2003s-41

On the Distribution of the Residual Cross-Correlations between Two Uncorrelated Infinite Order Vector Autoregressive Series

Chafik Bouhaddioui, Roch Roy

Série Scientifique
Scientific Series

Montréal
Juin 2003

CIRANO
Centre interuniversitaire de recherche en analyse des organisations

CIRANO

ISSN 1198-8177

On the Distribution of the Residual Cross-Correlations between Two Uncorrelated Infinite Order Vector Autoregressive Series *

Chafik Bouhaddioui,[†] *Roch Roy*[‡]

Résumé / *Abstract*

Dans ce travail, nous obtenons la loi asymptotique d'un vecteur quelconque de corrélations croisées résiduelles résultant de l'ajustement d'autorégressions finies à deux séries non corrélées décrites par des processus autorégressifs multivariés d'ordre infini. La loi asymptotique est la même que celle du vecteur correspondant de corrélations croisées entre les deux séries d'innovations correspondantes qui est une loi multinormale. Nous discutons aussi l'application de ce résultat afin de tester l'hypothèse de non corrélation de deux séries multivariées.

Mots clés : Autorégression finie, corrélations croisées résiduelles, distribution asymptotique, tests d'indépendance, statistique Portmanteau.

Here we derive the asymptotic distribution of an arbitrary vector of residual cross-correlations resulting from the fitting of finite autoregressions to two uncorrelated infinite order vector autoregressive series. Its asymptotic distribution is the same multivariate normal as the one of the corresponding vector of cross-correlations between the two innovation series. The application of that result for testing the uncorrelatedness of two series is also discussed.

Keywords: *Finite autoregression; residual cross-correlations; asymptotic distribution; tests for independence; portmanteau statistics.*

Mathematics subject classification (2000): Primary 62M10; secondary 62M15.

---

* This work was supported by grants to the second named author from the Natural Science and Engineering Research Council of Canada, the Network of Centers of Excellence on the Mathematics of Information Technology and Complex Systems (MITACS) and the Fonds FQRNT (Government of Quebec).

[†] CIRANO and Département de mathématiques et de statistique, Université de Montréal, e-mail: Chafik.Bouhaddioui@cirano.qc.ca.

‡ CRM and Département de mathématiques et de statistique, Université de Montréal, Case postale 6128; succursale Centre-ville, Montréeal, Québec, Canada, H3C 3J7, e-mail: royr@dms.umontreal.ca.

1 Introduction

In the context of multivariate time series, many recent papers address the problem of testing independence or non-correlation between two observed series. El Himdi and Roy (1997) generalized the procedure developed by Haugh (1976) for univariate time series to the case of two multivariate invertible and stationary vector autoregressive moving average (VARMA) series. They proposed a test statistic based on the residual cross-correlation matrices $\mathbf{R}_{\hat{\boldsymbol{a}}}^{(12)}(j)$, $|j| \leq M$ for a given $M < N$, (N being the sample size), between the two residual series $\{\hat{\boldsymbol{a}}_t^{(1)}\}$ and $\{\hat{\boldsymbol{a}}_t^{(2)}\}$ resulting from fitting the *true* VARMA models to each of the original series $\{\mathbf{X}_t^{(1)}\}$ and $\{\mathbf{X}_t^{(2)}\}$. Under the hypothesis of non-correlation between the two series, they showed in particular that an arbitrary vector of residual cross-correlations asymptotically follows the same multivariate normal distribution as the corresponding vector of cross-correlations between the two innovation series. Hallin and Saidi (2002) used that result to develop a test statistic that takes into account a possible pattern in the signs of cross-correlations at different lags. They generalized to the multivariate case the procedure introduced by Koch and Yang (1986). Also, in the univariate context, Hong (1996) proposed a modification of Haugh's procedure for stationary infinite order autoregressive (AR(∞)) series. For a given sample size N, the true AR(∞) model of each series is approximated by a finite order AR(p) model. To derive the statistical properties of the test statistics, it is assumed that the orders of the autoregressions are functions of N that tend to infinity as N increases. In practice, such an approach protects us against misspecifications of the true underlying ARMA models that may lead to misleading conclusions because they invalidate the asymptotic theory of the test statistic. Hong's test statistic is a standardized version of the sum of weighted squared residual cross-correlations at all possible lags. Bouhaddioui (2002) generalized that procedure to the multivariate case.

In this work, we adopt the semiparametric approach taken by Hong for testing non-correlation between two AR(∞) series. We focus on the asymptotic properties of an arbitrary finite vector of residual cross-correlations.

The main contribution of this paper is to show that in the case of two uncorrelated infinite order vector autoregressive VAR(∞) time series, an arbitrary vector of residual cross-correlations asymptotically follows the same multivariate normal distribution as the corresponding vector of cross-correlations between the two innovation series. That result allows us to test the null hypothesis of non-correlation against the alternative of serial cross-correlation at a particular lag as described in El Himdi and Roy (1997). In a sense, it complements the analysis that can be done with Hong's global statistic.

The organization of the paper is as follows. Section 2 contains preliminary results. The main result and its proof are presented in Section 3. In Section 4, we describe two test procedures for the null hypothesis of non-correlation. The first one is based, for a particular lag j, on the cross-correlation matrix $\mathbf{R}^{(12)}_{\hat{\boldsymbol{a}}}(j)$, whilst the second one is based on a global statistic that takes into account all possible lags from $-M$ to $+M$, say. The results of a small Monte Carlo experiment investigating the finite sample properties of the tests are also presented.

2 Preliminaries

Let $\mathbf{X} = \{\mathbf{X}_t \ , \ t \in \mathbb{Z}\}$ be a multivariate second-order stationary process of dimension m. Without loss of generality, we can assume that $\mathbb{E}(\mathbf{X}_t) = \mathbf{0}$. The autocovariance matrix at lag j, $j \in \mathbb{Z}$, is given by

$$\boldsymbol{\Gamma}_{\mathbf{X}}(j) \quad = \quad \mathbb{E}(\mathbf{X}_t \mathbf{X}_{t-j})^T = (\gamma_{uv}(j))_{m \times m},$$

with $\boldsymbol{\Gamma}_{\mathbf{X}}(j) = \boldsymbol{\Gamma}_{\mathbf{X}}(-j)^T$. The autocorrelation matrix at lag j, $j \in \mathbb{Z}$, is denoted by

$$\boldsymbol{\rho}_{\mathbf{X}}(j) = (\rho_{uv}(j))_{m \times m}, \qquad \rho_{uv}(j) = \gamma_{uv}(j)\{\gamma_{uu}(0)\gamma_{vv}(0)\}^{-1/2},$$

with $\boldsymbol{\rho}_{\mathbf{X}}(j) = \boldsymbol{\rho}_{\mathbf{X}}(-j)^T$. If we denote by $\mathbf{D}\{b_i\}$ a diagonal matrix whose elements are $b_1, ..., b_m$, the matrix $\boldsymbol{\rho}_{\mathbf{X}}(j)$ can be written in the following way:

$$\boldsymbol{\rho}_{\mathbf{X}}(j) \quad = \quad \mathbf{D}\{\gamma_{ii}(0)^{-1/2}\}\boldsymbol{\Gamma}_{\mathbf{X}}(j)\mathbf{D}\{\gamma_{ii}(0)^{-1/2}\} \ , \ j \in \mathbb{Z}. \tag{2.1}$$

Throughout the paper, $\|A\| = \{tr(A^T A)\}^{1/2}$ represents the Euclidean norm of the matrix A, $\otimes$ stands for the Kronecker product of matrices, the symbols $\xrightarrow{p}$, $\xrightarrow{L}$ are used for convergence in probability and in distribution respectively, and $\mathbb{I}_m$ denotes the identity matrix of dimension m.

In the sequel, we suppose that the stationary process $\mathbf{X}$ follows an infinite order autoregressive model, VAR(∞), i.e., there exists $m \times m$ matrices $\boldsymbol{\Phi}_l$, $l \in \mathbb{N}$, such that

$$\mathbf{X}_t - \sum_{l=1}^{\infty} \boldsymbol{\Phi}_l \mathbf{X}_{t-l} \quad = \quad \boldsymbol{a}_t, \ t \in \mathbb{Z}, \tag{2.2}$$

where $\sum_{l=1}^{\infty} \|\boldsymbol{\Phi}_l\| < \infty$, $\boldsymbol{\Phi}(z) = \boldsymbol{I}_m - \sum_{l=1}^{\infty} \boldsymbol{\Phi}_l z^l$ and $\det\{\boldsymbol{\Phi}(z)\} \neq 0$, $|z| \leq 1$. The process $\boldsymbol{a} = \{\boldsymbol{a}_t; t \in \mathbb{Z}\}$ is a strong white noise that is, a sequence of independent identically distributed random vectors with mean $\mathbf{0}$ and regular covariance matrix $\boldsymbol{\Sigma}$. The stationarity assumption ensures that the process $\mathbf{X}$ also admits a causal linear representation.

Based on a realization $\mathbf{X}_1, ..., \mathbf{X}_N$ of length N, we fit an autoregressive model of order p, VAR(p), whose coefficients are denoted by $\mathbf{\Phi}_{1,p}, ..., \mathbf{\Phi}_{p,p}$ and we write $\mathbf{\Phi}(p) = (\mathbf{\Phi}_{1,p}, ..., \mathbf{\Phi}_{p,p})$. The corresponding Yule-Walker estimator $\hat{\mathbf{\Phi}}(p) = (\hat{\mathbf{\Phi}}_{1,p}, ..., \hat{\mathbf{\Phi}}_{p,p})$ is given by

$$\hat{\mathbf{\Phi}}(p) \quad = \quad \hat{\mathbf{A}}_{1,p}^T \hat{\mathbf{A}}_p^{-1}, \tag{2.3}$$

where $\hat{\mathbf{A}}_{1,p} = (N-p)^{-1} \sum_{t=p+1}^N \mathbf{X}_t(p) \mathbf{X}_t^T$, $\hat{\mathbf{A}}_p = (N-p)^{-1} \sum_{t=p+1}^N \mathbf{X}_t(p) \mathbf{X}_t^T(p)$ and $\mathbf{X}_t(p) = (\mathbf{X}_{t-1}^T, \mathbf{X}_{t-2}^T, ..., \mathbf{X}_{t-p}^T)^T$. To obtain a consistent estimator $\hat{\mathbf{\Phi}}(p)$, we must let p tends to infinity as N increases but not too fast. The following assumption on the noise process is also needed.

Assumption A *The m-dimensional strong white noise $\boldsymbol{a} = \{\boldsymbol{a}_t = (a_{1t}, .., a_{mt})^T\}$ is such that $\mathbb{E}(\boldsymbol{a}) = \mathbf{0}$, its covariance matrix $\mathbf{\Sigma}$ is regular and*

$$\mathbb{E}|a_{i,t} a_{j,t} a_{k,t} a_{l,t}| \quad < \quad \gamma_4 < \infty \ , \ i, j, k, l \in \{1, ..., m\} \text{ and } t \in \mathbb{Z}.$$

The following proposition that gives the consistency rate of $\hat{\mathbf{\Phi}}(p)$ is a multivariate generalization of a univariate result presented in Berk (1974). It follows from Eq. (2.8) of Lewis and Reinsel (1985, p. 397), see also Theorem 2.1 in Paparoditis (1996).

Proposition 2.1 *Let $\{\mathbf{X}_t\}$ be a VAR(∞) process given by (2.2) and satisfying Assumption A. Also, suppose that the following two conditions are verified:*
(i) p is chosen as a function of N such that $p \to \infty$ and $p^2/N \to 0$ as $N \to \infty$;
(ii) $\sqrt{p} \sum_{j=p+1}^{\infty} \|\mathbf{\Phi}_j\| \to 0$ as $N \to \infty$.
Then, the estimator $\hat{\mathbf{\Phi}}(p)$ defined by (2.3) is such that

$$\|\hat{\mathbf{\Phi}}(p) - \mathbf{\Phi}(p)\| \quad = \quad O_p(\frac{p^{1/2}}{N^{1/2}}). \tag{2.4}$$

In this result, the condition $p = o(N^{1/2})$ for the rate of increase of p ensures that asymptotically, enough sample information is available for the estimators to have standard limiting distributions. The condition $\sqrt{p} \sum_{j=p+1}^{\infty} \|\mathbf{\Phi}_j\| \to 0$ imposes a lower bound on the growth rate of p, which ensures that the approximation error of the true underlying model by a finite order autoregression gets small when the sample size increases. A more detailed discussion of these conditions is available in Lütkepohl (1991), see also Hong (1996).

3 Asymptotic distribution of a vector of residual cross-correlations

From now, suppose that the process $\mathbf{X}$ is partitioned into two subprocesses $\mathbf{X}^{(h)} = \{\mathbf{X}_t^{(h)}, \ t \in \mathbb{Z}\}$, $h = 1, 2$, with m_1 and m_2 components respectively ($m_1 + m_2 = m$), that is $\mathbf{X}_t = (\mathbf{X}_t^{(1)^T}, \mathbf{X}_t^{(2)^T})^T$.

The partition of $\mathbf{X}_t$ induces the following partition of the autocovariance matrix $\mathbf{\Gamma_X}(j)$:

$$\mathbf{\Gamma_X}(j) \;=\; \begin{pmatrix} \mathbf{\Gamma}^{(11)}_{\mathbf{X}}(j) & \mathbf{\Gamma}^{(12)}_{\mathbf{X}}(j) \\ \mathbf{\Gamma}^{(21)}_{\mathbf{X}}(j) & \mathbf{\Gamma}^{(22)}_{\mathbf{X}}(j) \end{pmatrix}, \; j \in \mathbb{Z},$$

where $\mathbf{\Gamma}^{(hh)}_{\mathbf{X}}(j)$ is the autocovariance matrix at lag j of the process $\mathbf{X}^{(h)}$, $h = 1, 2$, and $\mathbf{\Gamma}^{(12)}_{\mathbf{X}}(j)$ is the cross-covariance matrix at lag j between $\{\mathbf{X}^{(1)}_t\}$ and $\{\mathbf{X}^{(2)}_t\}$ with $\mathbf{\Gamma}^{(21)}_{\mathbf{X}}(j) = \mathbf{\Gamma}^{(12)}_{\mathbf{X}}(-j)^T$. The autocorrelation matrix $\boldsymbol{\rho}_{\mathbf{X}}(j)$ is also partitioned in a similar way. Given a realization of length N of the process $\mathbf{X}$, the sample cross-covariance matrix at lag j is defined by

$$\boldsymbol{C}^{(12)}_{\mathbf{X}}(j) \;=\; N^{-1} \sum_{t=j+1}^{N} \left(\mathbf{X}^{(1)}_t - \overline{\mathbf{X}}^{(1)}\right) \left(\mathbf{X}^{(2)}_{t-j} - \overline{\mathbf{X}}^{(2)}\right)^T, \; 0 \le j \le N-1. \tag{3.1}$$

Also, for $-N+1 \le j \le 0$, $\boldsymbol{C}^{(12)}_{\mathbf{X}}(-j) = \boldsymbol{C}^{(21)}_{\mathbf{X}}(j)^T$ and $\boldsymbol{C}^{(12)}_{\mathbf{X}}(j) = 0$ for $|j| \ge N$. The sample cross-correlation matrix at lag j is given by

$$\mathbf{R}^{(12)}_{\mathbf{X}}(j) \;=\; \mathbf{D}_1 \boldsymbol{C}^{(12)}_{\mathbf{X}}(j) \mathbf{D}_2 \tag{3.2}$$

where D_h is a diagonal matrix whose elements are the square root of the elements on the main diagonal of $\boldsymbol{C}^{(hh)}_{\boldsymbol{X}}(0)$, the sample autocovariance at lag j of $\boldsymbol{X}$, $h = 1, 2$. We write $\boldsymbol{r}^{(12)}_{\mathbf{X}}(j) = vec(\mathbf{R}^{(12)}_{\mathbf{X}}(j))$ where the symbol *vec* stands for the usual operator that transforms a matrix into a vector by stacking its columns on top of each other.

In the sequel, we suppose that for $h = 1, 2$, $\mathbf{X}^{(h)}$ satisfy (2.2) and Assumption A and we want to test the null hypothesis that they are uncorrelated (or independent in the Gaussian case), that is, $\boldsymbol{\rho}^{(12)}_{\mathbf{X}}(j) = \mathbf{0}$, $j \in \mathbb{Z}$. As in El Himdi and Roy (1997), that hypothesis is equivalent to

$$\mathcal{H}_0 : \; \boldsymbol{\rho}^{(12)}_{\boldsymbol{a}}(j) \;=\; \mathbf{0}\,, \qquad j \in \mathbb{Z}. \tag{3.3}$$

Each series $\mathbf{X}^{(h)}_1, ..., \mathbf{X}^{(h)}_N$ is described by a finite-order autoregressive model VAR(p_h). The order p_h depends on N. The resulting residuals are given by

$$\hat{\boldsymbol{a}}^{(h)}_t = \begin{cases} \mathbf{X}^{(h)}_t - \sum_{l=1}^{p_h} \hat{\mathbf{\Phi}}_{l,p_h} \mathbf{X}^{(h)}_{t-l} & \text{if } t = p_h + 1, ..., N\,, \\ \mathbf{0} & \text{if } t \le p_h, \end{cases} \tag{3.4}$$

where the $\hat{\mathbf{\Phi}}_{l,p_h}$ are the Yule-Walker estimators defined by (2.3). The residual cross-covariance matrix $\boldsymbol{C}^{(12)}_{\hat{\mathbf{a}}}(j)$ is given by

$$\boldsymbol{C}^{(12)}_{\hat{\mathbf{a}}}(j) = \begin{cases} N^{-1} \sum_{t=j+1}^{N} \hat{\boldsymbol{a}}^{(1)}_t \hat{\boldsymbol{a}}^{(2)^T}_{t-j} & \text{if } 0 \le j \le N-1\,, \\ N^{-1} \sum_{t=-j+1}^{N} \hat{\boldsymbol{a}}^{(1)}_{t+j} \hat{\boldsymbol{a}}^{(2)^T}_t & \text{if } -N+1 \le j \le 0\,. \end{cases} \tag{3.5}$$

The residual cross-correlation matrix at lag j is given by $\mathbf{R}_{\hat{\boldsymbol{a}}}^{(12)}(j) = \mathbf{D}_1 \boldsymbol{C}_{\hat{\boldsymbol{a}}}^{(12)}(j)\mathbf{D}_2$ where D_1 and D_2 are defined as in (3.2).

Let $d_1, ..., d_L$ be a finite set of lags such that $|d_i| \leq D < N$ where D is independent of N. Denote by $\boldsymbol{r}_{\hat{\boldsymbol{a}}}^{(12)}$ the vector

$$\boldsymbol{r}_{\hat{\boldsymbol{a}}}^{(12)} = \left(vec(\mathbf{R}_{\hat{\boldsymbol{a}}}^{(12)}(d_1))^T, ..., vec(\mathbf{R}_{\hat{\boldsymbol{a}}}^{(12)}(d_L))^T \right)^T, \tag{3.6}$$

of dimension Lm_1m_2 and by $\boldsymbol{r}_{\boldsymbol{a}}^{(12)}$ the corresponding vector of cross-correlations between the two innovation series. Under some general assumptions, it follows from Roy (1989), see also El Himdi and Roy (1997), that

$$\sqrt{N}\boldsymbol{r}_{\boldsymbol{a}}^{(12)} \xrightarrow{L} N(\mathbf{0}, \mathbb{I}_L \otimes (\boldsymbol{\rho}_2 \otimes \boldsymbol{\rho}_1)) \tag{3.7}$$

where $\boldsymbol{\rho}_h$ is the correlation matrix of the innovation process $\boldsymbol{a}^{(h)}$, $h = 1, 2$. The asymptotic distribution of the vector $\boldsymbol{r}_{\hat{\boldsymbol{a}}}^{(12)}$ is provided by the following theorem.

Theorem 3.1 *Let $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ be two multivariate stationary processes which satisfy the multivariate infinite order autoregressive model* (2.2)*. Suppose that the corresponding innovation processes satisfy Assumption A and that all their fourth-order cumulants are zero. Let p_h, $h = 1, 2$, satisfy the following conditions*
(i) p_h, $h = 1, 2$, are chosen as a function of N such that $p_h \to \infty$ and $p_h^2/N \to 0$ as $N \to \infty$;
(ii) $\sqrt{N}\sum_{j=p_h+1}^{\infty} \|\boldsymbol{\Phi}_j^{(h)}\| \to 0$ as $N \to \infty$.
Then, under the null hypothesis of non-correlation between the two innovation processes $\boldsymbol{a}^{(1)}$ and $\boldsymbol{a}^{(2)}$,

$$\sqrt{N}\boldsymbol{r}_{\hat{\boldsymbol{a}}}^{(12)} \xrightarrow{L} N(\mathbf{0}, \mathbb{I}_L \otimes (\boldsymbol{\rho}_2 \otimes \boldsymbol{\rho}_1)).$$

PROOF

From (3.7), it suffices to prove that $\sqrt{N}\left(\boldsymbol{r}_{\hat{\boldsymbol{a}}}^{(12)} - \boldsymbol{r}_{\boldsymbol{a}}^{(12)}\right) \xrightarrow{p} \mathbf{0}$, or equivalently, that for any $j \in \{j_1, .., j_L\}$,

$$\sqrt{N}\left[\mathbf{R}_{\hat{\boldsymbol{a}}}^{(12)}(j) - \mathbf{R}_{\boldsymbol{a}}^{(12)}(j)\right] \xrightarrow{p} \mathbf{0}. \tag{3.8}$$

Using (3.2), we have that $\mathbf{R}_{\boldsymbol{a}}^{(12)}(j) = \mathbf{D}_1^{-1}\mathbf{C}_{\boldsymbol{a}}^{(12)}(j)\mathbf{D}_2^{-1}$, where $\mathbf{D}_h$ is a diagonal matrix whose elements are the square root of the elements on the main diagonal of $\boldsymbol{C}_{\boldsymbol{a}}^{(hh)}(0)$, $h = 1, 2$. Thus, we have

$$\mathbf{R}_{\hat{\boldsymbol{a}}}^{(12)}(j) - \mathbf{R}_{\boldsymbol{a}}^{(12)}(j) = \hat{\mathbf{D}}_1^{-1}\left[\mathbf{C}_{\hat{\boldsymbol{a}}}^{(12)}(j) - \mathbf{C}_{\boldsymbol{a}}^{(12)}(j)\right]\hat{\mathbf{D}}_2^{-1} + \hat{\mathbf{D}}_1^{-1}\mathbf{D}_1\mathbf{R}_{\boldsymbol{a}}^{(12)}(j)\mathbf{D}_2\hat{\mathbf{D}}_2^{-1} - \mathbf{R}_{\boldsymbol{a}}^{(12)}(j).$$

From Lütkepohl (1991, p. 309), we know that $\hat{\mathbf{D}}_h^{-1} - \mathbf{D}_h^{-1} = O_p(N^{-1/2})$, $h = 1, 2$, and therefore, $\hat{\mathbf{D}}_h^{-1}\mathbf{D}_h = \mathbb{I}_{m_h} + O_p(N^{-1/2})$. Also, from (3.7), $\sqrt{N}\mathbf{R}_{\boldsymbol{a}}^{(12)}(j)$ converges in distribution and it follows that

$$\sqrt{N}\left[\hat{\mathbf{D}}_1^{-1}\mathbf{D}_1\mathbf{R}_{\boldsymbol{a}}^{(12)}(j)\mathbf{D}_2\hat{\mathbf{D}}_2^{-1} - \mathbf{R}_{\boldsymbol{a}}^{(12)}(j)\right] \xrightarrow{p} \mathbf{0}.$$

Since $\hat{\mathbf{D}}_h^{-1} \xrightarrow{p} \mathbf{D}_h^{-1}$, $h = 1, 2$, to prove (3.8), it is sufficient to show that

$$\sqrt{N}\left(\mathbf{C}_{\hat{\boldsymbol{a}}}^{(12)}(j) - \mathbf{C}_{\boldsymbol{a}}^{(12)}(j)\right) \xrightarrow{p} \mathbf{0}. \tag{3.9}$$

If we note $\hat{\boldsymbol{\eta}}_t = \hat{\boldsymbol{a}}_t^{(1)} - \boldsymbol{a}_t^{(1)}$ and $\hat{\boldsymbol{\delta}}_t = \hat{\boldsymbol{a}}_t^{(2)} - \boldsymbol{a}_t^{(2)}$, we can write

$$\sqrt{N}\left(\mathbf{C}_{\hat{\boldsymbol{a}}}^{(12)}(j) - \mathbf{C}_{\boldsymbol{a}}^{(12)}(j)\right) \quad = \quad \mathcal{C}_1 + \mathcal{C}_2 + \mathcal{C}_3, \tag{3.10}$$

where $\mathcal{C}_1 = N^{-1/2}\sum_{t=j+1}^{N} \hat{\boldsymbol{\eta}}_t \boldsymbol{a}_{t-j}^{(2)^T}$, $\mathcal{C}_2 = N^{-1/2}\sum_{t=j+1}^{N} \boldsymbol{a}_t^{(1)} \hat{\boldsymbol{\delta}}_{t-j}^T$ and $\mathcal{C}_3 = N^{-1/2}\sum_{t=j+1}^{N} \hat{\boldsymbol{\eta}}_t \hat{\boldsymbol{\delta}}_{t-j}^T$ and it remains to show that $\mathcal{C}_i \xrightarrow{p} 0$ for $i = 1, 2, 3$.

From (2.3), we have

$$\hat{\boldsymbol{\eta}}_t \quad = \quad \left\{\boldsymbol{\Phi}(p_1) - \hat{\boldsymbol{\Phi}}(p_1)\right\}\mathbf{X}_t^{(1)}(p_1) + \boldsymbol{\xi}_{t,p_1}, \tag{3.11}$$

where $\boldsymbol{\xi}_{t,p_1} = \sum_{l=p_1+1}^{\infty} \boldsymbol{\Phi}_l^{(1)}\mathbf{X}_{t-l}^{(1)}$ represents the bias of the VAR(p_1) approximation of $\{\mathbf{X}_t^{(1)}\}$ and we can write,

$$\|\mathcal{C}_1\| \quad = \quad \|N^{-1/2}\sum_{t=j+1}^{N}\{\boldsymbol{\Phi}(p_1) - \hat{\boldsymbol{\Phi}}(p_1)\}\mathbf{X}_t^{(1)}(p_1)\boldsymbol{a}_{t-j}^{(2)^T} + N^{-1/2}\sum_{t=j+1}^{N}\boldsymbol{\xi}_t(p_1)\boldsymbol{a}_{t-j}^{(2)^T}\|$$

Using the triangular inequality and the property $\|\mathbf{AB}\| \leq \|\mathbf{A}\|\|\mathbf{B}\|$, we have

$$\|\mathcal{C}_1\| \quad \leq \quad \|\boldsymbol{\Phi}(p_1) - \hat{\boldsymbol{\Phi}}(p_1)\|\|N^{-1/2}\sum_{t=j+1}^{N}\mathbf{X}_t^{(1)}(p_1)\boldsymbol{a}_{t-j}^{(2)^T}\| + \|N^{-1/2}\sum_{t=j+1}^{N}\boldsymbol{\xi}_t(p_1)\boldsymbol{a}_{t-j}^{(2)^T}\|. \tag{3.12}$$

If we denote by $\mathcal{C}_{11} = \mathbb{E}\|N^{-1/2}\sum_{t=j+1}^{N}\mathbf{X}_t^{(1)}(p_1)\boldsymbol{a}_{t-j}^{(2)^T}\|^2$, and $\mathcal{C}_{12} = \mathbb{E}\|N^{-1/2}\sum_{t=j+1}^{N}\boldsymbol{\xi}_t(p_1)\boldsymbol{a}_{t-j}^{(2)^T}\|^2$, we can write

$$\begin{aligned}
\mathcal{C}_{11} \quad &= \quad N^{-1}\mathbb{E}\{tr[(\sum_{t=j+1}^{N}\mathbf{X}_t^{(1)}(p_1)\boldsymbol{a}_{t-j}^{(2)^T})(\sum_{t=j+1}^{N}\mathbf{X}_t^{(1)}(p_1)\boldsymbol{a}_{t-j}^{(2)^T})^T]\} \\
&= \quad N^{-1}\mathbb{E}\{tr[\sum_{t=j+1}^{N}\mathbf{X}_t^{(1)}(p_1)\boldsymbol{a}_{t-j}^{(2)^T}\boldsymbol{a}_{t-j}^{(2)}\mathbf{X}_t^{(1)}(p_1)^T]\} \\
&+ \quad N^{-1}\mathbb{E}\{tr[\sum_{t=j+1}^{N}\sum_{\substack{t_1=j+1 \\ t_1\neq t}}^{N}\mathbf{X}_t^{(1)}(p_1)\boldsymbol{a}_{t-j}^{(2)^T}\boldsymbol{a}_{t_1-j}^{(2)}\mathbf{X}_{t_1}^{(1)}(p_1)^T]\}
\end{aligned}$$

Using properties of the trace operator and under the hypothesis of non-correlation between the two processes, we obtain that

$$\begin{aligned}\mathcal{C}_{11} &= N^{-1}tr\{\sum_{t=j+1}^{N}\mathbb{E}(\boldsymbol{a}_{t-j}^{(2)^T}\boldsymbol{a}_{t-j}^{(2)})\mathbb{E}(\mathbf{X}_{t-j}^{(1)}(p_1)^T\mathbf{X}_{t-j}^{(1)}(p_1))\}\\ &+ N^{-1}tr\{\sum_{t=j+1}^{N}\sum_{\substack{t_1=j+1\\ t_1\neq t}}^{N}\mathbb{E}(\boldsymbol{a}_{t-j}^{(2)^T}\boldsymbol{a}_{t_1-j}^{(2)})\mathbb{E}(\mathbf{X}_{t_1}^{(1)}(p_1)^T\mathbf{X}_{t}^{(1)}(p_1))\}.\end{aligned}$$

It follows that

$$\mathcal{C}_{11} = \mathbb{E}\|N^{-1/2}\sum_{t=j+1}^{N}\mathbf{X}_t^{(1)}(p_1)\boldsymbol{a}_{t-j}^{(2)^T}\|^2 \leq \Delta p_1, \tag{3.13}$$

since $\mathbb{E}(\boldsymbol{a}_{t-j}^{(2)^T}\boldsymbol{a}_{t_1-j}^{(2)}) = 0$, $t_1 \neq t$, and

$$\mathbb{E}(\mathbf{X}_t^{(1)}(p_1)^T\mathbf{X}_t^{(1)}(p_1)) = \mathbb{E}(\sum_{i=1}^{p_1}\mathbf{X}_{t-i}^{(1)^T}\mathbf{X}_{t-i}^{(1)}) \leq \Delta p_1. \tag{3.14}$$

By a similar argument, we can write

$$\mathcal{C}_{12} = N^{-1}\mathbb{E}[tr\{\sum_{t=j+1}^{N}\boldsymbol{\xi}_t(p_1)^T\boldsymbol{\xi}_t(p_1)\boldsymbol{a}_{t-j}^{(2)^T}\boldsymbol{a}_{t-j}^{(2)}\}],$$

and from the definition of $\boldsymbol{\xi}_t(p_1)$, we have

$$\mathbb{E}\|\boldsymbol{\xi}_t(p_1)\|^2 = \sum_{l=p_1+1}^{\infty}\sum_{l_1=p_1+1}^{\infty}tr(\boldsymbol{\Phi}_l^{(1)}\mathbb{E}(\mathbf{X}_{t-l}^{(1)}\mathbf{X}_{t-l_1}^{(1)^T})\boldsymbol{\Phi}_{l_1}^{(1)^T}).$$

Using the Cauchy-Schwarz inequality for $tr(\boldsymbol{\Phi}_l^{(1)}\boldsymbol{\Phi}_{l_1}^{(1)^T}) = \langle\boldsymbol{\Phi}_l^{(1)}, \boldsymbol{\Phi}_{l_1}^{(1)}\rangle$ and since $\|\boldsymbol{\Gamma}_{\mathbf{X}}^{(11)}(l)\|$ is uniformly bounded, we obtain

$$\mathbb{E}\|\boldsymbol{\xi}_t(p_1)\|^2 \leq \Delta\sum_{l=p_1+1}^{\infty}\sum_{l_1=p_1+1}^{\infty}\|\boldsymbol{\Phi}_l^{(1)}\|\|\boldsymbol{\Phi}_{l_1}^{(1)}\| \leq \Delta(\sum_{l=p_1+1}^{\infty}\|\boldsymbol{\Phi}_l^{(1)}\|)^2. \tag{3.15}$$

Under the hypothesis of non-correlation of the two processes, it follows that

$$\mathcal{C}_{12} \leq \Delta(\sum_{l=p_1+1}^{\infty}\|\boldsymbol{\Phi}_l^{(1)}\|)^2 = O(\frac{1}{N}). \tag{3.16}$$

From the assumptions (i) and (ii) and using (3.12), (2.4), (3.13) and (3.16), we finally obtain that

$$\|\mathcal{C}_1\| = O_p(\frac{p_1^{1/2}}{N^{1/2}})O_p(p_1^{1/2}) + O_p(N^{-1/2}) = O_p(\frac{p_1}{N^{1/2}}), \tag{3.17}$$

and consequently $\|\mathcal{C}_1\| \xrightarrow{p} 0$. By symmetry, we obtain a similar result for $\mathcal{C}_2$.

For the third term $\mathcal{C}_3$ of (3.10), we have

$$\begin{aligned}
\|\mathcal{C}_3\| &= \|N^{-1/2}\sum_{t=j+1}^{N}\hat{\boldsymbol{\eta}}_t\hat{\boldsymbol{\delta}}_{t-j}^T\| \\
&\leq \|\boldsymbol{\Phi}(p_1)-\hat{\boldsymbol{\Phi}}(p_1)\|\|\boldsymbol{\Phi}(p_2)-\hat{\boldsymbol{\Phi}}(p_2)\|\|N^{-1/2}\sum_{t=j+1}^{N}\mathbf{X}_t^{(1)}(p_1)\mathbf{X}_{t-j}^{(2)}(p_2)^T\| \\
&+ \|\boldsymbol{\Phi}(p_1)-\hat{\boldsymbol{\Phi}}(p_1)\|\|N^{-1/2}\sum_{t=j+1}^{N}\mathbf{X}_t^{(1)}(p_1)\boldsymbol{\xi}_{t-j}(p_2)^T\| \\
&+ \|\boldsymbol{\Phi}(p_2)-\hat{\boldsymbol{\Phi}}(p_2)\|\|N^{-1/2}\sum_{t=j+1}^{N}\boldsymbol{\xi}_t(p_1)\mathbf{X}_{t-j}^{(2)}(p_2)^T\| \\
&+ \|N^{-1/2}\sum_{t=j+1}^{N}\boldsymbol{\xi}_t(p_1)\boldsymbol{\xi}_{t-j}(p_2)^T\|. \qquad (3.18)
\end{aligned}$$

Let us denote $\mathcal{C}_{31} = \|N^{-1/2}\sum_{t=j+1}^{N}\mathbf{X}_t^{(1)}(p_1)\mathbf{X}_{t-j}^{(2)}(p_2)^T\|^2$, $\mathcal{C}_{32} = \|N^{-1/2}\sum_{t=j+1}^{N}\mathbf{X}_t^{(1)}(p_1)\boldsymbol{\xi}_{t-j}(p_2)^T\|^2$, $\mathcal{C}_{33} = \|N^{-1/2}\sum_{t=j+1}^{N}\boldsymbol{\xi}_t(p_1)\mathbf{X}_{t-j}^{(2)}(p_2)^T\|^2$ and $\mathcal{C}_{34} = \|N^{-1/2}\sum_{t=j+1}^{N}\boldsymbol{\xi}_t(p_1)\boldsymbol{\xi}_{t-j}(p_2)^T\|^2$. Arguing as in the proof of Theorem 1 of Lewis and Reinsel (1985), we can show that

$$\mathcal{C}_{31} = \mathbb{E}\|N^{-1/2}\sum_{t=j+1}^{N}\mathbf{X}_t^{(1)}(p_1)\mathbf{X}_{t-j}^{(2)}(p_2)^T\|^2 \leq \Delta p_1 p_2. \qquad (3.19)$$

For the second term $\mathcal{C}_{32}$, define $\boldsymbol{Y}_t = \mathbf{X}_t^{(1)}(p_1)\boldsymbol{\xi}_{t-j}(p_2)^T$. Under the hypothesis of non-correlation of the two processes, we have that

$$\begin{aligned}
tr(\mathbb{E}(\boldsymbol{Y}_t\boldsymbol{Y}_t^T)) &= tr(\mathbb{E}(\mathbf{X}_t^{(1)}(p_1)\boldsymbol{\xi}_{t-j}(p_2)^T\boldsymbol{\xi}_{t-j}(p_2)\mathbf{X}_t^{(1)}(p_1)^T)) \\
&= \mathbb{E}(\|\mathbf{X}_t^{(1)}(p_1)\|^2)\mathbb{E}(\|\boldsymbol{\xi}_{t-j}(p_2)\|^2).
\end{aligned}$$

Using (3.14) and (3.15), we have that $\mathbb{E}(\boldsymbol{Y}_t\boldsymbol{Y}_t^T) \leq \Delta p_1(\sum_{l=p_2+1}^{\infty}\|\boldsymbol{\Phi}_l^{(2)}\|)^2$. By a similar argument to the one made to bound $\mathcal{C}_{11}$, we obtain that

$$\mathcal{C}_{32} = N^{-1}\mathbb{E}\{tr[(\sum_{t=j+1}^{N}\boldsymbol{Y}_t)(\sum_{t=j+1}^{N}\boldsymbol{Y}_t^T)]\} \leq \Delta p_1(\sum_{l=p_2+1}^{\infty}\|\boldsymbol{\Phi}_l^{(2)}\|)^2. \qquad (3.20)$$

By symmetry, we also get

$$\mathcal{C}_{33} \leq \Delta p_2(\sum_{l=p_1+1}^{\infty}\|\boldsymbol{\Phi}_l^{(1)}\|)^2. \qquad (3.21)$$

Finally, under the non-correlation hypothesis, by a similar development to the one for $\mathcal{C}_{32}$, it follows that

$$\mathcal{C}_{34} \leq \Delta(\sum_{l=p_1+1}^{\infty}\|\boldsymbol{\Phi}_l^{(1)}\|)^2(\sum_{l=p_2+1}^{\infty}\|\boldsymbol{\Phi}_l^{(2)}\|)^2. \qquad (3.22)$$

Combining the results (3.19)-(3.22), we have

$$\begin{aligned}\|\mathcal{C}_3\| &= O_p(\frac{p_1^{1/2}}{N^{1/2}})O_p(\frac{p_2^{1/2}}{N^{1/2}})O_p(p_1^{1/2}p_2^{1/2}) \\ &+ O_p(\frac{p_1^{1/2}}{N^{1/2}})O_p(p_1^{1/2}\sum_{l=p_2+1}^{\infty}\|\mathbf{\Phi}_l^{(2)}\|) + O_p(\frac{p_2^{1/2}}{N^{1/2}})O_p(p_2^{1/2}\sum_{l=p_1+1}^{\infty}\|\mathbf{\Phi}_l^{(1)}\|) \\ &+ O_p(\sum_{l=p_1+1}^{\infty}\|\mathbf{\Phi}_l^{(1)}\|\sum_{l=p_2+1}^{\infty}\|\mathbf{\Phi}_l^{(2)}\|),\end{aligned}$$

Using once again the assumptions (i) and (ii), it follows that $\|\mathcal{C}_3\| \xrightarrow{p} 0$ and the proof of the theorem is completed. □

4 Application

4.1 Tests procedures for non-correlation

The null hypothesis that the two VAR(∞) processes $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ are uncorrelated, that is, $\boldsymbol{\rho}_{\mathbf{X}}^{(12)}(j) = \mathbf{0}$, $j \in \mathbb{Z}$, is equivalent under the assumptions made to

$$\mathcal{H}_0 \quad : \boldsymbol{\rho}_{\boldsymbol{a}}^{(12)}(j) = \mathbf{0} \ , \ \forall j \in \mathbb{Z}.$$

One can now construct the test statistics in a similar way as in El Himdi and Roy (1997). Two types of test statistics will be considered.

The first type of test is based on the residual cross-correlations at individual lags. Since $\hat{\boldsymbol{\rho}}_h = \mathbf{R}_{\hat{\boldsymbol{a}}}^{(hh)}(0)$, $h = 1, 2$, is a consistent estimator of $\boldsymbol{\rho}_h$, we define the following statistic

$$\mathcal{Q}(j) \quad = \quad N\mathbf{r}_{\hat{\boldsymbol{a}}}^{(12)}(j)^T \left(\hat{\boldsymbol{\rho}}_2^{-1} \otimes \hat{\boldsymbol{\rho}}_1^{-1}\right) \mathbf{r}_{\hat{\boldsymbol{a}}}^{(12)}(j), \tag{4.1}$$

where $\mathbf{r}_{\hat{\boldsymbol{a}}}^{(12)}(j) = vec(\mathbf{R}_{\hat{\boldsymbol{a}}}^{(12)}(j))$ and under $\mathcal{H}_0$, $\mathcal{Q}(j)$ is asymptotically distributed as a $\chi^2_{m_1m_2}$ variable by Theorem 3.1. Thus, at a given significance level α, $\mathcal{H}_0$ is rejected if $\mathcal{Q}(j) > \chi^2_{m_1m_2,1-\alpha}$, where $\chi^2_{m,p}$ denotes the p-quantile of the χ^2_m distribution.

In practice, we usually want to consider simultaneously many lags, for example all lags not greater than M in absolute value. A global test based on the statistics $\mathcal{Q}(j)$, $|j| \leq M$, consists in rejecting $\mathcal{H}_0$ if for at least one $j \in \{-M, \ldots, M\}$, $\mathcal{Q}(j) > \chi^2_{m_1m_2,1-\alpha_0}$. Since the statistics $\mathcal{Q}(j)$ are asymptotically independent, the use of the marginal level $\alpha_0 = 1 - (1-\alpha)^{1/(2M+1)}$ for each j assures us that the asymptotic global level is α.

The second type of test is the generalization of Haugh's global test described in El Himdi and Roy (1997). It is based on the statistic

$$\mathcal{Q}_M = N\boldsymbol{r}_{\hat{\mathbf{a}},M}^{(12)^T}\left(\mathbf{I}_{2M+1}\otimes\hat{\boldsymbol{\rho}}_2^{-1}\otimes\hat{\boldsymbol{\rho}}_1^{-1}\right)\boldsymbol{r}_{\hat{\mathbf{a}},M}^{(12)} = \sum_{j=-M}^{M}\mathcal{Q}(j), \tag{4.2}$$

where $\boldsymbol{r}_{\hat{\mathbf{a}},M}^{(12)} = (\mathbf{r}_{\hat{\mathbf{a}}}^{(12)}(-M)^T,..,\mathbf{r}_{\hat{\mathbf{a}}}^{(12)}(M)^T)^T$, and $M \leq N-1$ is fixed with respect to N. Under $\mathcal{H}_0$, $\mathcal{Q}_M$ obeys asymptotically a $\chi^2_{(2M+1)m_1m_2}$ distribution from Theorem 3.1 and $\mathcal{H}_0$ is rejected for large values of $\mathcal{Q}_M$. This statistic can be also expressed in term of the residual autocovariances $\mathbf{C}_{\hat{\mathbf{a}}}^{(hh)}(0)$, $h = 1,2$, and of cross-covariances $\mathbf{C}_{\hat{\mathbf{a}}}^{(12)}(j)$, see Lemma 4.1 of El Himdi and Roy (1997). As in El Himdi and Roy (1997), the modified statistics

$$\mathcal{Q}^*(j) = \frac{N}{N-|j|}\mathcal{Q}(j), \qquad \mathcal{Q}^*_M = \sum_{j=-M}^{M}\mathcal{Q}^*(j), \tag{4.3}$$

will be considered as they seem to be better approximated by a chi-square distribution, for small sample sizes.

From Theorem 3.1, we can also derive the asymptotic distribution of a more general test statistic proposed by Hallin and Saidi (2002) that takes into account a possible pattern in the sign of the cross-correlation coefficients. Define the vector

$$\boldsymbol{\nu}_M = \sqrt{N}\boldsymbol{I}_{2M+1}\otimes\left(\hat{\boldsymbol{\rho}}_2^{-\frac{1}{2}}\otimes\hat{\boldsymbol{\rho}}_1^{-\frac{1}{2}}\right)\boldsymbol{r}_{\hat{\mathbf{a}},M}^{(12)},$$

and the statistic

$$\mathcal{Q}_{i,M} = \sum_{k=1}^{(2M+1)m_1m_2-i}\left[\sum_{l=0}^{i}\boldsymbol{\nu}_M(k+l)\right]^2, \; i = 0,1,..,Mm_1m_2-1,$$

where $\boldsymbol{\nu}_M(j)$ denotes the j^{th} element of the vector $\boldsymbol{\nu}_M$. When $m_1 = m_2 = 1$, we retrieve Koch and Yang's (1986) statistic for univariate series. The case $i = 0$ corresponds to the statistic $\mathcal{Q}_M$ defined by (4.2). Under the assumptions of Theorem 3.1, the vector $\boldsymbol{\nu}_M$ is asymptotically $\mathcal{N}(\mathbf{0},\mathbb{I})$ and since $\mathcal{Q}_{i,M}$ can be written as the following quadratic form

$$\mathcal{Q}_{i,M} = \boldsymbol{\nu}_M^T\boldsymbol{D}_{i,M}\boldsymbol{\nu}_M,$$

where the matrix $\boldsymbol{D}_{i,M}$ is defined in Proposition 2.1 of Hallin and Saidi (2002), we have the following result.

Proposition 4.1 *Under the assumptions of Theorem 3.1 and under the null hypothesis that* $\mathbf{X}^{(1)}$ *and* $\mathbf{X}^{(2)}$ *are uncorrelated, we have that*

$$\mathcal{Q}_{i,M} \quad \xrightarrow{L} \quad \mathcal{Q}(\boldsymbol{D}_{i,M}) = \sum_{j=1}^{(2M+1)m_1m_2} \lambda_{i,M}(j)\mathcal{Z}_j^2$$

where the coefficients $\lambda_{i,M}(j)$ *are the eigenvalues of the matrix* $\boldsymbol{D}_{i,M}$ *and the* $\mathcal{Z}_j$ *are independent and identically distributed* $\mathcal{N}(0,1)$ *random variables.*

The matrix $\boldsymbol{D}_{i,M}$ is completely determined by i and M and the simulation results of Hallin and Saidi (2002) show that in some situations, the test $\mathcal{Q}_{i,M}$ with $i > 0$ is more powerful than $\mathcal{Q}_{0,M} = \mathcal{Q}_M$.

4.2 Simulation study

Here, we report the results of a small simulation experiment conducted in order to compare the exact distribution of the statistics $\mathcal{Q}^*(j)$ and $\mathcal{Q}_M^*$ with their corresponding asymptotic chi-square distributions, under the null hypothesis for non-correlation. To do that, we examined the empirical frequencies of rejection of the null hypothesis with the proposed tests at three different nominal levels (1, 5 and 10 percent) for each of two series lengths ($N = 100$ and 200) and for two different global models for $\{\mathbf{X}_t^{(1)}\}$ and $\{\mathbf{X}_t^{(2)}\}$. These models are described in Table 1. The dimension of each of the two models is four and for each one, the submodels for $\{\mathbf{X}_t^{(1)}\}$ and $\{\mathbf{X}_t^{(2)}\}$ are bivariate. Also, with the considered values for the autoregressive and moving average parameters as well as for the covariance matrices of the innovations, the subprocesses $\{\mathbf{X}_t^{(1)}\}$ and $\{\mathbf{X}_t^{(2)}\}$ are uncorrelated and the corresponding submodels are stationary and invertible.

For each model, the experiment proceeded in the following way.

1. For each model and for each series length N, 5000 independent realizations were generated. First, pseudo-random $\mathcal{N}(0,1)$ variables were obtained with the S-plus pseudo-random normal generator and were transformed into independent $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{a}})$ vectors using the Cholesky transformation. Second, the $\mathbf{X}_t$ values were obtained by directly solving the difference equation that defines a VARMA model. For the AR(1) model, the first observation $\mathbf{X}_1$ was generated from the exact $\mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_{\mathbf{X}}(\mathbf{0}))$ distribution of the $\mathbf{X}_t$'s. The covariance $\boldsymbol{\Gamma}_{\mathbf{X}}(0)$ was obtained with Ansley's (1980) algorithm.

2. An autoregression was fitted to each series by conditional least squares estimation. The autoregressive order was obtained by minimizing the AIC criterion for $p \leq P$, where P was fixed to 12. The resulting residual series $\{\hat{\boldsymbol{a}}_t^{(h)}\}$, $h = 1, 2$, were cross-correlated by computing $\mathbf{R}_{\hat{\mathbf{a}}}^{(12)}(j)$ as defined by (3.2).

3. The values of the test statistic $\mathcal{Q}^*(j)$ were computed for $j = -12, ..., 12$ and those of $\mathcal{Q}_M^*$ for

Table 1: Time series models used in the simulation study

MODELS	EQUATIONS	Σ_a
AR(1)	$\begin{bmatrix} \mathbf{X}_t^{(1)} \\ \mathbf{X}_t^{(2)} \end{bmatrix} = \begin{bmatrix} \Phi^{(1)} & 0 \\ 0 & \Phi^{(2)} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{t-1}^{(1)} \\ \mathbf{X}_{t-1}^{(2)} \end{bmatrix} + \begin{bmatrix} \mathbf{a}_t^{(1)} \\ \mathbf{a}_t^{(2)} \end{bmatrix}$	$\begin{bmatrix} \Sigma_a^{(1)} & 0 \\ 0 & \Sigma_a^{(2)} \end{bmatrix}$
MA(1)	$\begin{bmatrix} \mathbf{X}_t^{(1)} \\ \mathbf{X}_t^{(2)} \end{bmatrix} = \begin{bmatrix} \Theta^{(1)} & 0 \\ 0 & \Theta^{(2)} \end{bmatrix} \begin{bmatrix} \mathbf{a}_{t-1}^{(1)} \\ \mathbf{a}_{t-1}^{(2)} \end{bmatrix} + \begin{bmatrix} \mathbf{a}_t^{(1)} \\ \mathbf{a}_t^{(2)} \end{bmatrix}$	$\begin{bmatrix} \Sigma_a^{(1)} & 0 \\ 0 & \Sigma_a^{(2)} \end{bmatrix}$
	PARAMETER VALUES	
$\Phi^{(1)} = \begin{bmatrix} 1.2 & -0.5 \\ 0.6 & 0.3 \end{bmatrix}$	$\Phi^{(2)} = \begin{bmatrix} -0.6 & 0.3 \\ 0.3 & 0.6 \end{bmatrix}$	$\Theta^{(1)} = \begin{bmatrix} -0.2 & 0.3 \\ -0.6 & 1.1 \end{bmatrix}$
$\Theta^{(2)} = \begin{bmatrix} 0.8 & 0.3 \\ 0.1 & 0.6 \end{bmatrix}$	$\Sigma_a^{(1)} = \begin{bmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{bmatrix}$	$\Sigma_a^{(2)} = \begin{bmatrix} 1.0 & 0.75 \\ 0.75 & 1.0 \end{bmatrix}$

$M = 1, .., 12$. For each test, the value of the statistic was compared with the critical value obtained from the corresponding chi-square distribution.

4. Finally, for each nominal level and for each series length N, we obtained from the 5000 realizations, the empirical frequencies of rejection of the null hypothesis of non-correlation. The standard error of the empirical levels based on 5000 realizations is 0.22% for the nominal level 1%, 0.49% for 5% and 0.67% for 10%.

The empirical levels of tests at individual lags based on $\mathcal{Q}^*(j)$ are presented in Table 5, for $|k| = 0, 1, 2, 4, 6, 8$ and 10. We make the following observations. For the series length $N = 100$, the chi-square distribution provides a relatively good approximation for all lags at the three significance levels since almost all corresponding empirical levels lie within the 5% significant limits. With $N = 200$, the approximation is very good for the three significant levels. The results for the portmanteau test $\mathcal{Q}_M^*$, $M = 1, .., 12$, are given in the second part of Table 5 for the two models. The chi-square approximation for all values of M is very good at the three significance levels especially with $N = 200$.

5 Conclusion

In this paper, we showed that the procedure for testing the uncorrelatedness of two multivariate stationary ARMA time series described in El Himdi and Roy (1997) remains valid in the more general context of two stationary VAR(∞) series. At the modeling stage, a possible high order autoregression is fitted to each series. The autoregression modeling protects us against misspecifications of the true underlying VARMA models that may invalidate the asymptotic theory and eventually lead to wrong conclusions.

The global statistics considered in this paper are based on the lags j such that $|j| \leq M$. They lead to consistent tests only if there is no cross-correlation at lags j such that $|j| > M$. In order to have a consistent test for an arbitrary cross-correlation structure (possible at lags j such that $|j| > M$), we may consider a statistic that takes into account all possible lags from $-N+1$ to $N-1$ (N representing the sample size) as in Hong (1996). A version of Hong's statistic for multivariate time series was proposed by Bouhaddioui (2002) and it will be the subject of a forthcoming paper.

References

Ansley, C. F. (1980), 'Computation of the theoretical autocovariance function for a vector ARMA process', *J. Statist. Comput. Simul.* **12**, 15–24.

Berk, K. N. (1974), 'Consistent autoregressive spectral estimates', *Ann. Statist.* **2**, 489–502.

Bouhaddioui, C. (2002), 'Tests d'indépendance de deux séries multivariées autorégressives d'ordre infini', *Unpublished Ph.D. Thesis, Département de mathématiques et de statistique, Université de Montréal.*

El Himdi, K. and Roy, R. (1997), 'Tests for noncorrelation of two multivariate ARMA time series', *Canad. J. Statist.* **25**, 233–256.

Hallin, M. and Saidi, A. (2002), 'Testing independence and causality between multivariate ARMA times series', *J. Time Ser. Anal.,* forthcoming.

Haugh, L. D. (1976), 'Checking the independence of two covariance-stationary time series: a univariate residual cross correlation approch', *J. Amer. Statist. Assoc.* **71**, 376–385.

Hong, Y. (1996), 'Testing for independence between two covariance stationary time series', *Biometrika* **83, 3**, 615–625.

Koch, P. and Yang, S.-S. (1986), 'A method for testing the independance of two time series that accounts for a potential pattern in the cross-correlation function', *J. Amer. Statist. Assoc.* **81**, 533–544.

Lewis, R. and Reinsel, G. (1985), 'Prediction of multivariate time series by autoregressive model fitting', *J. Mult. Anal.* **16**, 393–411.

Lütkepohl, H. (1991), *Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.

Paparoditis, E. (1996), 'Bootstrapping autoregressive and moving average parameter estimates of infinite order vector autoregressive processes', *J. Mult. Anal.* **57**, 277–296.

Roy, R. (1989), 'Asymptotic covariance structure of serial correlation in multivariate time series', *Biometrika* **76**, 824–827.

Table 2: Empirical level of the tests at individual lags $\mathcal{Q}^*(j)$ and of the global test $\mathcal{Q}^*_M$ defined by (4.3) for the AR(1) and MA(1) models.

	$\mathcal{Q}^*(j)$											
	AR(1)						MA(1)					
	N=100			N=200			N=100			N=200		
α	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
k=-12	0.90	5.15	10.15	1.20	5.80	10.90	0.70	4.80	10.15	1.10	5.60	10.20
-10	0.95	4.90	10.60	0.90	4.80	10.50	0.85	5.05	10.60	0.80	5.20	9.90
-8	0.85	4.40	9.00	0.50	4.60	9.90	0.85	5.20	9.00	0.95	4.80	10.10
-6	0.95	5.20	10.55	1.00	5.20	10.20	0.75	4.60	10.55	1.20	5.00	10.20
-4	1.10	5.20	9.70	1.10	4.40	9.60	1.00	4.80	9.70	1.20	4.60	9.30
-2	0.85	4.20	9.15	0.90	5.20	9.70	0.65	5.10	9.15	0.70	4.40	9.80
0	1.05	5.15	11.00	1.20	5.90	10.90	1.15	5.15	11.00	0.80	5.20	10.50
2	0.80	5.10	9.55	1.10	4.90	10.00	0.90	4.90	9.55	0.80	4.40	10.00
4	0.85	5.20	10.30	1.10	6.30	11.7	0.90	4.80	10.30	1.10	5.10	10.30
6	1.00	4.75	9.75	0.90	4.20	9.70	1.20	4.75	9.75	0.90	5.30	9.70
8	0.95	4.80	10.30	0.70	4.30	9.10	0.80	5.20	10.30	0.75	5.30	9.80
10	0.75	4.95	10.75	1.40	4.90	10.60	0.70	4.95	10.75	0.80	4.90	10.50
12	0.80	4.60	10.35	0.70	4.50	11.00	0.90	4.70	10.35	0.90	4.80	10.20
	$\mathcal{Q}^*_M$											
$M=1$	0.60	5.05	10.50	0.90	5.50	11.10	0.70	4.65	9.60	0.80	5.40	10.40
2	0.90	4.55	9.35	0.60	5.20	11.10	0.75	4.80	9.80	0.80	4.80	9.50
3	0.95	4.40	9.40	1.00	5.60	11.10	0.70	4.40	9.70	0.90	5.30	11.20
4	0.80	4.80	10.10	0.80	5.40	10.60	0.60	5.10	9.60	0.85	5.30	10.60
5	0.50	5.05	10.15	0.60	5.70	11.40	0.55	5.30	9.40	0.60	4.90	9.45
6	0.55	5.10	9.75	0.70	6.30	11.10	0.50	4.30	9.30	0.55	5.10	9.85
7	0.60	4.45	9.70	0.70	6.10	11.10	0.55	4.75	9.50	0.70	5.50	10.30
8	0.65	4.35	9.45	0.80	5.10	9.30	0.80	5.40	9.45	0.85	4.70	9.70
9	0.55	3.90	9.30	0.70	4.60	9.70	0.65	4.20	9.50	0.65	4.80	9.65
10	0.85	4.55	10.00	0.60	4.70	10.10	0.70	5.60	9.90	0.80	5.10	10.45
11	0.90	4.35	9.80	0.60	4.30	9.80	0.70	4.50	9.90	0.90	4.70	11.00
12	0.80	4.45	9.85	0.80	4.90	9.50	0.60	4.85	9.65	0.95	5.10	10.60