensitive to outliers and support the use of the new robust statistic when additive outliers are suspected. Our empirical results seem to indicate that the loss of power in using the new test rather than its unrobustified version is not important.

APPENDIX

Here is the proof of Theorem 2. In the following, Δ represents a finite positive constant that can differ from one place to another. Let $k_{jm} = k(j/m)$ and $\eta_{t,s} = \eta(u_t/\sigma_u, v_s/\sigma_v)$. First, we establish the following result:

Part 1 *Under the assumptions of Theorem 2, $\tilde{Q}_{Rn} \rightarrow_L N(0,1)$.*

Proof: Let $T_n = na^{-1}\sum_{j=-n+1}^{n-1} k_{jm}^2 \gamma_{uv}^2(j;\eta)$. We treat separately the two cases $j \geq 0$ and $j < 0$ and in both of them, we decompose $\gamma_{uv}^2(j;\eta)$ in two terms: For $j \geq 0$, we show that

$$\gamma_{uv}^2(j;\eta) = n^{-2}\sum_{t=j+1}^{n} \eta_{t,t-j}^2 + 2n^{-2}\sum_{t=j+2}^{n}\sum_{s=j+1}^{t-1} \eta_{t,t-j}\eta_{s,s-j},$$

and if $j < 0$, we have that

$$\gamma_{uv}^2(j;\eta) = n^{-2}\sum_{t=j+1}^{n} \eta_{t-j,t}^2 + 2n^{-2}\sum_{t=j+2}^{n}\sum_{s=j+1}^{t-1} \eta_{t-j,t}\eta_{s-j,s}.$$

We can write T_n in the following manner:

$$\begin{aligned} T_n &= na^{-1}\sum_{j=0}^{n-1} k_{jm}^2\gamma_{uv}^2(j;\eta) + na^{-1}\sum_{j=0}^{n-1} k_{jm}^2\gamma_{uv}^2(-j;\eta), \\ &= (H_{1n}+H_{2n}) + (W_{1n}+W_{2n}) = H_n + W_n, \end{aligned}$$

where

$$\begin{aligned} H_{1n} &= \frac{1}{na}\sum_{j=0}^{n-1} k_{jm}^2(\sum_{t=j+1}^{n}\eta_{t,t-j}^2), \quad H_{2n} = \frac{1}{na}\sum_{j=1}^{n-1} k_{jm}^2(\sum_{t=j+1}^{n}\eta_{t-j,t}^2), \\ W_{1n} &= \frac{2}{na}\sum_{j=0}^{n-2} k_{jm}^2(\sum_{t=j+2}^{n}\sum_{s=j+1}^{t-1}\eta_{t,t-j}\eta_{s,s-j}), \quad W_{2n} = \frac{2}{na}\sum_{j=1}^{n-2} k_{jm}^2(\sum_{t=j+2}^{n}\sum_{s=j+1}^{t-1}\eta_{t-j,t}\eta_{s-j,s}). \end{aligned}$$

Let us evaluate $E(H_n)$. Since $a = E(\eta_{s,t}^2)$, $\forall s,t$, we have that

$$E(H_{1n}) = \frac{1}{an}\sum_{j=0}^{n-1} k_{jm}^2[\sum_{t=j+1}^{n} E(\eta_{t,t-j}^2)] = \frac{1}{an}\sum_{j=0}^{n-1} k_{jm}^2(n-j)a.$$

Also, $E(H_{2n}) = \sum_{j=1}^{n-1} k_{jm}^2(1-j/n)$, and it follows that $E(H_n) = M_n(k)$. We further show that $E(W_n) = 0$. Note that for $s < t$, $E(\eta_{t-j,t}\eta_{s-j,s}) = 0$ and $E(\eta_{t,t-j}\eta_{s,s-j}) = 0$, since $\eta(\cdot,\cdot)$ is an odd function in each variable, that $\{u_t\}$ and $\{v_t\}$ are independent strong white noises with symmetric distributions. Then, we can affirm that

$$\sum_{j=-n+1}^{n-1} k_{jm}^2\gamma_{uv}^2(j;\eta) = O_p(m/n).$$

To establish Part 1, we will show the following two results:

$$\begin{aligned} (i)\ \frac{H_n - M_n(k)}{[2V_n(k)]^{1/2}} &\to_P \quad 0, \\ (ii)\ \frac{W_n}{[2V_n(k)]^{1/2}} &\to_L \quad N(0,1). \end{aligned}$$

We begin with (i). The result is established if we prove the next proposition:

Proposition 1 $E[H_{jn} - E(H_{jn})]^2 = O(m^2/n),\ j = 1,2.$

Proof:
We show the result for $j = 1$. The case $j = 2$ is similarly obtained. We note that

$$H_{1n} - E(H_{1n}) = \sum_{j=0}^{n-1} k_{jm}^2[n^{-1}\sum_{t=j+1}^{n}(\eta_{t,t-j}^2/a - 1)]$$

and using Minkovski's inequality, we have that

$$E[H_{1n} - E(H_{1n})]^2 \le \{\sum_{j=0}^{n-1} k_{jm}^2(E[n^{-1}\sum_{t=j+1}^{n}(\eta_{t,t-j}^2/a - 1)]^2)^{1/2}\}^2.$$

Proposition 1 is verified if we show that

$$E[n^{-1}\sum_{t=j+1}^{n}(\eta_{t,t-j}^2/a - 1)]^2 = O(n^{-1}), \text{uniformly in } j.$$

But we have

$$[n^{-1}\sum_{t=j+1}^{n}(\eta_{t,t-j}^2/a-1)]^2 = n^{-2}\sum_{t=j+1}^{n}(\eta_{t,t-j}^2/a-1)^2 + 2n^{-2}\sum_{t=j+2}^{n}\sum_{s=j+1}^{t-1}(\eta_{t,t-j}^2/a-1)(\eta_{s,s-j}^2/a-1),$$

and thus

$$E[n^{-1}\sum_{t=j+1}^{n}(\eta_{t,t-j}^2/a-1)]^2 = \frac{n-j}{n^2}E[\eta_{1,1}^2/a-1]^2 = O(n^{-1}).$$

Since $E[(H_n - M_n(k))/(2V_n(k))^{1/2}] = 0$ and $\text{var}[(H_n - M_n(k))/(2V_n(k))^{1/2}] = \text{var}(H_n)/(2V_n(k)) = O(m/n)$, and since by hypothesis $m/n \to 0$, we have shown (i).

We now show (ii). We do a change in the indices of summation noting that $\sum_{j=1}^{n-2}\sum_{t=j+2}^{n}\sum_{s=j+1}^{t-1} = \sum_{t=3}^{n}\sum_{s=2}^{t-1}\sum_{j=1}^{s-1}$. Also $\sum_{j=0}^{n-2}\sum_{t=j+2}^{n}\sum_{s=j+1}^{t-1} = \sum_{t=2}^{n}\sum_{s=1}^{t-1}\sum_{j=0}^{s-1}$. We can write

$$W_n = 2a^{-1}\eta_{1,1}(n^{-1}\sum_{t=2}^{n}\eta_{t,t}) + W_n',$$

where W_n' is defined by

$$\begin{aligned} W_n' &= n^{-1}\sum_{t=3}^{n}\sum_{s=2}^{t-1}\sum_{j=0}^{s-1}2a^{-1}k_{jm}^2\eta_{t,t-j}\eta_{s,s-j} + n^{-1}\sum_{t=3}^{n}\sum_{s=2}^{t-1}\sum_{j=1}^{s-1}2a^{-1}k_{jm}^2\eta_{t-j,t}\eta_{s-j,s}, \\ &= n^{-1}\sum_{t=3}^{n}(W_{1nt}+W_{2nt}) = n^{-1}\sum_{t=3}^{n}W_{nt}, \end{aligned}$$

with the following definitions for W_{1nt} and W_{2nt}:

$$\begin{aligned} W_{1nt} &= \sum_{s=2}^{t-1}\sum_{j=0}^{s-1}2a^{-1}k_{jm}^2\eta_{t,t-j}\eta_{s,s-j}, \\ W_{2nt} &= \sum_{s=2}^{t-1}\sum_{j=1}^{s-1}2a^{-1}k_{jm}^2\eta_{t-j,t}\eta_{s-j,s}. \end{aligned}$$

We remark that $2a^{-1}\eta_{1,1}(n^{-1}\sum_{t=2}^{n}\eta_{t,t}) = o_p(1)$ and that term can be neglected in the sequel. We note that $(W_{nt},\mathcal{F}_t)$ is a martingale difference sequence, where $\mathcal{F}_t = \sigma[(u_s,v_s), s \le t]$. Indeed, we verify that $E(W_{1nt}|\mathcal{F}_{t-1}) = 0$ since $E[\eta_{t,t-j}\eta_{s,s-j}|\mathcal{F}_{t-1}] = 0$. We proceed in a similar way for W_{2nt}. To prove (ii), we use a theorem of Brown (1971), and in our context that theorem says that $W_n/\sqrt{\text{var}(W_n)} \to_L N(0,1)$ if we verify the following two conditions:

$$(A1) \qquad \frac{n^{-2}}{\text{var}(W_n)}\sum_{t=3}^{n}E[W_{nt}^2\chi(|W_{nt}| > \epsilon n[\text{var}(W_n)]^{1/2})] \to 0,\ \forall\epsilon > 0,$$

$$(A2) \qquad \frac{n^{-2}}{\text{var}(W_n)}\sum_{t=3}^{n}\ddot{W}_{nt}^2 \to_P 1,$$

where $\ddot{W}_{nt}^2 = E(W_{nt}^2|\mathcal{F}_{t-1})$ and $\chi(A)$ denotes the indicator function of the set A. We begin with (A1). It is sufficient to verify the following Liapounov's condition:

$$\sigma^{-4}(n)n^{-4}\sum_{t=3}^{n}E(W_{nt}^4) \to 0,$$

where $\sigma^2(n) = 2V_n(k)$. The following proposition gives the variance of W_n'.

Proposition 2 $\text{var}(W'_n) = \sigma^2(n)$.

Proof:
$E(W_{nt}^2) = E(W_{1nt}^2) + E(W_{2nt}^2)$, since $E(W_{1nt}W_{2nt}) = 0$, this can be deduced from the fact that

$$E[\eta_{t,t-j_1}\eta_{s_1,s_1-j_1}\eta_{t-j_2,t}\eta_{s_2-j_2,s_2}] = 0, t > s_1, t > s_2.$$

Developing W_{1nt}^2, we obtain

$$\begin{aligned} W_{1nt}^2 &= 4a^{-2}\sum_{s=2}^{t-1}(\sum_{j=0}^{s-1} k_{jm}^2 \eta_{t,t-j}\eta_{s,s-j})^2 + \\ & \quad 8a^{-2}\sum_{s_1=3}^{t-1}\sum_{s_2=2}^{s_1-1}\sum_{j_1=0}^{s_1-1}\sum_{j_2=0}^{s_2-1} k_{j_1m}^2 k_{j_2m}^2 \eta_{t,t-j_1}\eta_{s_1,s_1-j_1}\eta_{t,t-j_2}\eta_{s_2,s_2-j_2} \end{aligned}$$

and $E(W_{1nt}^2) = 4a^{-2}\sum_{s=2}^{t-1} E[(\sum_{j=0}^{s-1} k_{jm}^2\eta_{t,t-j}\eta_{s,s-j})^2]$, since

$$E(\eta_{t,t-j_1}\eta_{s_1,s_1-j_1}\eta_{t,t-j_2}\eta_{s_2,s_2-j_2}) = 0,\ s_2 < s_1, s_1 < t, s_2 < t.$$

Developing the square sum in $E(W_{1nt}^2)$, we have that

$$(\sum_{j=0}^{s-1} k_{jm}^2\eta_{t,t-j}\eta_{s,s-j})^2 = \sum_{j=0}^{s-1} k_{jm}^4\eta_{t,t-j}^2\eta_{s,s-j}^2 + 2\sum_{j_1=1}^{s-1}\sum_{j_2=0}^{j_1-1} k_{j_1m}^2 k_{j_2m}^2 \eta_{t,t-j_1}\eta_{s,s-j_1}\eta_{t,t-j_2}\eta_{s,s-j_2},$$

and then $E(W_{1nt}^2) = 4\sum_{s=2}^{t-1}\sum_{j=0}^{s-1} k_{jm}^4$, since

$$E(\eta_{t,t-j_1}\eta_{s,s-j_1}\eta_{t,t-j_2}\eta_{s,s-j_2}) = 0,\ t > s, j_1 > j_2,$$

and also $E(\eta_{t,t-j}^2\eta_{s,s-j}^2) = a^2$. We show in a similar way that $E(W_{2nt}^2) = 4\sum_{s=2}^{t-1}\sum_{j=1}^{s-1} k_{jm}^4$ and it implies that $E(W_{nt}^2) = 4\sum_{s=2}^{t-1}\sum_{|j|=0}^{s-1} k_{jm}^4$. By calculations analogous to those in Hong (1996b), we show that

$$\begin{aligned} \text{var}(W'_n) &= n^{-2}\sum_{t=3}^{n} E(W_{nt}^2) = 4n^{-2}\sum_{t=3}^{n}\sum_{s=2}^{t-1}\sum_{|j|=0}^{s-1} k_{jm}^4, \\ &= 2\sum_{|j|=0}^{n-2}(1-(|j|+1)/n)(1-|j|/n)k_{jm}^4 = 2V_n(k). \end{aligned}$$

and the proof of Proposition 2 is completed. Writing W_{1nt} and W_{2nt} in the following manner:

$$\begin{aligned} W_{1nt} &= 2a^{-1}\sum_{s=2}^{t-1} G_{t,s}^{(1)}, \\ W_{2nt} &= 2a^{-1}\sum_{s=2}^{t-1} G_{t,s}^{(2)}, \end{aligned}$$

where $G_{t,s}^{(1)} = \sum_{j=0}^{s-1} k_{jm}^2\eta_{t,t-j}\eta_{s,s-j}$, $t > s$ and $G_{t,s}^{(2)} = \sum_{j=1}^{s-1} k_{jm}^2\eta_{t-j,t}\eta_{s-j,s}$, $t > s$, we prove the next lemma.

Lemma 1 $E(G_{ts}^{(i)4}) = O(m^2),\ t > s,\ i = 1, 2.$

Proof:
We present the proof for $G_{ts}^{(1)}$, the other case being similar. We first develop $G_{ts}^{(1)4}$ and evaluating the expectation, we obtain that

$$\begin{aligned} E(G_{ts}^{(1)4}) &= \sum_{j=0}^{s-1} k_{jm}^8 E[(\eta_{t,t-j}\eta_{s,s-j})^4] + \\ &\quad \Delta \sum_{j_1<j_2} k_{j_1m}^4 k_{j_2m}^4 E[(\eta_{t,t-j_1}\eta_{s,s-j_1})^2(\eta_{t,t-j_2}\eta_{s,s-j_2})^2], \\ &\leq \Delta(\sum_{j=0}^{s-1} k_{jm}^8 + \sum_{j_1<j_2} k_{j_1m}^4 k_{j_2m}^4) \leq \Delta(\sum_{j=0}^{n-1} k_{jm}^4)^2 = O(m^2). \end{aligned}$$

In the first equality, the other terms in the development of $G_{ts}^{(1)4}$ have 0 expectation; this property is obtained by conditioning on the $\{u_t\}$ process. The second inequality follows from the fact that the moments in that expression are finite. With the preceding lemma, we can show the following proposition.

Proposition 3 $E(W_{int}^4) = O(t^2m^2)$, $i = 1, 2$.

Proof:
We present the proof for W_{1nt}, the one for W_{2nt} being similar. Developing W_{1nt}^4 and evaluating the expectation, we obtain that

$$E(W_{1nt}^4) = 16a^{-4}\{\sum_{s=2}^{t-1} E(G_{ts}^{(1)4}) + \Delta \sum_{s_1<s_2} E(G_{ts_1}^{(1)2}G_{ts_2}^{(1)2})\}.$$

Since by conditioning on the $\{v_t\}$ process, it is easily seen that the other terms have 0 expectation. Then we obtain

$$\begin{aligned} E(W_{1nt}^4) &\leq \Delta \sum_{s_1=2}^{t-1}\sum_{s_2=2}^{t-1} E(G_{ts_1}^{(1)2}G_{ts_2}^{(1)2}) \leq \Delta \sum_{s_1=2}^{t-1}\sum_{s_2=2}^{t-1} [E(G_{ts_1}^{(1)4})E(G_{ts_2}^{(1)4})]^{1/2}, \\ &\leq \Delta\{\sum_{s=2}^{t-1}[E(G_{ts}^{(1)4})]^{1/2}\}^2 = O(t^2m^2). \end{aligned}$$

This completes the proof of the proposition and condition (A1) follows since

$$\sigma^{-4}(n)n^{-4}\sum_{t=3}^{n} E(W_{nt}^4) = O(n^{-1}).$$

It remains to establish the validity of condition (A2). First note that $\ddot{W}_{nt}^2 = \ddot{W}_{1nt}^2 + \ddot{W}_{2nt}^2$, where we let $\ddot{W}_{1nt}^2 = 4a^{-2}E[(\sum_{s=2}^{t-1} G_{ts}^{(1)})^2|\mathcal{F}_{t-1}]$ and $\ddot{W}_{2nt}^2 = 4a^{-2}E[(\sum_{s=2}^{t-1} G_{ts}^{(2)})^2|\mathcal{F}_{t-1}]$. Again, we give the details for $\ddot{W}_{1nt}^2$; the development for $\ddot{W}_{2nt}^2$ is similar. First, $\ddot{W}_{1nt}^2$ can be developed in the following way:

$$\ddot{W}_{1nt}^2 = E(\ddot{W}_{1nt}^2) + 4a^{-2}C_{1nt} + 4a^{-2}C_{2nt} + 4a^{-2}C_{3nt},$$

where

$$\begin{aligned} E(\ddot{W}_{1nt}^2) &= 4\sum_{s=2}^{t-1}\sum_{j=0}^{s-1} k_{jm}^4, \quad C_{1nt} = \sum_{s=2}^{t-1}\sum_{j=0}^{s-1} k_{jm}^4 E[\eta_{s,s-j}^2\eta_{t,t-j}^2 - a^2|\mathcal{F}_{t-1}], \\ C_{2nt} &= 2\sum_{s=2}^{t-1}\sum_{j_1=1}^{s-1}\sum_{j_2=0}^{j_1-1} k_{j_1m}^2 k_{j_2m}^2 E[\eta_{t,t-j_1}\eta_{t,t-j_2}\eta_{s,s-j_1}\eta_{s,s-j_2}|\mathcal{F}_{t-1}], \\ C_{3nt} &= 2\sum_{s_1=3}^{t-1}\sum_{s_2=2}^{s_1-1} E[G_{ts_1}^{(1)}G_{ts_2}^{(1)}|\mathcal{F}_{t-1}]. \end{aligned}$$

The following lemma will be useful in the sequel, and is a generalization of a result of Hong (p.4, 1996b).

Lemma 2 *For $t_2 > s_2$, $t_1 > s_1$,*

$$|E[E(G^{(1)}_{t_1s_1}G^{(1)}_{t_1s_2}|\mathcal{F}_{t_1-1})E(G^{(1)}_{t_2s_1}G^{(1)}_{t_2s_2}|\mathcal{F}_{t_2-1})]| \leq \begin{cases} \Delta m^2(m^{-1}\sum_{j=0}^{n-1}k^2_{jm})^2, & \textit{if } t_1 = t_2, \\ \Delta m(m^{-1}\sum_{j=0}^{n-1}k^2_{jm})^2, & \textit{if } t_1 < t_2. \end{cases}$$

Proof:
To prove the result, we proceed as in Hong (p.4, 1996b) using the Assumptions on the function η and the fact that $\{u_t\}$ and $\{v_t\}$ are independent with a symmetric distribution. The details are omitted. The terms C_{1nt} and C_{2nt} have the following properties.

Proposition 4 $E(C^2_{1nt}) = O(t^2m + tm^2)$, $E(C^2_{2nt}) = O(tm^3)$.

Proof:
We first establish the result for C_{1nt}. Let us write $C_{1nt} = \sum_{s=2}^{t-1} C_{1nst}$, where

$$C_{1nst} = \sum_{j=0}^{s-1} E[\eta^2_{s,s-j}\eta^2_{t,t-j} - a^2|\mathcal{F}_{t-1}].$$

Developing the square sum, we obtain

$$C^2_{1nt} = \sum_{s=2}^{t-1} C^2_{1nst} + 2\sum_{s_1=3}^{t-1}\sum_{s_2=2}^{s_1-1} C_{1ns_1t}C_{1ns_2t}$$

and we study the two parts separately. For the first term, we easily show that $E(\sum_{s=2}^{t-1} C^2_{1nst}) = O(tm^2)$. The study of the second term allows us to conclude that it is of order $O(t^2m)$ and it follows that $E(C^2_{1nt}) = O(t^2m + tm^2)$. Now let us study C_{2nt}. Making changes in the indices of summation, C_{2nt} can be written in the following manner:

$$\begin{aligned} C_{2nt} &= 2\sum_{j_1+1}^{t-2}\sum_{j_2=j_1+1}^{t-2}\sum_{s=j_2+1}^{t-1} k^2_{j_1m}k^2_{j_2m}E[\eta_{t,t-j_1}\eta_{t,t-j_2}\eta_{s,s-j_1}\eta_{s,s-j_2}|\mathcal{F}_{t-1}], \\ &= 2\sum_{j_1+1}^{t-2} k^2_{j_1m}C_{2nj_1t}, \end{aligned}$$

where

$$C_{2nj_1t} = \sum_{l_1=j_1+1}^{t-2}\sum_{s=l_1+1}^{t-1} k^2_{l_1m}E[\eta_{t,t-j_1}\eta_{t,t-j_2}\eta_{s,s-j_1}\eta_{s,s-j_2}|\mathcal{F}_{t-1}].$$

Developing C^2_{2nt}, we obtain

$$C^2_{2nt} = 4\sum_{j_1=1}^{t-2} k^4_{j_1m}C^2_{2nj_1t} + 8\sum_{j_1=2}^{t-2}\sum_{j_2=1}^{j_1-1} k^2_{j_1m}k^2_{j_2m}C_{2nj_1t}C_{2nj_2t},$$

and we analyze the two terms separately. The study of the first term allows us to show that $E(\sum_{j_1=1}^{t-2} k^4_{j_1m}C^2_{2nj_1t}) = O(m^3t)$ and the expectation of the second term is zero. Therefore, we have $E(C^2_{2nt}) = O(m^3t)$. We can now establish (A2). We want to show that $m^{-2}\operatorname{var}(n^{-2}\sum_{t=3}^n \ddot{W}^2_{1nt}) \to 0$, and similarly for $\ddot{W}_{2nt}$. Since

$$\begin{aligned} m^{-2}\operatorname{var}(n^{-2}\sum_{t=3}^n \ddot{W}^2_{1nt}) &= m^{-2}E\{(n^{-2}\sum_{t=3}^n[\ddot{W}_{1nt} - E(\ddot{W}_{1nt})])^2\}, \\ &= m^{-2}E[(n^{-2}\sum_{i=1}^3\sum_{t=3}^n 4a^{-2}C_{int})^2], \end{aligned}$$

it is sufficient to show that

$$m^{-2}E[(n^{-2}\sum_{t=3}^{n}C_{int})^2]\to 0,$$

for $i=1,2,3$. Using Proposition 4, we have that

$$m^{-2}E[n^{-2}\sum_{t=3}^{n}C_{1nt}]^2\leq m^{-2}n^{-4}\{\sum_{t=3}^{n}E(C_{1nt}^2)^{1/2}\}^2=O(m^{-1}+n^{-1}).$$

Also, we obtain that

$$m^{-2}E[n^{-2}\sum_{t=3}^{n}C_{2nt}]^2\leq m^{-2}n^{-4}\{\sum_{t=3}^{n}E(C_{2nt}^2)^{1/2}\}^2=O(m/n)$$

and finally,

$$m^{-2}E[n^{-2}\sum_{t=3}^{n}C_{3nt}]^2=m^{-2}n^{-4}\sum_{t=3}^{n}E(C_{3nt}^2)+2m^{-2}n^{-4}\sum_{t_2=3}^{n}\sum_{t_1=2}^{t_2-1}E(C_{3nt_1}C_{3nt_2}).$$

But we verify that

$$E(C_{3nt_1}C_{3nt_2})=4\sum_{s_2=3}^{t_1-1}\sum_{s_1=2}^{s_2-2}E[E(G_{t_1s_1}G_{t_1s_2}|\mathcal{F}_{t_1-1})E(G_{t_2s_1}G_{t_2s_2}|\mathcal{F}_{t_2-1})].$$

Thus $E(C_{3nt}^2)\leq\Delta m^2t^2$ and $E(C_{3nt_1}C_{3nt_2})\leq\Delta mt_1^2$ by Lemma 2. Therefore, $m^{-2}E[n^{-2}\sum_{t=3}^{n}C_{3nt}]^2=O(n^{-1}+m^{-1})$, since $m/n\to 0$, $m\to\infty$, and condition (A2) is valid. This completes the proof of step 1.

Part 2 *We have to prove that $\tilde{Q}_{Rn}-Q_{Rn}=o_p(1)$. Then we have to show that*

$$\sum_{j=1-n}^{n-1}k_{jm}^2[\gamma_{uv}^2(j;\eta)-\gamma_{\hat{u}\hat{v}}^2(j;\eta)]=o_p(\sqrt{m}/n),$$

since $V_n(k)=O(m)$.

Using the identity $a^2-b^2=(a-b)^2+2b(a-b)$, we can write $\gamma_{\hat{u}\hat{v}}^2(j;\eta)-\gamma_{uv}^2(j;\eta)=[\gamma_{\hat{u}\hat{v}}(j;\eta)-\gamma_{uv}(j;\eta)]^2+2\gamma_{uv}(j;\eta)[\gamma_{\hat{u}\hat{v}}(j;\eta)-\gamma_{uv}(j;\eta)]$ and it is enough to show the two following results:

$$(B1)\qquad\sum_{j=1-n}^{n-1}k_{jm}^2[\gamma_{\hat{u}\hat{v}}(j;\eta)-\gamma_{uv}(j;\eta)]^2=o_p(\sqrt{m}/n),$$

$$(B2)\qquad\sum_{j=1-n}^{n-1}k_{jm}^2\gamma_{uv}(j;\eta)[\gamma_{\hat{u}\hat{v}}(j;\eta)-\gamma_{uv}(j;\eta)]=o_p(\sqrt{m}/n).$$

We write $\sum_{j=1-n}^{n-1}=\sum_{j<0}+\sum_{j\geq 0}$ and we present the details for $j\geq 0$, the other case being similar.

In the sequel, we make use of Taylor expansions of the residuals $\hat{u}_t$ and $\hat{v}_t$. Such expansions are given in Box and Pierce (1970) for univariate ARMA processes and in Hosking (1980) for the multivariate case.

Proposition 5

$$\begin{aligned}\hat{u}_t &= u_t+\sum_{r=1}^{p_1}\sum_{i=0}^{\infty}\Pi_{1i}(\hat{\phi}_{1r}-\phi_{1r})X_{t-r-i}-\sum_{r=1}^{q_1}\sum_{i=0}^{\infty}\Pi_{1i}(\hat{\theta}_{1r}-\theta_{1r})u_{t-r-i}+\lambda_{unt},\\ \hat{v}_t &= v_t+\sum_{r=1}^{p_2}\sum_{i=0}^{\infty}\Pi_{2i}(\hat{\phi}_{2r}-\phi_{2r})Y_{t-r-i}-\sum_{r=1}^{q_2}\sum_{i=0}^{\infty}\Pi_{2i}(\hat{\theta}_{2r}-\theta_{2r})v_{t-r-i}+\lambda_{vnt}.\end{aligned}$$

For a proof, see for example Lemma 1 of Hosking (1980). Remark that Hosking (1980) writes $\lambda_{unt} = O_p(n^{-1})$. Here, we prefer to retain the eventual dependence with respect to t of the remainders and we write λ_{unt} and λ_{vnt}. However, from Hosking's proof, we can conclude that $n^{-1}\sum_{t=1}^{n}\lambda_{unt} = O_p(n^{-1})$ and $n^{-1}\sum_{t=1}^{n}\lambda_{vnt} = O_p(n^{-1})$. We also have that $n^{-1}\sum_{t=1}^{n}\lambda_{unt}^2 = O_p(n^{-2})$ and $n^{-1}\sum_{t=1}^{n}\lambda_{vnt}^2 = O_p(n^{-2})$. The following lemma which gives us results on the orders in probability of the differences $\delta_{ut} = u_t - \hat{u}_t$ and $\delta_{vt} = v_t - \hat{v}_t$ directly follows from Proposition 5.

Lemma 3 *$n^{-1}\sum_{t=1}^{n}\delta_{ut}^2 = O_p(n^{-1})$, $n^{-1}\sum_{t=1}^{n}\delta_{vt}^2 = O_p(n^{-1})$.*

To show (B1) and (B2) we need results on $\gamma_{\hat{u}\hat{v}}(j;\eta) - \gamma_{uv}(j;\eta)$ and a Taylor series expansion of the η function will be used.

Proposition 6 *Let $\eta_1(u,v) = \partial\eta(u,v)/\partial u$ and $\eta_2(u,v) = \partial\eta(u,v)/\partial v$. Then we have*

$$\begin{aligned}\gamma_{\hat{u}\hat{v}}(j;\eta) - \gamma_{uv}(j;\eta) &= -n^{-1}\sigma_u^{-1}\sum_{t=j+1}^{n}\eta_1(u_t/\sigma_u, v_{t-j}/\sigma_v)\delta_{ut} - \\ &\quad n^{-1}\sigma_v^{-1}\sum_{t=j+1}^{n}\eta_2(u_t/\sigma_u, v_{t-j}/\sigma_v)\delta_{v,t-j} + R_{nt},\end{aligned}$$

where the remainder R_{nt} is such that $R_{nt} = O_p(n^{-1})$.

Proof:
This result is established in Li and Hui (p. 106, 1994) for the autoregressive case and can be easily generalized to the ARMA case. It is sufficient to write a Taylor expansion of η, to sum on t and to divide by n. We also need that $\sigma_u - \hat{\sigma}_u = O_p(n^{-1/2})$ and $\sigma_v - \hat{\sigma}_v = O_p(n^{-1/2})$.

We now derive (B1). Using the inequality $(a_1+a_2+a_3)^2 \leq 4[a_1^2+a_2^2+a_3^2]$ and Proposition 6, we have that

$$\sum_{j=0}^{n-1} k_{jm}^2[\gamma_{\hat{u}\hat{v}}(j;\eta) - \gamma_{uv}(j;\eta)]^2 \leq 4(\sigma_u^{-2}T_{1n} + \sigma_v^{-2}T_{2n} + T_{3n}),$$

where

$$\begin{aligned}T_{1n} &= \sum_{j=0}^{n-1} k_{jm}^2[n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}\delta_{ut}]^2, T_{2n} = \sum_{j=0}^{n-1} k_{jm}^2[n^{-1}\sum_{t=j+1}^{n}\eta_{2t,t-j}\delta_{v,t-j}]^2, \\ T_{3n} &= \sum_{j=0}^{n-1} k_{jm}^2 R_{nt}^2,\end{aligned}$$

with $\eta_{1t,t-j} = \eta_1(u_t/\sigma_u, v_{t-j}/\sigma_v)$ and $\eta_{2t,t-j} = \eta_2(u_t/\sigma_u, v_{t-j}/\sigma_v)$.

We begin with T_{1n}. Replacing δ_{ut} by the Taylor expansion given in Proposition 5, we can write

$$\begin{aligned}n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}\delta_{ut} &= -\sum_{r=1}^{p_1}(\hat{\phi}_{1r} - \phi_{1r})\{\sum_{i\geq 0}\Pi_{1i}[n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}X_{t-r-i}]\} + \\ &\quad \sum_{r=1}^{q_1}(\hat{\theta}_{1r} - \theta_{1r})\{\sum_{i\geq 0}\Pi_{1i}[n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}u_{t-r-i}]\} - n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}\lambda_{unt}.\end{aligned} \quad (23)$$

We can majorate $(n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}\delta_{ut})^2$ in the following way:

$$\begin{aligned}(n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}\delta_{ut})^2 &\leq \Delta[\sum_{r=1}^{p_1}(\hat{\phi}_{1r} - \phi_{1r})^2 A_{rjn} + \\ &\quad \sum_{r=1}^{q_1}(\hat{\theta}_{1r} - \theta_{1r})^2 B_{rjn} + (n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}\lambda_{unt})^2],\end{aligned}$$

where

$$A_{rjn} = (\sum_{i\geq 0}\Pi_{1i}[n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}X_{t-r-i}])^2 = \{n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}(\Pi_1(B)X_{t-r})\}^2,$$
$$B_{rjn} = (\sum_{i\geq 0}\Pi_{1i}[n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}u_{t-r-i}])^2 = \{n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}(\Pi_1(B)u_{t-r})\}^2.$$

Thus we have that

$$\begin{aligned} T_{1n} \leq\ & \Delta[\sum_{r=1}^{p_1}(\hat{\phi}_{1r}-\phi_{1r})^2(\sum_{j=0}^{n-1}k_{jm}^2A_{rjn}) + \\ & \sum_{r=1}^{q_1}(\hat{\theta}_{1r}-\theta_{1r})^2(\sum_{j=0}^{n-1}k_{jm}^2B_{rjn}) + \sum_{j=0}^{n-1}k_{jm}^2(n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}\lambda_{unt})^2]. \end{aligned}$$

We show that $\sum_{j=0}^{n-1}k_{jm}^2A_{rjn} = O_p(m/n)$ and that $\sum_{j=0}^{n-1}k_{jm}^2B_{rjn} = O_p(m/n)$. Indeed, we have that

$$\begin{aligned} \{n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}(\Pi_1(B)u_{t-r})\}^2 =\ & n^{-2}\sum_{t=j+1}^{n}\eta_{1t,t-j}^2(\Pi_1(B)X_{t-r})^2 + \\ & 2n^{-2}\sum_{t=j+2}^{n}\sum_{s=j+1}^{t-1}\eta_{1t,t-j}\eta_{1s,s-j}(\Pi_1(B)X_{t-r})(\Pi_1(B)X_{s-r}). \end{aligned}$$

Since $E[\eta_{1t,t-j}\eta_{1s,s-j}(\Pi_1(B)X_{t-r})(\Pi_1(B)X_{s-r})] = 0$, $t > s$, and that $E[\eta_{1t,t-j}^2(\Pi_1(B)X_{t-r})^2]$ is finite, it follows that $\sum_{j=0}^{n-1}k_{jm}^2A_{rjn} = O_p(m/n)$. The development for $\sum_{j=0}^{n-1}k_{jm}^2B_{rjn}$ is similar. We also have that

$$(n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}\lambda_{unt})^2 \leq (n^{-1}\sum_{t=1}^{n}\eta_{1t,t-j}^2)(n^{-1}\sum_{t=1}^{n}\lambda_{unt}^2).$$

This shows that $T_{1n} = O_p(m/n^2) = o_p(\sqrt{m}/n)$, since the estimators are such that $\hat{\phi}_{1r}-\phi_{1r} = O_p(n^{-1/2})$ and $\hat{\theta}_{1r}-\theta_{1r} = O_p(n^{-1/2})$, that $n^{-1}\sum_{t=1}^{n}\lambda_{unt}^2 = O_p(n^{-2})$, and since $m/n \to 0$. This show the result for T_{1n}. The proof for T_{2n} is similar. Also, we have that $T_{3n} = O_p(m/n^2)$ since $R_{nt}^2 = O_p(n^{-2})$ and condition (B1) is verified.

To establish (B2), let us remark that we can write

$$\sum_{j=0}^{n-1}k_{jm}^2\gamma_{uv}(j;\eta)[\gamma_{\hat{u}\hat{v}}(j;\eta)-\gamma_{uv}(j;\eta)] = -\sigma_u^{-1}T_{4n} - \sigma_v^{-1}T_{5n} + T_{6n},$$

where

$$T_{4n} = \sum_{j=0}^{n-1}k_{jm}^2\gamma_{uv}(j;\eta)[n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}\delta_{ut}],\ T_{5n} = \sum_{j=0}^{n-1}k_{jm}^2\gamma_{uv}(j;\eta)[n^{-1}\sum_{t=j+1}^{n}\eta_{2t,t-j}\delta_{v,t-j}],$$
$$T_{6n} = \sum_{j=0}^{n-1}k_{jm}^2\gamma_{uv}(j;\eta)R_{nt}.$$

Using one again (23), we can write T_{4n} in the following manner:

$$\begin{aligned} T_{4n} =\ & -\sum_{r=1}^{p_1}(\hat{\phi}_{1r}-\phi_{1r})C_{rn} + \sum_{r=1}^{q_1}(\hat{\theta}_{1r}-\theta_{1r})D_{rn} \\ & -\sum_{j=0}^{n-1}k_{jm}^2\gamma_{uv}(j;\eta)[n^{-1}\sum_{t=j+1}^{n}\eta_{1t,t-j}\lambda_{unt}], \end{aligned}$$

where

$$\begin{aligned}
C_{rn} &= \sum_{j=0}^{n-1} k_{jm}^2 \gamma_{uv}(j;\eta)[n^{-1} \sum_{t=j+1}^{n} \eta_{1t,t-j}(\Pi_1(B)X_{t-r})], \\
D_{rn} &= \sum_{j=0}^{n-1} k_{jm}^2 \gamma_{uv}(j;\eta)[n^{-1} \sum_{t=j+1}^{n} \eta_{1t,t-j}(\Pi_1(B)u_{t-r})].
\end{aligned}$$

Using Minkovski's inequality, we have that

$$\begin{aligned}
E|C_{rn}| &\leq \sum_{j=0}^{n-1} k_{jm}^2 E|\gamma_{uv}(j;\eta)[n^{-1} \sum_{t=j+1}^{n} \eta_{1t,t-j}(\Pi_1(B)X_{t-r})]| \\
&\leq \sum_{j=0}^{n-1} k_{jm}^2 [E(\gamma_{uv}^2(j;\eta))]^{1/2} [E(n^{-1} \sum_{t=j+1}^{n} \eta_{1t,t-j}(\Pi_1(B)X_{t-r}))^2]^{1/2}.
\end{aligned}$$

We study the square of $n^{-1}\sum_{t=j+1}^{n} \eta_{1t,t-j}(\Pi_1(B)X_{t-r})$. We have that

$$\begin{aligned}
(n^{-1} \sum_{t=j+1}^{n} \eta_{1t,t-j}(\Pi_1(B)X_{t-r}))^2 &= n^{-2} \sum_{t=j+1}^{n} \eta_{1t,t-j}^2 (\Pi_1(B)X_{t-r})^2 + \\
&\quad 2n^{-2} \sum_{t=j+1}^{n} \sum_{s=j+2}^{t-1} (\eta_{1s,s-j}\eta_{1t,t-j})(\{\Pi_1(B)X_{s-r}\}\{\Pi_1(B)X_{t-r}\}).
\end{aligned}$$

Taking expectation, we show that $E[n^{-1}\sum_{t=j+1}^{n} \eta_{1t,t-j}(\Pi_1(B)X_{t-r})]^2 = O(n^{-1})$ uniformly in j. We proceed in a similar way for the term involving D_{rn}. It is easily shown that

$$\sum_{j=0}^{n-1} k_{jm}^2 \gamma_{uv}(j;\eta)[n^{-1} \sum_{t=j+1}^{n} \eta_{1t,t-j}\lambda_{unt}] = o_p(\sqrt{m}/n).$$

and thus $T_{4n} = O_p(m/n^{3/2}) = o_p(\sqrt{m}/n)$ since $m/n \to 0$. The order in probability of T_{5n} is obtained in a similar way and the proof of part 2 is completed.

REFERENCES

Akaike, H. and Kitagawa, G. (1999), *The Practice of Time Series Analysis*, Springer:Berlin.

Ansley, C. F. (1980), 'Computation of the theoretical autocovariance function for a vector ARMA process', *Journal of Statistical Computation and Simulation* **12**, 15–24.

Beaton, A. E. and Tukey, J. W. (1974), 'The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data', *Technometrics* **16**, 147–185.

Ben, M. G., Martinez, E. J. and Yohai, V. J. (1999), 'Robust estimation in vector autoregressive moving-average models', *Journal of Time Series Analysis* **20**, 381–399.

Boudjellaba, H., Dufour, J.-M. and Roy, R. (1992), 'Testing causality between two vectors in multivariate autoregressive moving average models', *Journal of American Statistical Association* **87**, 1082–1090.

Boudjellaba, H., Dufour, J.-M. and Roy, R. (1994), 'Simplified conditions for noncausality between vectors in multivariate ARMA models', *Journal of Econometrics* **63**, 271–287.

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1994), *Time Series Analysis. Forecasting and Control*, third edn, Prentice Hall, Inc., Englewood Cliffs, NJ.

Box, G. E. P. and Pierce, D. A. (1970), 'Distribution of residual autocorrelations in autoregressive-integrated moving average time series models', *Journal of American Statistical Association* **65**, 1509–1526.

Brockwell, P. J. and Davis, R. A. (1996), *Introduction to Time Series and Forecasting*, Springer:Berlin.

Brown, B. M. (1971), 'Martingale central limit theorems', *Annals of Mathematical Statistics* **42**, 59–66.

Bustos, O. H. and Yohai, V. J. (1986), 'Robust estimates for ARMA models', *Journal of American Statistical Association* **81**, 155–168.

Bustos, O., Fraiman, R. and Yohai, V. J. (1984), 'Asymptotic Behaviour of the Estimates Based on Residual Autocovariances for ARMA Models', in *Robust and Nonlinear Time Series Analysis*, edited by Franke, J., Haerdle, W. and Martin, D., Springer:Berlin, 26–49.

Derby, L. and Martin, R. D. (1979), 'Robust estimation of the first-order autoregressive parameter', *Journal of American Statistical Association* **74**, 140–146.

Duchesne, P. (2000), 'Quelques contributions en théorie de l'échantillonnage et dans l'analyse des séries chronologiques multidimensionnelles', Thèse de doctorat, département de mathématiques et statistique, Université de Montréal.

El Himdi, K. and Roy, R. (1997), 'Tests for noncorrelation of two multivariate ARMA time series', *The Canadian Journal of Statistics* **25**, 233–256.

Fox, A. J. (1972), 'Outliers in time series', *Journal of the Royal Statistical Society, B* **34**, 350–363.

Fuller, W. A. (1996), *Introduction to Statistical Time Series*, second edn, Wiley:New York.

Granger, C. W. J. (1969), 'Investigating causal relations by econometric models and cross-spectral methods', *Econometrica* **37**, 424–438.

Hallin, M., Jurevcková, J., Picek, J. and Zahaf, T. (1999), 'Nonparametric tests of independence of two autoregressive time series based on autoregression rank scores', *Journal of Statistical Planning and Inference* **75**, 319–330.

Hallin, M. and Saidi, A. (2001), 'Testing independence and causality between multivariate ARMA time series', Working paper, ISRO, Université Libre de Bruxelles.

Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press:Princeton.

Hampel, F.R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986), *Robust Statistics: The Approach Based on Influence Functions*, Wiley:New York.

Haugh, L. D. (1976), 'Checking the independence of two covariance-stationary time series: a univariate residual cross-correlation approach', *Journal of American Statistical Association* **71**, 378–385.

Hong, Y. (1996*a*), 'Testing for independence between two covariance stationary time series', *Biometrika* **83**, 615–625.

Hong, Y. (1996*b*), 'A separate mathematical appendix for 'Testing for independence between two covariance stationary time series'', Mimeo, Department of Economics and Department of Statistical Sciences, Cornell University.

Hong, Y. (1996*c*), 'Consistent testing for serial correlation of unknown form', *Econometrica* **64**, 837–864.

Hosking, J. (1980), 'The multivariate portmanteau statistic', *Journal of American Statistical Association* **75**, 602–608.

Judge, G. G., Hill, R. C., Griffiths, W. E., Lütkepohl, H. and Lee, T.-C. (1985), *The Theory and Practice of Econometrics*, second edn, Wiley:New York.

Koch, P. D. and Yang, S. S. (1986), 'A method for testing the independence of two time series that accounts for a potential pattern in the cross-correlation function', *Journal of the American Statistical Association*, **81**, 533–544.

Li, W. K. (1988), 'A goodness-of-fit test in robust time series modelling', *Biometrika* **75**, 355–361.

Li, W. K. and Hui, Y. V. (1994), 'Robust residual cross correlation tests for lagged relations in time series', *Journal Statistical Computation and Simulation* **49**, 103–109.

Lütkepohl, H. (1993), *Introduction to Multiple Time Series Analysis*, second edn, Springer:Berlin.

Martin, R. D. and Yohai, V. J. (1985), 'Robustness in time series and estimating ARMA models', in *Handbook of Statistics, 5, Time Series in the Time Domain*, eds. Hannan, E.J, Krishnaiah, P.R. and Rao, M.M., North-Holland, Amsterdam, 119–155.

McLeod, A. I. (1979), 'Distribution of the residual cross-correlation in univariate ARMA time series models', *Journal of American Statistical Association* **74**, 849–855.

Pham, D.T., Roy, R. and Cédras, L. (2001), 'Tests for non-correlation of two cointegrated ARMA time series'. To appear in *The Journal of Time Series Analysis*.

Pierce, D. A. (1977), 'Relationships – and the lack thereof – between economic time series, with special reference to money and interest rates', *Journal of the American Statistical Association* **72**, 11–22. (C/R: p22-26).

Pierce, D. A. and Haugh, L. D. (1977), 'Causality in temporal systems: characterizations and survey', *Journal of Econometrics* **5**, 265–293.

Priestley, M. B. (1981), *Spectral Analysis and Time Series: Univariate Series*, Vol. 1, Academic Press:London.

Rousseeuw, P. J. and Leroy, A. M. (1987), *Robust Regression and Outlier Detection*, Wiley:New York.

Taniguchi, M. and Kakizawa, Y. (2000), 'Asymptotic Theory of Statistical Inference for Time Series', Springer:Berlin.