Semiparametric independence tests between two infinite-order cointegrated series *

Chafik Bouhaddioui † United Arab Emirates University
Jean-Marie Dufour ‡ McGill University
Masaya Takano § McGill University

June 2022

ABSTRACT

We propose a semiparametric approach for testing independence between two infinite-order cointegrated vector autoregressive series [IVAR(∞)]. The procedures considered can be viewed as extensions of classical methods proposed by Haugh (1976, *JASA*) and Hong (1996, *Biometrika*) for testing independence between stationary univariate time series. The tests are based on the residu-als of long autoregressions, hence allowing for computational simplicity, weak assumptions on the form of the underlying process, and a direct interpretation of the results in terms of innovations (or shocks). The test statistics are standardized versions of the sum of weighted squares of residual cross-correlation matrices. The weights depend on a kernel function and a truncation parameter. Multivariate portmanteau statistics can be viewed as a special case of our procedure based on the truncated uniform kernel. The asymptotic distributions of the test statistics under the null hypoth-esis are derived, and consistency is established against fixed alternatives of serial cross-correlation of unknown form. A simulation study is presented which indicates that the proposed tests have good size and power properties in finite samples.

Keywords: Infinite-order cointegrated vector autoregressive process; independence; causality; residual cross-correlation; consistency; asymptotic power.

---

* The authors thank Yoosoon Chang, Kilani Ghoudi, Ian McLeod, Abderrahim Taamouti, and an anonymous ref-eree for several useful comments. Earlier versions of this paper were presented at the Statistical Society of Canada annual meeting and the University of Western Ontario (Department of Statistical and Actuarial Sciences). This work was supported by the William Dow Chair in Political Economy (McGill University), the Bank of Canada (Research Fellowship), the Toulouse School of Economics (Pierre-de-Fermat Chair of excellence), the Universitad Carlos III de Madrid (Banco Santander de Madrid Chair of excellence), a Guggenheim Fellowship, a Konrad-Adenauer Fellowship (Alexander-von-Humboldt Foundation, Germany), the Canadian Network of Centres of Excellence [program on *Math-ematics of Information Technology and Complex Systems* (MITACS)], the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Fonds de recherche sur la société et la culture (Québec).

† Department of Statistics, United Arab Emirates. Mailing address: College of Business and Economics, Department of Statistics, United Arab Emirates University, P.O. Box 17555 Al Ain, United Arab Emirates. TEL: 971 3 713 3345; FAX 971 3 762 4385; e-mail: chafikb@uaeu.ac.ae .

‡ William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en économie quantitative (CIREQ). Mailing ad-dress: Department of Economics, McGill University, Leacock Building, Room 414, 855 Sherbrooke Street West, Mon-tréal, Québec H3A 2T7, Canada. TEL: (1) 514 398 6071; FAX: (1) 514 398 4800; e-mail: jean-marie.dufour@ mcgill.ca . Web page: http://www.jeanmariedufour.com

§ McGill University, Leacock Building, Room 414 855 Sherbrooke Street West Montreal, Quebec, H3A 2T7. Tel: +1-514-398-3030; e-mail: masaya.takano@mail.mcgill.ca.

1. Introduction

Studying the dynamic relationship between two multivariate series is a fundamental objective of time series analysis in statistics and econometrics. For example, in econometrics, this can help one to understand the associated economic mechanisms. In this context, a basic problem consists in testing independence (or the absence of serial cross-correlation) between two vector processes. The seminal paper on this problem is due to Haugh (1976), who proposed a general procedure for testing independence between two covariance-stationary ARMA time series. His method is based on considering cross-correlations between residuals obtained after fitting univariate ARMA mod-els on each series. Since the innovations of an ARMA model follow a white noise by assumption, this considerably simplifies the underlying distributional theory, and the corresponding tests are relatively simple to apply. Further, the corresponding statistics have a direct interpretation in terms of process innovations (or reduced-form shocks), a feature of interest in econometrics since inno-vations can often be interpreted as "shocks" to economic systems. Consequently, the possibility of focusing on "shock cross-correlations" should be useful in econometric research.

The work of Haugh (1976) has been extended by several authors; see Hong (1996*a*), El Himdi and Roy (1997), Pham, Roy and Cédras (2003), Hallin and Saidi (2005), Bouhaddioui and Roy (2006*a*, 2006*b*), Hallin and Saidi (2007), Saidi (2007), and Bouhaddioui and Dufour (2008). Most of these studies focus on independence between two multivariate finite-order vector autoregressive (VAR) or vector autoregressive moving-average (VARMA) models. El Himdi and Roy (1997) extended the procedure developed by Haugh (1976) in order to test non-correlation between two time series in the context of multivariate stationary and invertible VARMA models. This result was used by Hallin and Saidi (2005) to develop a test which takes into account a possible pattern in the signs of cross-correlations at different lags. In a nonparametric setup, Hallin, Jurecková,ˇ Picek and Zahaf (1999) proposed a test for independence between two autoregressive time series which is based on autoregressive rank scores, while Hong (1998) proposed a test based on empirical distribution functions.

The stationarity condition is often unrealistic and constitutes a strong limitation. Even though stationarity may be achieved in many cases by differencing each series (so that distributional com-plications are avoided), this type of transformation can distort our ability to identify or accurately measure parameters and relations of interest. It is typically more interesting to be able to work with the original series without prefiltering (like differencing). This is especially important if we wish to study cointegrating relationships.

Engle and Granger (1987) introduced the concept of cointegration, which is used in many stud-ies across several fields. In the case of a finite-order autoregressive cointegrated vector, Ahn and Reinsel (1990) developed an efficient estimation method for Gaussian processes. Yap and Reinsel (1995) proposed full- and reduced-rank Gaussian estimation procedures for cointegrated VARMA processes. For a good discussion of the related models, see Lütkepohl (2001). By exploiting the estimation methods proposed by Yap and Reinsel (1995), Pham et al. (2003) generalized the main result of El Himdi and Roy (1997) to the case of two cointegrated (or partially nonstationary) VARMA series. They proposed test statistics based on residual cross-correlation matrices $\mathbf{Ra}^{(12)}\hat{}(j)$, $|j| \leq M$ [where M does not depend on the sample size n] between the two residual series $\mathbf{a}\hat{}^{(1)}{}_t$ and $\mathbf{a}\hat{}^{(2)}{}_t$ 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 show that an arbitrary vector of residual cross-correlations asymptotically follows a multivariate normal distribution.

In practice, a finite-order VAR model can be a rough approximation to the true data generating process of a multivariate time series. The “true” model may easily not be reducible to a parsimo-nious model with a small number of unknown parameters. From this perspective, a more flexible alternative approach assumes that the data are generated by an infinite-order autoregressive process. Such models lead one to consider a truncated (potentially long) autoregression as an approximation of the underlying process. In statistics and econometrics, one typically derives the properties of es-timators and test criteria under the assumption of correct specification, even if model assumptions are clearly not fulfilled. For example, in VARMA estimation, it is well known that misspecification of the AR or MA orders can lead to inconsistent estimators. Further, the estimation of VARMA models is highly nonlinear and raises difficult identification complications (in the sense of multiple observationally equivalent representations).

The autoregressive model fitting approach has been successfully applied by several authors: Akaike (1969), Berk (1974) and Parzen (1974) for spectral density estimation, Parzen (1974), Lütkepohl (1985), Lewis and Reinsel (1985) and Bhansali (1996) for prediction, Park (1990) and

Saikkonen (1992) for inference in cointegrated systems; see also Reinsel (1997), Lütkepohl (2005) and Park, Shin and Wang (2010) . In previous work [Bouhaddioui and Roy (2006*b*)], we have generalized the work of El Himdi and Roy (1997) to the case of two stationary multivariate infinite-order autoregressive series VAR(∞). This result allows one to develop tests against serial cross-correlation at a particular lag or at a fixed number of lags j such as $|j| \le M$, where M does not depend on the sample size n.

In the univariate stationary case, Hong (1996*b*) introduced an important extension of Haugh's procedure by proposing a class of spectral test statistics. His approach is semiparametric and valid for two infinite-order autoregressive series AR(∞). It is based on fitting an autoregressive model of order p to a series of n observations from each infinite-order autoregressive process. Following Berk (1974), the order p of the fitted autoregression is a function of the sample size. This approach was also used by Hong (1999), Duchesne and Roy (2003), Duchesne (2005) and Shao (2009) for the case of two univariate long memory processes. In Bouhaddioui and Roy (2006*a*), it is extended to VAR(∞) models, hence protecting against misspecification of the underlying VARMA model. In contrast with Haugh's test, which is based on the residual cross-correlations at lag j such that $|j| \le M$, the portmanteau test Q_n is consistent for a large class of serial cross-correlations alternatives of an arbitrary form between the two series.

The main objective of this paper is to propose a semiparametric approach to test independence between two infinite-order cointegrated autoregressive [IVAR(∞)] models against alternatives where they would be correlated. These models were introduced by Saikkonen (1992) and involve much weaker conditions than those considered by Yap and Reinsel (1995), Pham et al. (2003), Hallin and Saidi (2005) and Saidi (2007); for further, discussion of this setup, see Saikkonen and Lütkepohl (1996), Saikkonen and Luukkonen (1997), and Lütkepohl and Saikkonen (1997). The problem of testing the absence of correlation between two IVAR(∞) was first considered in Bouhaddioui and Dufour (2008), where the asymptotic distribution of an arbitrary vector of residual cross-correlations and partial cross-correlations under the hypothesis of non-correlation of the two series is derived under the assumption that innovations are a strong white noise. However, the test statistics proposed in the latter paper only consider one lag at a time or a fixed number of lags j (for example $|j| \le M$).

In this article, we propose a multivariate version of the weighted portmanteau statistic Q_n, based on the sample cross-correlation matrices $\mathbf{R_a}^{(12)\hat{}}(j)$, $|j| \leq n - 1$, between the residuals $\hat{\mathbf{a}}^{(1)}_t$ and $\hat{\mathbf{a}}^{(2)}_t$. The latter are obtained by approximating two multivariate IVAR(∞) series with finite-order autoregressions whose order increases with the sample size at an appropriate rate. The test statistics continue to have a N (0, 1) asymptotic distribution under the hypothesis of independence of the two series. The tests are consistent against serial cross-correlation of arbitrary form.

The paper is organized as follows. Section 2 describes the statistical framework as well as some preliminary results. The new test statistics are introduced in Section 3. We show that their asymp-totic distributions under the null hypothesis are N (0, 1). In section 4, we establish the consistency of the tests. In Section 6, we present the results of a small Monte Carlo experiment on the level and power of the tests in finite samples, including the effect of the kernel. We conclude in Section 7. The proofs of all results are given in Appendix.

2. Framework and preliminary results

Following the notations of Saikkonen (1992), Saikkonen and Lütkepohl (1996) and Bouhaddioui and Dufour (2008), we consider a d-dimensional process $\mathbf{X} = \{X_t : t \in \mathbf{Z}\}$ partitioned into two subprocesses $X_i = \{X_{it} : t \in \mathbf{Z}\}$, $i = 1, 2$, with d_1 and d_2 components respectively ($d_1 + d_2 = d$). The data generating process has the form:

$$\mathbf{X}_{1t} = \mathbf{C}_1 \mathbf{X}_{2t} + \varepsilon_{1t} , \tag{2.1}$$

$$\Delta \mathbf{X}_{2t} = \varepsilon_{2t} , \tag{2.2}$$

where $\mathbf{C}_1$ is a fixed $d_1 \times d_2$ matrix, Δ is the usual difference operator, and $\varepsilon_t = (\varepsilon'_{1t}, \varepsilon'_{2t})'$ is a stationary process with zero mean and continuous spectral density matrix positive definite at frequency zero. X_{2t} is an integrated vector process of order one (with no cointegrating relationship), while X_{1t} and X_{2t} are cointegrated. By taking first differences in (2.1), we see that

$$\Delta \mathbf{X}_t = \begin{bmatrix} -\mathbf{I}_{d_1} & \mathbf{C}_1 \end{bmatrix} \mathbf{X}_{t-1} + v_t = \mathbf{J}\, \mathbf{\Theta}' \mathbf{X}_{t-1} + v_t \tag{2.3}$$

where $\mathbf{I}_d$ is the identity matrix of order d, $\mathbf{J}' = [-\mathbf{I}_{d_1} \vdots\,]$, $\mathbf{\Theta}' = [\mathbf{I}_{d_1} \vdots -\mathbf{C}_1]$, $v_t = (v'_{1t}, v'_{2t})'$ is a nonsingular transformation of ε_t defined by

$$v_{1t} = \varepsilon_{1t} + \mathbf{C}_1 \varepsilon_{2t} , \quad v_{2t} = \varepsilon_{2t} , \tag{2.4}$$

$$\mathbf{X}_t := \mathbf{X}_{1t} \quad , \quad v_t := v_{1t} \quad . \tag{2.5}$$

The notation $\mathbf{A} = [\mathbf{A}_1 \; \vdots \; \mathbf{A}_2]$ means that the matrix $\mathbf{A}$ is partitioned into a matrix $\mathbf{A}_1$ consisting of the first d_1 columns and a matrix $\mathbf{A}_2$ with d_2 columns.

We suppose that v_t (hence also ε_t) has an infinite-order autoregressive representation

$$\sum_{l=0}^{\infty} \mathbf{G}_l v_{t-l} = \mathbf{a}_t$$

where $\mathbf{G}_0 = \mathbf{I}_d$, $\mathbf{a}_t$ is a sequence of independent and identically distributed ran
that $\mathbf{E}(\mathbf{a}_t) = 0$ and $\mathbf{E}(\mathbf{a}_t\mathbf{a}_t') = \mathbf{\Sigma_a}$ is positive definite, and the roots of the e

$$\det\{\mathbf{I}_d - \sum_{l=1}^{\infty} \mathbf{G}_l z^l\} = 0$$

all lie outside the unit circle $|z| = 1$; $\det\{\mathbf{A}\}$ denotes the determinant of the square matrix $\mathbf{A}$. We also assume that the following summability condition holds:

$$\sum_{l=1}^{\infty} l^{\delta} \left\| \mathbf{G}_l \right\| < \infty \quad \text{for some } \delta \geq 1 \tag{2.8}$$

where $\|\cdot\|$ is the Euclidean matrix norm defined by $\|\mathbf{A}\|^2 = \mathrm{tr}(\mathbf{A}'\mathbf{A})$. This is a standard condition for weakly stationary processes, which ensures that the process is well defined. It also implies that the process v_t and, consequently X_t, can be approximated by an autoregression of finite order $p_n = p(n)$ where n is the sample size and p_n can grow with n. More explicitly, we assume that p_n satisfies the following condition.

Assumption 2.1 *There is a sequence of positive integers* p_n *such that*

$$n^{-1/3} p_n \to 0 \quad \textit{and} \quad \sqrt{p_n} \sum_{l=p_n+1}^{\infty} \left\| \mathbf{G}_l \right\| \to 0 \quad \textit{as } n \to \infty \, . \tag{2.9}$$

The condition $p_n = o(n^{1/3})$ for the rate of increase of p_n ensures that enough sample informa-tion is asymptotically available for estimators to have standard limiting distributions. The condition $\sqrt{p_n} \sum_{j=p_n+1}^{\infty} \|\mathbf{G}_j\| \to 0$ imposes a lower bound on the growth rate of p_n, 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 Burnham and Anderson (2002) and Lütkepohl (2005).

Using the equations (2.3) - (2.6) and rearranging terms, we obtain the autoregressive *error* *correction model* (ECM) representation

$$\Delta \mathbf{X}_t = \mathbf{\Psi}\,\mathbf{\Theta}'\mathbf{X}_{t-1} + \sum_{l=1}^{p_n} \mathbf{\Pi}_l \Delta \mathbf{X}_{t-l} + \mathbf{e}_t(n),\ t = p_n + 1, p_n + 2, \ldots, \tag{2.10}$$

$$\mathbf{e}_t(n) = \mathbf{a}_t - \sum_{l=p_n+1}^{\infty} \mathbf{G}_l v_{t-l}, \quad \mathbf{\Psi} = -\sum_{l=0}^{p_n} \mathbf{G}_l \mathbf{J}, \tag{2.11}$$

where $\mathbf{\Psi}$ is a $d \times d_1$ a full-column rank matrix (at least for p_n large enough). Details for this derivation can be found in Saikkonen and Lütkepohl (1994) and Saikkonen and Luukkonen (1997). Note the coefficient matrices $\mathbf{\Pi}_l$ ($l = 1, \ldots, p_n$) are functions of $\mathbf{\Theta}$ and $\mathbf{G}_l$ ($l = 1, 2, \ldots$), and they depend on p_n. Furthermore, the sequence $\mathbf{\Pi}_l$ ($l = 1, \ldots, p_n$) is absolutely summable as $p_n \to \infty$.

The autoregressive ECM in (2.10) can be rewritten in a pure vector autoregressive (VAR) form

$$\mathbf{X}_t = \sum_{l=1}^{p_n+1} \mathbf{\Phi}_l \mathbf{X}_{t-l} + \mathbf{e}_t(n) \tag{2.12}$$

where $\mathbf{\Phi}_1 = \mathbf{I}_d + \mathbf{\Psi}\,\mathbf{\Theta}' + \mathbf{\Pi}_1$, $\mathbf{\Phi}_l = \mathbf{\Pi}_l - \mathbf{\Pi}_{l-1}$, $l = 2, \ldots, p_n$ and $\mathbf{\Phi}_{p_n+1} = -\mathbf{\Pi}_{p_n}$. Although the $\mathbf{\Pi}_l$ depend on p_n, the same is not true for the $\mathbf{\Phi}_l$ except for $\mathbf{\Phi}_{p_n+1}$.

Saikkonen and Lütkepohl (1996) derived the asymptotic properties of the multivariate least square (LS) estimators of the VAR coefficients under a standard assumption. Let

$$\mathbf{\Phi}(p_n) = [\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_{p_n}] \tag{2.13}$$

be the matrix of the first p_n autoregressive parameter matrices in the representation (2.12), and denote by $\hat{\mathbf{\Phi}}(p_n) = [\hat{\mathbf{\Phi}}_1, \ldots, \hat{\mathbf{\Phi}}_{p_n}]$ the corresponding LS estimator. The following proposition gives a direct result on the asymptotic properties of the estimator $\hat{\mathbf{\Phi}}(p_n)$. It can be proved using the techniques very similar to those used by Saikkonen (1992, part (i) of Theorem 3.2); see also Saikkonen and Lütkepohl (1996, Theorem 2).

Proposition 2.1 ASYMPTOTIC PROPERTIES OF THE AUTOREGRESSIVE PARAMETER ESTIMATORS. *Let* $\{X_t\}$ *be a process which satisfies* (2.3) - (2.6) *with*

$$\mathbf{E}|a_{it}a_{jt}a_{kt}a_{lt}| < \gamma_4 < \infty, \qquad 1 \le i, j, k, l \le d. \tag{2.14}$$

where $\mathbf{a}_t := (a_{1t}, \ldots, a_{dt})'$. *If Assumption* 2.1 *holds, then*

$$\hat{\mathbf{\Phi}}(p_n) - \mathbf{\Phi}(p_n) = O_p(p_n^{1/2}/n^{1/2}). \tag{2.15}$$

This proposition is formulated for the first p_n coefficient matrices, whereas the fitted model is a VAR($p_n + 1$) where p_n goes to infinity with the sample size n. Dropping the last lag in deriving the consistency of the estimators will not affect the asymptotic distribution of the test statistic; see Lütkepohl (2005). Details on the estimates of the $\mathbf{\Phi}_l$ matrices are given in Saikkonen and Lütkepohl (1996). This result can be viewed as a generalization of Theorem 1 in Lewis and Reinsel (1985) to infinite-order stationary vector autoregressive processes.

Let us now consider two processes $\mathbf{X}^{(h)} = \{\mathbf{X}_t^{(h)} : t \in \mathbf{Z}\}$, $h = 1, 2$, with m_1 and m_2 components respectively, each of which satisfies an IVAR(∞) model of the form (2.3) - (2.6) with $m_h = d_1^{(h)} + d_2^{(h)}$, $h = 1, 2$, where $d_1^{(h)}$ and $d_2^{(h)}$ replace d_1 and d_2 for $\mathbf{X}^{(h)}$. The coefficients of the two processes may differ. We wish to decide whether $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ are independent against an alternative where they are correlated at some lag. Following Pham et al. (2003), the independence between $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ can be tested by testing non-correlation between the corresponding innovation processes $\mathbf{a}^{(1)}$ and $\mathbf{a}^{(2)}$. This leads one to consider the hypothesis:

$$H_0 : \rho_{\mathbf{a}}^{(12)}(j) = 0 \text{ , for all } j \in \mathbf{Z} \text{ ,} \tag{2.16}$$

where

$$\rho_{\mathbf{a}}^{(12)}(j) = [\mathbf{D}(\mathbf{\Sigma})]^{-1/2} \ {}_{\mathbf{a}}^{(12)}(j)[\mathbf{D}(\mathbf{\Sigma})]^{-1/2}, \quad {}_{\mathbf{a}}^{(hi)}(j) = \mathbf{E}\, \mathbf{a}_t^{(h)} \ \mathbf{a}_{t-j}^{(i)} \ ', \ j \in \mathbf{Z} \text{ ,} \tag{2.17}$$

$$\mathbf{\Sigma}_h = \ {}_{\mathbf{a}}^{(hh)}(0) \text{ , } \mathbf{D}(\mathbf{\Sigma}_h) = \operatorname{diag}\{\mathbf{\Sigma}_h\} \text{ , } \mathbf{\Sigma} = \mathbf{\Sigma}_1 \qquad \text{, } h, i = 1, 2 \text{ .} \tag{2.18}$$

$\rho^{(12)}_{\mathbf{a}}(j)$ represents the cross-correlation matrix at lag j between the two innovation processes. On setting

$$\mathbf{b}_t := \mathbf{\Sigma}^{-1/2}\mathbf{a}_t = \begin{pmatrix} \mathbf{\Sigma}_1^{-1/2} & \\ & \mathbf{\Sigma}_2^{-1/2} \end{pmatrix} \begin{pmatrix} \mathbf{a}_{1t} \\ \mathbf{a}_{2t} \end{pmatrix} = \begin{pmatrix} \mathbf{\Sigma}_1^{-1/2}\mathbf{a}_{1t} \\ \mathbf{\Sigma}_2^{-1/2}\mathbf{a}_{2t} \end{pmatrix} = \begin{pmatrix} \mathbf{b}_{1t} \\ \mathbf{b}_{2t} \end{pmatrix} \text{ ,} \tag{2.19}$$

we see that $\rho_{\mathbf{b}}^{(12)}(j) = {}_{\mathbf{b}}^{(12)}(j) = \rho_{\mathbf{b}}^{(12)}(j)$, for all $j \in \mathbf{Z}$, so that H_0 is equivalent to

$$\rho_{\mathbf{b}}^{(12)}(j) = 0 \text{ , for all } j \in \mathbf{Z} \text{ .} \tag{2.20}$$

This equivalence plays a central role for proving the required distributional results stated below.

3. Test statistics and asymptotic null distributions

Based on a realization $X_1^{(h)}, \ldots, X_n^{(h)}$ of length n, for h = 1, 2, a finite-order autoregressive model VAR($p_n^{(h)} + 1$) is fitted to each one of these two series. The order $p_n^{(h)}$ depends on the sample size n. The resulting residuals are given by

$$\hat{a}^{(h)}_t(n) = \begin{cases} X^{(h)}_t - \sum_{l=1}^{p^{(h)}_n} \hat{\Phi}^{(h)}_l(n) X^{(h)}_{t-l} & \text{if } t = p^{(h)}_n + 2, \ldots, n, \\ 0 & \text{if } t \le p^{(h)}_n + 1, \end{cases} \tag{3.1}$$

where the matrices $\hat{\Phi}^{(h)}_l(n)$ are the OLS estimators of $\Phi^{(h)}_l(n)$, and h = 1, 2. We can also use the conditional maximum likelihood estimator of the error correction form of the model as discussed by Ahn and Reinsel (1990) and Reinsel (1993), or some other estimator with the same rate of convergence. We now consider the residual sample (cross-)covariance matrices

$$\mathbf{C}^{(hi)}_{\hat{a}}(j) = \begin{cases} n^{-1} \sum_{t=j+1}^{n} \hat{\mathbf{a}}^{(h)}_t (\hat{\mathbf{a}}^{(i)}_{t-j})' & \text{if } 0 \le j \le n-1 \\ n^{-1} \sum_{t=-j+1}^{n} \hat{\mathbf{a}}^{(h)}_{t+j} (\hat{\mathbf{a}}^{(i)}_t)' & \text{if } -n+1 \le j \le 0 \end{cases} \tag{3.2}$$

where h, i = 1, 2, and the corresponding cross-correlation matrices

$$\mathbf{R}^{(hi)}_{\hat{a}}(j) = [\mathbf{D}\,\mathbf{C}^{(hh)}_{\hat{a}}(0)]^{-1/2}\, \mathbf{C}^{(hi)}_{\hat{a}}(j)\, [\mathbf{D}\,\mathbf{C}^{(ii)}_{\hat{a}}(0)]^{-1/2} \tag{3.3}$$

where $\mathbf{D}\, C^{(hh)}_{\hat{\mathbf{a}}}(0) = \mathrm{diag}\{C^{(hh)}_{\hat{\mathbf{a}}}(0)\}$. The orthogonality tests we consider are based on $C^{(12)}_{\hat{\mathbf{a}}}(j)$ and $\mathbf{R}^{(12)}_{\hat{\mathbf{a}}}(j)$. In the sequel, we suppose that $X^{(h)}$ satisfies (2.3) for h = 1, 2. We wish to test the null hypothesis H_0 using the cross-correlation matrices $\mathbf{R}^{(hi)}_{\hat{\mathbf{a}}}(j)$, $j \in \mathbf{Z}$.

In the univariate case, Hong (1996*b*) proposed a portmanteau-type statistic based on the sum of the weighted squared cross-correlations $r^{(12)}_{\hat{a}}(j)$ at all possible lags between the residual series:

$$Q_n = \frac{n \sum_{j=1-n}^{n-1} k^2(j/M)\, r^{(12)}_{\hat{a}}(j)^2 - S_n(k)}{\{2D_n(k)\}^{1/2}} \tag{3.4}$$

where k(·) is an arbitrary kernel function [see Table 2 for examples] and M is a smoothing param-eter, while $S_n(k)$ and $D_n(k)$ are normalization coefficients which depend on the kernel k(·):

$$S_n(k) = \sum_{j=1-n}^{n-1} \left(1 - \frac{|j|}{n}\right) k^2(j/M), \quad D_n(k) = \sum_{j=2-n}^{n-2} \left(1 - \frac{|j|}{n}\right)\left(1 - \frac{|j|+1}{n}\right) k^4(j/M). \tag{3.5}$$

They correspond to the asymptotic mean and variance of the weighted sum. In multivariate time series, the squared cross-correlation $r^{(12)}_{\hat{a}}(j)^2$ in (3.4) is replaced by a quadratic form in the vector $\mathbf{r}^{(12)}_{\hat{\mathbf{a}}}(j) = \mathrm{vec}[\mathbf{R}^{(12)}_{\hat{\mathbf{a}}}(j)]$. For H_0, the test statistic is based on the following sum of weighted quadratic forms at all possible lags:

$$T(\hat{\mathbf{a}}, \hat{\boldsymbol{\Sigma}}) = \sum_{j=1-n}^{n-1} k^2(j/M)\, Q^{(12)}_{\hat{\mathbf{a}}}(j), \tag{3.6}$$

$$Q^{(12)}_{\hat{a}}(j) := n\, \mathbf{r}^{(12)}_{\hat{a}}(j)' \left[\mathbf{R}^{(22)}_{\hat{a}}(0)^{-1} \otimes \mathbf{R}^{(11)}_{\hat{a}}(0)^{-1}\right] \mathbf{r}^{(12)}_{\hat{a}}(j), \quad \hat{\boldsymbol{\Sigma}} := \hat{\boldsymbol{\Sigma}}_1, \tag{3.7}$$

where $\mathbf{R}_{\hat{\mathbf{a}}}^{(hh)}(0)$ is a consistent estimator of the correlation matrix $\rho_{\mathbf{a}}^{(h)}$ of the process $\mathbf{a}^{(h)}$, and $k(\cdot)$ is a suitable kernel function. The parameter M is a truncation point when the kernel has compact support, or a smoothing parameter when the kernel support is unbounded. We suppose that M is function of n ($M = M_n$) such that $M_n \to \infty$ and $M_n/n \to 0$ as $n \to \infty$. The most commonly used kernels typically give more weight to lower lags and less weight to higher ones. An exception is the truncated uniform kernel $k_T(z) = \mathbf{I}[|z| \le 1]$, where $\mathbf{I}(A)$ represents the indicator function of the set A, which gives the same weight to all lags. The asymptotic distribution of $Q_{\hat{\mathbf{a}}}(j)$ is given in Bouhaddioui and Dufour (2008). In the sequel, we suppose that the kernel function k and the order $p_n^{(h)}$ respectively satisfy the following assumptions.

Assumption 3.1 *The kernel function* $k : \mathbf{R} \to [-1, 1]$ *is a symmetric function, continuous at zero, with at most a finite number of discontinuity points, such that* $k(0) = 1$ *and* $\int_{-\infty}^{+\infty} k^2(z)dz < \infty$.

Assumption 3.2 *The orders* $p_n^{(h)}$, $h = 1, 2$*, satisfy the following conditions:*

$$\text{(i)}\quad p_n^{(h)} = o\left(n^{1/2}/M^{1/4}\right), \quad \text{(ii)}\quad n\sum_{j=p_n^{(h)}+1}^{\infty} \|\mathbf{\Phi}_j^{(h)}\|^2 = o\left(n^{1/2}/M^{1/4}\right). \tag{3.8}$$

Note that the two conditions (i) and (ii) imply that the order $p_n^{(h)}$ satisfies Assumption 2.1. The property $k(0) = 1$ implies that the weights assigned to the lower lags are close to unity. The square integrability of $k(\cdot)$ implies that $k(z) \to 0$ as $|z| \to \infty$, so that less weight is given to $\hat{\mathbf{R}}_{\mathbf{a}}^{(12)}(j)$ as j increases. Note that all the kernels used in spectral analysis satisfy Assumption 3.1; see Priestley (1981, Section 6.2.3). For hypothesis H_0, the test statistic is a standardized version of $T(\hat{\mathbf{a}}, \hat{\mathbf{\Sigma}})$:

$$Q_n = \frac{T(\hat{\mathbf{a}}, \hat{\mathbf{\Sigma}}) - m_1 m_2 S_n(k)}{2 m_1 m_2 D_n(k)}, \tag{3.9}$$

where the smoothing parameter $M_n \to \infty$ and $M_n/n \to 0$ when $n \to \infty$.

This test statistic can be viewed as a normalized version of the L_2-norm of a kernel-based estimator of the cross-coherency function between the two innovation series. $S_n(k)$ and $D_n(k)$ represent the asymptotic mean and variance of $T(\hat{\mathbf{a}}, \hat{\mathbf{\Sigma}})$ under H_0. If $k(\cdot)$ is the truncated uniform kernel, apart from the standardization factors $S_n(k)$ and $D_n(k)$, Q_n corresponds to the multivariate version of Haugh's statistic used by Pham et al. (2003) for the finite-order cointegrated case, and by Bouhaddioui and Dufour (2008) for the infinite-order case, namely

$$P_M = \sum_{j=-M}^{M} Q_{\hat{a}}(j) . \qquad (3.10)$$

In this case, M is a fixed integer that does not depend on the sample size n. The properties of P_M in the stationary VAR(∞) context and cointegrated IVAR(∞) are studied respectively in Bouhaddioui and Roy (2006*b*) and Bouhaddioui and Dufour (2008). As it will be seen below, many kernels k yield tests that are more powerful than P_M .

In the case of testing independence, under some conditions on the smoothing parameter M and if the kernel k verifies Assumption 3.1, one sees easily that

$$M^{-1}S_n(k) \to S(k) , \; M^{-1}D_n(k) \to D(k), \qquad (3.11)$$

where

$$S(k) = \int_{-\infty}^{+\infty} k^2(z)dz \;, \; D(k) = \int_{-\infty}^{+\infty} k^4(z)dz . \qquad (3.12)$$

An alternative statistic is obtained by replacing $S_n(k)$ and $D_n(k)$ by their asymptotic approximations M S(k) an M D(k) respectively and is defined by

$$Q_n^* = \frac{T(\hat{a}, \hat{\Sigma}) - M m_1 m_2 S(k)}{\sqrt{2M m_1 m_2 D(k)}} . \qquad (3.13)$$

Both Q_n and Q^*_n have the same asymptotic null distribution and power properties.

The statistic Q_n can also be expressed in term of the autocovariances $C_{\hat{a}}^{(hh)}(0)$ and the cross-covariances $C_{\hat{a}}^{(12)}(j)$ of the same residual series. Invoking Lemma 4.1 of El Himdi and Roy (1997), the quadratic form $T(\hat{a}, \hat{\Sigma})$ can be written as follows in terms of the residual covariances:

$$T(\hat{a}, \hat{\Sigma}) = n \sum_{j=1-n}^{n-1} k^2(j/M) c_{\hat{a}}^{(12)}(j)' \left[C_{\hat{a}}^{(22)}(0)^{-1} \otimes C_{\hat{a}}^{(11)}(0)^{-1} \right] c_{\hat{a}}^{(12)}(j) \qquad (3.14)$$

with $c_{\hat{a}}^{(12)}(j) = \mathrm{vec}[C_{\hat{a}}^{(12)}(j)]$. Let us now consider the "pseudo-statistic"

$$T(a, \Sigma) = n \sum_{j=1-n}^{n-1} k^2(j/M) c_a^{(12)}(j)' \left[\Sigma_2^{-1} \otimes \Sigma_1^{-1} \right] c_a^{(12)}(j) \qquad (3.15)$$

where $c_a^{(12)}(j)$ is defined as $c_{\hat{a}}^{(12)}(j)$ with the residuals $\hat{a}_t^{(1)}$ and $\hat{a}_t^{(2)}$ replaced by the unobservable innovation series $a_t^{(1)}$ and $a_t^{(2)}$, t = 1, . . . , n, and

$$T(\hat{a}, \Sigma) = n \sum_{j=1-n}^{n-1} k^2(j/M) c_{\hat{a}}^{(12)}(j)' (\Sigma_2^{-1} \otimes \Sigma_1^{-1}) c_{\hat{a}}^{(12)}(j) . \qquad (3.16)$$

Thus, with $\hat{\Sigma}_h = C_{\hat{a}}^{(hh)}(0)$, h = 1, 2, we can write the statistic Q_n as

$$Q_n = \frac{T(\hat{a}, \hat{\Sigma}) - m_1 m_2 S_n(k)}{\sqrt{2 m_1 m_2 D_n(k)}}$$

$$= \frac{T(\mathbf{a}, \mathbf{\Sigma}) - m_1 m_2 S_n(k)}{2m_1 m_2 D_n(k)} + \frac{T(\hat{\mathbf{a}}, \mathbf{\Sigma}) - T(\mathbf{a}, \mathbf{\Sigma})}{2m_1 m_2 D_n(k)} + \frac{T(\hat{\mathbf{a}}, \hat{\mathbf{\Sigma}}) - T(\hat{\mathbf{a}}, \mathbf{\Sigma})}{2m_1 m_2 D_n(k)} . \quad (3.17)$$

Since the quantity $T(\mathbf{a}, \mathbf{\Sigma})$ depends only on the stationary process $\mathbf{a}$, the result of Lemma 3.1 in Bouhaddioui and Roy (2006*a*) is still valid. We conclude that

$$\frac{T(\mathbf{a}, \mathbf{\Sigma}) - m_1 m_2 S_n(k)}{2m_1 m_2 D_n(k)} \xrightarrow{L} N$$

The asymptotic distribution of Q_n follows from the next two propositions.

Proposition 3.1 APPROXIMATION OF THE PSEUDO-STATISTIC. *Suppose* $X^{(1)} = \{X^{(1)}_t : t \in \mathbf{Z}\}$ *and* $X^{(2)} = \{X^{(2)}_t : t \in \mathbf{Z}\}$ *satisfy the IVAR*(∞) *model* (2.3) - (2.6) *along with Assumption* 3.1 *and the bounded moment condition*

$$\mathbf{E}|a_{it}^{(h)} a_{jt}^{(h)} a_{kt}^{(h)} a_{lt}^{(h)}| < \gamma_4 < \infty, \ 1 \le i, j, k, l \le m_h . \quad (3.19)$$

Let $M = M_n$, *with* $M_n \to \infty$ *and* $M_n/n \to 0$ *as* $n \to \infty$, *and suppose that* $p_n^{(h)}$, $h = 1, 2$, *satisfy Assumption* 3.2. *If the processes* $\mathbf{a}^{(1)}$ *and* $\mathbf{a}^{(2)}$ *are independent, then*

$$T(\hat{\mathbf{a}}, \mathbf{\Sigma}) - T(\mathbf{a}, \mathbf{\Sigma}) = o_p(M^{1/2}) . \quad (3.20)$$

Proposition 3.2 ASYMPTOTIC EQUIVALENCE OF THE TEST STATISTIC. *Under the assumptions of Proposition* 3.1, *we have*

$$\frac{T(\hat{\mathbf{a}}, \hat{\mathbf{\Sigma}}) - T(\hat{\mathbf{a}}, \mathbf{\Sigma})}{2m_1 m_2 D_n(k)} \to 0 . \quad (3.21)$$

Our main result is a simple consequence of Propositions 3.1 - 3.2, as follows.

Theorem 3.3 NULL ASYMPTOTIC DISTRIBUTION. *Under the assumptions of Proposition* 3.1, *the statistic* Q_n *defined by* (3.9) *has an asymptotic* $N(0, 1)$ *distribution,* i.e. $Q_n \xrightarrow{L} N(0, 1)$.

4. Consistency of the generalized tests

We now investigate the asymptotic power of the test Q_n under fixed alternatives. We consider a fixed alternative H_1 of serial cross-correlation between the two innovation processes $\mathbf{a}^{(1)}$ and $\mathbf{a}^{(2)}$ with the following assumption.

Assumption 4.1 *The two innovation processes*

$$\mathbf{a}_t^{(1)} = (a_{1,t}^{(1)}, \ldots, a_{m_1,t}^{(1)})' \text{ and } \mathbf{a}_t^{(2)} = (a_{1,t}^{(2)}, \ldots, a_{m_2,t}^{(2)})', \quad t \in \mathbf{Z}, \quad (4.1)$$

are jointly fourth-order stationary, and their cross-correlation structure is such that $\mathbf{a}^{(12)}(j) =$ *for at least one value of* j, *with*

$$\sum_{j=-\infty}^{+\infty} \left\| \mathbf{a}^{(12)}(j) \right\|^2 < \infty , \qquad \sum_{i=-\infty}^{+\infty}\sum_{j=-\infty}^{+\infty}\sum_{l=-\infty}^{+\infty} |\kappa_{uvuv}(0, i, j, l)| < \infty, \tag{4.2}$$

where $\kappa_{uvuv}(0, i, j, l)$ *is the fourth cumulant of the joint distribution of* $a^{(1)}_{u,t}, a^{(2)}_{v,t+i}, a^{(1)}_{u,t+j}, a^{(2)}_{v,t+l}$.

The following theorem gives conditions for the consistency of Q_n under a fixed alternative.

Theorem 4.1 GLOBAL POWER. *Let* $X^{(1)}$ *and* $X^{(2)}$ *be two multivariate processes which follow the IVAR*(∞) *model* (2.3) - (2.6)*, and suppose that their innovation processes* $\mathbf{a}^{(1)}$ *and* $\mathbf{a}^{(2)}$ *satisfy Assumption* 4.1*. If the kernel* $k(\cdot)$ *satisfies Assumption 3.1 and if* $p_n^{(h)}$, $h = 1, 2$*, satisfy*

$$p_n^{(h)2} = o\left(\frac{n}{M}\right) , \qquad \sum_{j=p_n^{(h)}+1}^{\infty} \left\| \Phi_j^{(h)} \right\|^2 = o\left(M^{-1}\right) , \tag{4.3}$$

then, for any sequence of constants $C(n, M)$ *such that* $C(n, M) = o(n/M^{1/2})$,

$$P\left[Q_n > C(n, M)\right] \to 1 . \tag{4.4}$$

This theorem entails that the test based on Q_n is consistent against the general class of dependence alternatives described by Assumption 4.1. The slower M grows, the faster Q_n goes to infinity. To investigate the relative efficiency of Q_n, one can use the Bahadur's asymptotic slope criterion defined in Bahadur (1960); see also Hong (1996*a*, 1996*b*) and Bouhaddioui and Roy (2006*a*). As in Bouhaddioui and Roy (2006*a*), we can show that the relative efficiency of the kernel $k_2(\cdot)$ with respect to $k_1(\cdot)$ when $M = n^{\nu}$ is given by

$$ARE_B(k_2, k_1) = \left(\frac{D(k_1)}{D(k_2)}\right)^{1/(2-\nu)} . \tag{4.5}$$

We can then proceed like Bouhaddioui (2002) and Hong (1996*a*, 1996*b*) to derive the kernel that maximizes the asymptotic slope over appropriate classes of kernel functions. For example, consider the following class of kernels:

$$\kappa(\tau) = \{k(\cdot) : \text{Assumption 3.1 is satisfied, } k^{(2)} = \tau^2/2,\ K(\lambda) \geq 0 \text{ for } \lambda \in (-\infty, +\infty)\} \tag{4.6}$$

where

$$k^{(2)} = \lim_{z\to 0} [1 - k(z)]/z^2 \text{ and } K(\lambda) = \frac{1}{2\pi}\int_{-\infty}^{\infty} k(z)e^{-iz\lambda}\, dz . \tag{4.7}$$

This class contains the Daniell, Parzen and quadratic-spectral kernels (among others). Using Theorem 1 of Ghosh and Huang (1991), we can see that the Daniell kernel [see Table 2] maximizes the asymptotic slope of Q_n over $\kappa(\tau)$; for a similar argument, see Bouhaddioui (2002). As mentioned in Bouhaddioui and Roy (2006*a*), a test with a greater asymptotic slope may be expected to have a greater power for a fixed alternative than one with a smaller asymptotic slope. However, there is no clear analytical relationship between the slope of a test and its power function. For a specific alternative, we cannot conclude that a test with greater asymptotic slope should be automatically preferred to one with a smaller asymptotic slope without further analysis of the finite-sample prop-erties of the two test statistics.

5. Local power analysis

In this section ,we study the power of the test proposed above against a class of local alternatives of the form

$$H_{\mathbf{a}}(\boldsymbol{\Lambda}^{(12)}_{\mathbf{b}}) : \quad \mathbf{b}^{(12)}(j) = \frac{M^{1/4}}{n^{1/2}} \boldsymbol{\Lambda}^{(12)}_{\mathbf{b}}(j) \text{ , for all } j \in \mathbf{Z} \text{ ,}$$

where $\boldsymbol{\Lambda}^{(12)}_{\mathbf{b}} = \{\boldsymbol{\Lambda}^{(12)}_{\mathbf{b}}(j)\}_{j\in\mathbf{Z}}$ is a sequence of $m_1 \times m_2$ cross-correlation matrices such that only finite elements of $\boldsymbol{\Lambda}^{(12)}_{\mathbf{b}}$ are non-zero elements. Let

$$\lambda_{\mathbf{b}}^{(12)}(j) = \operatorname{vec}[\boldsymbol{\Lambda}_{\mathbf{b}}^{(12)}(j)] \text{ ,} \tag{5.1}$$

$$\beta(\boldsymbol{\Lambda}_{\mathbf{b}}^{(12)}) = \sum_{j=-\infty}^{\infty} \lambda_{\mathbf{b}}^{(12)}(j)' \lambda_{\mathbf{b}}^{(12)}(j). \tag{5.2}$$

The following theorem establishes the asymptotic distribution of Q_n under the local alternative $H_{\mathbf{a}}(\boldsymbol{\Lambda}^{(12)}_{\mathbf{b}})$.

Theorem 5.1 LOCAL POWER. *Let* $X^{(1)}$ *and* $X^{(2)}$ *be two multivariate processes which follow the IVAR*(∞) *model* (2.3) - (2.6)*, and suppose that their innovation processes* $\mathbf{a}^{(1)}$ *and* $\mathbf{a}^{(2)}$ *satisfy Assumption* 4.1*. If the kernel* $k(\cdot)$ *satisfies Assumption* 3.1 *and if* $p_n^{(h)}$, $h = 1, 2$*, satisfy*

$$\left(p_n^{(h)}\right)^2 = o(n/M) \text{ ,} \qquad \sum_{j=p_n^{(h)}+1}^{\infty} \left\|\boldsymbol{\Phi}_j^{(h)}\right\|^2 = o\left(M^{-1}\right) \text{ ,} \tag{5.3}$$

then, under $H_{\mathbf{a}}(\boldsymbol{\Lambda}_{\mathbf{b}}^{(12)})$*,*

$$Q_n \xrightarrow{L} N\left[\beta(\boldsymbol{\Lambda}_{\mathbf{b}}^{(12)})/\sqrt{2m_1m_2D(k)}, 1\right]. \tag{5.4}$$

where $\beta(\boldsymbol{\Lambda}^{(12)}_{\mathbf{b}})$ *is defined in* (5.2)*.*

Theorem 5.1 shows that the test Q_n has non-trivial power against a class of local alternatives converging to H_0 at the rate of $\frac{M^{1/4}}{n^{1/2}}$. The power depends on the kernel function k through D(k).

Similarly, we note that increasing slowly the parameter M , the divergence of the test statistic to infinity is faster and consequently, the test is more powerful.

6. Simulation study

In the previous sections, we have studied the asymptotic distribution of the test statistics. Here we investigate the finite-sample properties of the proposed test statistics, in particular their exact level and power. To do this, we performed a small Monte Carlo study. In addition to the test statistics discussed in the preceding sections, we also consider the nonstationary multivariate version of the Haugh statistic [previously studied by Pham et al. (2003)]:

$$P_M^* = \sum_{j=-M}^{M} \frac{n}{n-|j|} Q_{\hat{\mathbf{a}}}(j) \tag{6.1}$$

where $Q^{(12)}_{\hat{\mathbf{a}}}(j)$ is given by (3.7). P_M^* is a slightly modified version of P_M defined by (3.10).

6.1. Description of the experiment

In the simulation experiment, we considered bivariate series $\{X^{(1)}_t\}$ and $\{X^{(2)}_t\}$ generated from (joint) 4-dimensional **VAR**(2), **VARMA**(1,1) and VAR_δ (1) models (see Table 1). In the first two models, the two subprocesses $X^{(1)}$ and $X^{(2)}$ are independent bivariate **VAR**(2) or **VARMA**(1,1) and served for the level study and the corresponding submodels are partially nonstationary and invertible. The third one, in which there is instantaneous correlation between the two innovation series, was used for the power study. The correlation depends on a parameter δ and the values δ = 1.0, 1.5 and 2 were chosen. For each model, two series lengths (n = 100, 200) were considered. With the statistics Q_n and Q^*_n defined by (3.9) and (3.13), we used the four kernels described in Table 2. For each kernel, the following three truncation values M were employed: M = [ln(n)], $[3n^{0.2}]$ and $[3n^{0.3}]$ ([a] denotes the integer part of a). These rates are discussed in Hong (1996*a*, p. 849). They lead respectively to M = 5, 8, 12 for the series length n = 100, and to M = 5, 9, 15 for n = 200. The same truncation values were used for P_M^* .

In the level study, 5000 independent realizations were generated from both models **VAR**(2) and **VARMA**(1,1) for each series length n. Computations were made in the following way.

(1) First, pseudo-random variables from the N (0, 1) distribution were obtained with the pseudo-

Table 1. Time series models used in the simulation study

Models	Equations
VAR(2)	$\begin{pmatrix} x_t^{(1)} \\ x_t^{(2)} \end{pmatrix} = \begin{pmatrix} \Phi_1^{(1)} & \\ & \Phi_1^{(2)} \end{pmatrix} \begin{pmatrix} x_{t-1}^{(1)} \\ x_{t-1}^{(2)} \end{pmatrix} + \begin{pmatrix} \Phi_2^{(1)} & \\ & \Phi_2^{(2)} \end{pmatrix} \begin{pmatrix} x_{t-2}^{(1)} \\ x_{t-2}^{(2)} \end{pmatrix} + \begin{pmatrix} a_t^{(1)} \\ a_t^{(2)} \end{pmatrix}$
	$\mathbf{X}_t \quad \Phi_1 \quad \mathbf{X}_{t-1} \quad \Phi_2 \quad \mathbf{X}_{t-2} \quad \mathbf{a}_t$
VARMA(1, 1)	$\begin{pmatrix} x_t^{(1)} \\ x_t^{(2)} \end{pmatrix} = \begin{pmatrix} \Phi_1^{(1)} & \\ & \Phi_1^{(2)} \end{pmatrix} \begin{pmatrix} x_{t-1}^{(1)} \\ x_{t-1}^{(2)} \end{pmatrix} + \begin{pmatrix} \Psi_{(1)} & \\ & \Psi_{(2)} \end{pmatrix} \begin{pmatrix} a_{t-1}^{(1)} \\ a_{t-1}^{(2)} \end{pmatrix} + \begin{pmatrix} a_t^{(1)} \\ a_t^{(2)} \end{pmatrix}$
	$\mathbf{X}_t \quad \Phi_1 \quad \mathbf{X}_{t-1} \quad \Psi \quad \mathbf{a}_{t-1} \quad \mathbf{a}_t$
VARδ (1)	$\begin{pmatrix} x_t^{(1)} \\ x_t^{(2)} \end{pmatrix} = \begin{pmatrix} \Psi_1^{(1)} & \\ & \Phi_1^{(2)} \end{pmatrix} \begin{pmatrix} x_{t-1}^{(1)} \\ x_{t-1}^{(2)} \end{pmatrix} + \begin{pmatrix} a_t^{(1)} \\ a_t^{(2)} \end{pmatrix}$
	$\mathbf{X}_t \quad \Phi_1 \quad \mathbf{X}_{t-1} \quad \mathbf{a}_t$

noise covariance matrices

$$\boldsymbol{\Sigma}_A = \begin{pmatrix} \boldsymbol{\Sigma}_{A(1)} & \\ & \boldsymbol{\Sigma}_A^{(2)} \end{pmatrix} \qquad \boldsymbol{\Sigma}_{A,\boldsymbol{\delta}} = \begin{pmatrix} \boldsymbol{\Sigma}_{A(1)} & \boldsymbol{\Sigma}_{A,\boldsymbol{\delta}}^{(12)} \\ \boldsymbol{\Sigma}_{A,\boldsymbol{\delta}}^{(21)} & \boldsymbol{\Sigma}_A^{(2)} \end{pmatrix}$$

Parameters values

$\Phi_1^{(1)} = \begin{pmatrix} 0.4 & 0.0 \\ -1.0 & 1.0 \end{pmatrix}$ $\quad \Phi_1^{(2)} = \begin{pmatrix} 1.0 & 0.0 \\ -0.8 & 0.5 \end{pmatrix}$ $\quad \Phi_2^{(1)} = \begin{pmatrix} 0.6 & -0.5 \\ 0.3 & 0.4 \end{pmatrix}$ $\quad \Phi_2^{(2)} = \begin{pmatrix} -0.5 & -0.8 \\ -0.4 & 0.2 \end{pmatrix}$

$\Psi_{(1)} = \begin{pmatrix} -0.2 & 0.3 \\ -0.6 & 1.1 \end{pmatrix}$ $\quad \Psi_{(2)} = \begin{pmatrix} 0.8 & 0.3 \\ 0.1 & 0.6 \end{pmatrix}$ $\quad \Sigma_{\mathbf{a}}^{(1)} = \begin{pmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{pmatrix}$ $\quad \Sigma_{\mathbf{a}}^{(2)} = \begin{pmatrix} 1.0 & 0.75 \\ 0.75 & 1.0 \end{pmatrix}$

$\Sigma_{\mathbf{a},\boldsymbol{\delta}}^{(12)} = \begin{pmatrix} 0.1\delta & 0 \\ 0 & 0.05\delta \end{pmatrix}$

Table 2. Kernels used with the test statistics Q_n and Q^*_n

Truncated Uniform (TR): $k(z) = \begin{cases} 1, & |z| \le 1, \\ 0, & \text{otherwise.} \end{cases}$

Bartlett (BAR): $k(z) = \begin{cases} 1 - |z|, & |z| \le 1, \\ 0, & \text{otherwise .} \end{cases}$

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

Parzen (PAR): $k(z) = \begin{cases} 1 - 6z^2 + 6|z|^3, & \text{if } |z| \le 0.5, \\ \ldots & \text{if } 0.5 \le |z| \le 1, \\ 0, & \text{otherwise.} \end{cases}$

Bartlett-Priestley (BP): $= \dfrac{3}{\ldots} \left(\dfrac{\sin(\pi z)}{\ldots} - \cos \ldots \right) \quad \mathbf{R}$

random normal generator of the S-plus package and were transformed into independent N [, $\mathbf{\Sigma_A}$] pseudo-random vectors using the Cholesky decomposition. Second, the $\mathbf{X}_t$ values were obtained by directly solving the model difference equation.

(2) For the **VAR**(2) model, the least squares estimates of the coefficients of the true models were obtained using the procedure described in Reinsel (1993). The autoregressive order was obtained by minimizing the AIC criterion for $p \le P$, where P is set to $n^{1/3}$. We chose AIC criterion which seems behave better than other criteria such HQ or SC specifically in LR cointegration tests, see Lütkepohl and Saikkonen (1999). For the **VARMA**(1,1), each subseries was approximated by a possible high-order **VAR** model. From Pham et al. (2003), the value of the **VAR** order was obtained by minimizing the Hannan-Quinn criterion using conditional least square estimation. The residual cross-correlation matrix $\mathbf{R_a}^{(12)}\hat{}(j)$'s as defined by (3.3) is then computed.

(3) For each realization, the test statistics Q_n and Q^*_n were compared for each of the four kernels and the three values of M . The same values of M were used for the statistic P_M^* . The values of the statistics Q_n and Q^*_n were compared with the N (0, 1) critical values and those of P_M^* to the $\chi^2_{4(2M\ +1)}$ critical values.

(4) Finally, for each model, each series length and nominal level, the empirical frequencies of rejection of the null hypothesis of non-correlation were obtained from the 5000 realizations. The results in percentage are reported in Table 3. The standard error of the empirical level is 0.14% for the nominal level 1%, 0.31% for 5% and 0.42% for 10%.

Computations for the power analysis were made in a similar way using the VAR_δ (1) model with different values of δ.

6.2. Level

6.2.1. Gaussian innovations

Results from the level study are presented in Table 3. We make the following observations. The asymptotic N (0, 1) distribution provides a good approximation of the exact distributions of Q_n and Q^*_n at all nominal levels considered, kernels and truncation values. Almost all empirical levels are within three standard errors of the corresponding nominal levels and the majority are within two standard errors. The statistic Q^*_n is slightly better approximated than Q_n since most of its empirical

Table 3. Empirical level (in percentage) of the Q_n, Q^*_n and $P_M{}^*$ tests,
with different kernels and truncation values.
Gaussian **VAR**(2) and **VARMA**(1, 1) models. Number of realizations: 5000

				Q_n					$Q_n{}^*$					$P_M{}^*$
	n	M	α%	DAN	PAR	BAR	BP	TR	DAN	PAR	BAR	BP	TR	
VAR(2)	100	5	1	0.7	0.6	0.8	0.7	0.6	1.2	0.9	0.7	0.7	1.3	0.7
			5	5.8	3.9	5.7	5.2	4.4	5.9	4.3	5.8	6.1	3.7	4.2
			10	9.6	8.0	9.5	10.6	8.3	10.3	8.8	9.4	10.7	9.0	8.8
		8	1	1.3	0.6	0.9	1.2	0.7	1.4	1.2	1.0	1.5	0.6	0.8
			5	5.6	4.1	5.9	5.6	4.0	5.4	4.0	5.2	4.8	4.0	4.3
			10	10.7	9.2	10.8	10.7	7.4	10.6	9.6	11.0	10.4	8.2	8.4
		12	1	0.8	0.7	0.8	1.2	0.6	1.3	0.8	1.4	1.5	0.7	0.8
			5	5.4	4.8	5.3	5.4	4.2	5.6	4.5	4.9	5.7	4.2	4.5
			10	10.4	8.7	11.2	10.8	7.8	10.8	10.4	11.2	10.5	8.1	8.4
	200	5	1	0.8	1.2	0.8	1.2	0.8	0.7	0.8	1.2	1.3	0.7	0.9
			5	5.7	5.2	5.8	5.5	4.1	5.5	4.2	5.9	5.7	4.4	4.2
			10	9.1	9.2	10.4	10.6	8.3	8.4	10.2	10.6	10.2	8.7	8.9
		9	1	1.2	1.1	0.9	0.8	0.7	1.4	0.9	0.8	1.2	0.7	0.7
			5	6.1	4.3	5.5	5.7	4.4	6.3	4.6	5.5	5.9	4.5	4.1
			10	10.9	9.5	10.5	11.0	7.6	11.2	9.3	10.6	10.7	8.6	9.2
		15	1	1.4	0.8	1.2	1.4	1.2	0.9	1.2	1.4	0.8	0.6	0.6
			5	6.0	4.5	6.2	5.4	4.1	5.8	4.7	5.8	5.6	4.3	4.5
			10	10.6	10.3	11.2	10.6	7.9	11.0	10.5	10.8	10.4	8.2	8.9
VARMA(1,1)	100	5	1	1.3	1.1	0.7	0.8	0.7	1.2	0.7	1.4	1.2	0.6	0.8
			5	5.7	4.7	6.2	4.5	4.3	5.8	4.4	5.8	4.6	3.9	4.3
			10	9.6	8.6	9.3	10.4	8.3	9.6	9.0	9.5	10.8	8.2	8.4
		8	1	1.4	0.7	0.8	1.2	0.7	1.3	0.8	1.2	0.9	0.8	1.3
			5	5.6	4.4	5.9	5.6	3.9	5.4	4.1	5.5	5.5	4.3	5.6
			10	10.6	8.5	11.3	10.6	7.3	9.4	9.0	11.0	10.7	8.0	9.4
		12	1	0.9	1.2	0.7	0.8	0.6	1.1	0.9	0.9	1.3	0.7	1.4
			5	5.4	5.1	6.0	5.6	4.2	5.6	5.4	5.8	5.6	4.1	4.5
			10	9.4	8.8	10.4	10.2	7.9	9.1	8.2	9.1	10.6	7.5	8.3
	200	5	1	0.8	1.3	0.7	0.9	0.7	1.2	0.8	1.2	1.2	0.7	1.3
			5	5.6	4.7	5.4	5.9	4.0	6.2	4.8	5.7	6.3	4.6	5.9
			10	9.0	9.3	10.6	11.0	8.9	10.5	9.2	10.5	9.6	8.2	8.9
		9	1	1.3	0.7	1.2	1.1	0.8	0.9	0.8	1.3	0.8	0.8	0.9
			5	6.1	5.2	4.2	6.1	4.3	5.7	5.1	5.5	6.3	4.3	5.6
			10	9.4	10.5	11.0	10.7	8.4	10.7	9.5	10.8	10.3	8.7	8.9
		15	1	1.4	1.1	0.8	0.9	0.7	1.3	0.9	0.9	0.8	0.7	0.8
			5	6.2	4.6	5.2	6.0	4.3	5.3	5.1	5.3	6.0	4.6	5.5
			10	10.3	10.5	10.8	10.6	7.9	10.7	10.2	11.2	10.7	8.4	9.1

levels are within two standard errors of the nominal level.

These results are similar to those obtained for orthogonality tests between stationary series; see Bouhaddioui and Roy (2006*a*). At the 1% and 10% nominal levels, both statistics have a small tendency to under or over-reject. There is no significant difference between the kernels. The best approximations are obtained with the Bartlett and Bartlett-Priestley kernels, while the performance of the Parzen kernel is inferior. With the Bartlett kernel, the empirical size is always within two standard errors of the nominal size. For the truncated uniform kernel, the size of Q_n and Q^*_n are very close to the size of $P_M{}^*$, which is normal since Q_n and Q^*_n are linear transformations of P_M

and P_M^* is a finite-sample version of P_M . For the models considered, the values of the truncation parameter M has no significant effect on the size of the tests. Finally, when the series length n goes from 100 to 200, the approximation improves very slightly.

6.2.2. Non-Gaussian innovations

We now examine simulation results where innovations follow a multivariate contaminated normal distribution. We consider the distribution

$$pN_m[\ ,] + (1 - p)N_m[\ , \mathbf{\Lambda}]$$

to denote the m-dimensional contaminated normal distribution in which the $N_m(0,\ )$ distribution is contaminated with probability 1 − p, by the $N_m[0, \mathbf{\Lambda}]$ distribution. We can verify that the fourth-order cumulants of this distribution depend on p, and $\mathbf{\Lambda}$. Thus, we consider in this part of the simulation two innovations series $\mathbf{a}^{(1)}{}_t$ and $\mathbf{a}^{(2)}{}_t$ generated independently according to the following two distributions:

$$p_1N_{m_1}[\ , \mathbf{I}_{m_1}] + (1 - p_1)N_{m_1}[\ , \mathbf{\Omega a}^{(1)}],\ p_2N_{m_2}[\ , \mathbf{I}_{m_2}] + (1 - p_2)N_{m_2}[\ , \mathbf{\Omega a}^{(2)}],$$

$$\mathbf{\Omega}_{\mathbf{a}^{(1)}} = \begin{matrix} 25 & 5 \\ 5 & 4 \end{matrix} \quad \text{and } \mathbf{\Omega}_{\mathbf{a}^{(2)}} = \begin{matrix} 25 & 7.5 \\ 7.5 & 4 \end{matrix}. \tag{6.2}$$

Simulations were made for different values of the pair (p_1, p_2) and for two models of Table 1, where $\Sigma_{\mathbf{a}}^{(1)}$ and $\Sigma_{\mathbf{a}}^{(2)}$ are now the covariance matrices of the two contaminated normal distributions in (6.2). The results in Table 4 are obtained by using (p_1, p_2) = (0.7, 0.9); the results for the other values of (p_1, p_2) are similar. From Table 4, we see that the non-normality of the innovations does not significantly affect the behavior of the test statistic Q_n with the associate kernel function and truncation parameter for the two sizes n = 100 and n = 200.

6.3. Power

Results on power are presented in Table 5. In VARδ (1), the cross-correlation at lag 0 between the two innovation series increases with δ and, as expected, the powers of the three tests increase with δ. Since the relative behaviors of the various tests are similar for the three values of δ considered [δ = 1, 1.5, 2], we only present the results for δ = 2. Similarly, we only present results for $Q^*{}_n$, since Q_n and $Q^*{}_n$ have exhibit similar behaviors with respect to kernels and truncation values.

Table 4. Empirical level (in percentage) of the test Q_n , Q^*_n and $P_M{}^*$

VAR(2) and **VARMA**(1, 1) models with non-Gaussian innovations. Number of realizations: 5000

				Q_n					Q_n^*					P_M^*
	n	M	α%	DAN	PAR	BAR	BP	TR	DAN	PAR	BAR	BP	TR	
VAR(2)	100	5	1	1.3	0.7	1.2	1.3	0.6	0.8	1.3	0.9	0.8	1.4	1.3
			5	5.4	4.6	5.8	5.3	4.1	5.5	4.4	5.9	5.8	4.0	4.2
			10	9.8	8.4	10.5	10.7	8.2	10.5	9.0	9.1	9.3	8.5	8.9
		8	1	0.7	1.2	0.8	1.3	0.7	1.2	0.8	0.8	1.3	0.7	0.8
			5	6.0	5.4	4.6	5.8	3.8	5.7	4.2	5.6	4.4	4.0	4.2
			10	11.0	9.4	10.6	9.5	8.2	10.8	9.4	10.8	10.6	8.4	8.8
		12	1	1.2	0.9	0.7	1.3	0.7	1.4	1.2	0.8	1.3	0.6	0.8
			5	5.8	5.6	5.2	5.6	4.0	4.6	4.8	5.3	5.4	3.8	4.2
			10	11.3	10.9	11.0	10.6	8.4	10.6	9.8	10.8	9.5	8.3	8.8
	200	5	1	1.2	0.9	0.8	1.3	0.7	0.8	1.3	1.1	0.8	0.8	1.2
			5	6.0	5.8	5.4	5.6	3.9	6.1	5.9	5.5	5.3	4.0	4.4
			10	10.6	9.0	10.2	10.4	8.4	9.4	10.8	11.0	10.6	8.4	9.2
		9	1	0.7	0.9	0.7	0.8	0.8	1.3	0.7	0.7	1.1	0.8	0.8
			5	5.8	5.6	5.2	4.7	4.2	6.0	4.8	5.8	5.8	4.2	4.6
			10	11.2	9.3	9.6	10.6	8.8	11.4	9.7	10.3	10.9	8.6	9.4
		15	1	1.3	1.1	0.8	0.7	0.7	1.1	1.3	0.9	0.8	0.6	0.7
			5	5.6	5.8	6.0	5.6	4.2	5.6	4.4	6.0	6.2	4.1	4.6
			10	11.2	10.6	10.2	10.8	8.6	11.0	10.8	10.3	10.2	8.6	9.0
VARMA(1,1)	100	5	1	0.8	1.2	1.3	0.7	0.6	1.1	0.8	1.2	1.2	0.6	0.7
			5	5.9	6.1	5.6	4.4	4.0	5.7	5.9	4.8	4.8	4.0	4.4
			10	10.6	9.2	9.6	11.0	8.5	10.9	10.4	9.2	11.0	8.0	9.0
		8	1	1.4	1.2	1.2	0.8	0.7	1.2	1.4	1.3	0.8	0.7	1.4
			5	6.0	4.2	5.6	5.8	3.8	6.2	4.0	6.1	6.3	4.2	6.0
			10	11.6	9.6	10.4	10.8	8.0	11.2	9.4	11.2	10.6	8.0	9.6
		12	1	0.8	1.3	0.8	0.9	0.7	1.2	1.1	0.9	1.1	0.8	1.3
			5	5.8	5.3	5.8	6.0	4.4	6.0	5.2	5.4	5.8	4.0	5.8
			10	10.8	9.2	11.4	10.6	8.1	11.2	9.4	9.3	11.0	8.4	8.8
	200	5	1	1.1	1.2	0.9	1.3	0.7	1.2	1.3	1.1	0.8	0.8	1.2
			5	6.1	5.4	4.8	6.1	4.2	5.9	4.7	5.4	6.0	4.4	5.8
			10	10.6	10.3	11.3	11.5	8.4	11.3	10.4	11.0	10.8	8.4	9.2
		9	1	1.3	1.2	0.9	1.2	0.8	1.2	1.3	1.1	0.9	0.7	1.3
			5	5.9	5.9	4.6	5.4	4.1	5.7	6.1	5.2	5.8	4.4	5.8
			10	11.4	10.8	10.6	10.6	8.8	11.2	10.8	10.4	9.8	8.6	9.3
		15	1	0.9	1.3	0.8	1.2	0.8	1.3	1.2	1.3	1.1	0.7	1.3
			5	5.4	5.8	6.2	5.6	4.0	5.5	5.6	5.8	5.4	4.2	5.8
			10	11.0	10.8	9.8	10.2	8.2	10.6	10.6	10.2	10.4	8.6	9.3

From Table 5, we draw the following observations. First, power decreases as M increases. Indeed, the model considered here is characterized by the lag 0 serial correlation. In such a sit-uation, we expect that the tests assigning more weight to small lags will be more powerful than those assigning weights to a large number of lags. For the three significance levels and the three truncation values, the Daniell kernel provides the most powerful test, while the Parzen, Bartlett and Bartlett-Priestley kernels yield similar powers for Q^*_n. However, the power of Q^*_n with the truncated uniform kernel is much smaller and is comparable to the power of $P_M{}^*$. For the chosen model, the new tests Q_n or Q^*_n with kernels other than the truncated uniform are preferable to the

Table 5. Power of the tests Q_n, Q^*_n and P_M^* based on asymptotic critical values with different kernels and different truncation values.
VAR_δ (1) model with δ = 2.

			Q_n^*					P_M^*
n	M	α%	DAN	PAR	BAR	BP	TR	
100	5	1	57.3	53.5	54.6	52.6	35.3	24.6
		5	63.2	60.1	56.4	58.6	36.8	26.8
		10	72.6	70.8	62.5	64.3	38.2	27.5
	8	1	49.6	46.1	51.4	48.0	27.5	22.6
		5	58.4	53.2	55.8	51.6	31.2	23.8
		10	63.7	60.8	62.6	61.7	34.6	25.8
	12	1	43.6	38.5	41.8	42.6	23.3	18.9
		5	50.2	44.7	40.3	43.0	26.4	21.2
		10	56.8	50.6	48.8	46.5	28.8	23.7
200	5	1	78.4	74.5	74.8	76.2	54.8	50.6
		5	85.6	82.6	81.6	85.8	56.4	54.1
		10	93.4	89.5	87.5	90.2	60.4	56.8
	9	1	69.5	65.2	63.0	66.8	42.4	40.7
		5	75.6	76.6	72.4	78.2	46.2	44.6
		10	80.8	78.5	77.6	82.8	50.4	46.4
	15	1	56.8	52.4	54.8	56.1	36.8	32.8
		5	60.1	57.4	53.2	60.1	40.2	35.0
		10	64.8	54.4	54.2	62.6	44.8	40.4

nonstationary multivariate version of Haugh's test P_M^* . Finally, the powers of all tests increase when the sample size varies from 100 to 200.

7. Conclusion

In this paper, we have proposed a semiparametric approach to test the non-correlation (or independence in the Gaussian case) between infinite-order cointegrated series IVAR(∞). The approach is semiparametric in the sense that if the two series are VARMA, we do not need to separately esti-mate the *true* model for each of the series. Instead, we fit a vector autoregression to each series, and the test statistics are based on residual cross-correlations at all possible lags. The weights assigned to the lags are defined by a kernel function and a smoothing parameter. Under the hypothesis of independence or non-causality of the two series, the asymptotic normality of the tests statistics are established. The finite-sample properties of the test were investigated by a Monte Carlo experiment which shows that the level is reasonably well controlled for both series lengths 100 and 200. Furthermore, with the model considered, the four kernels DAN, PAR, BAR, BP lead to similar powers and are more powerful than the truncated uniform kernel which corresponds to the multivariate version of the portmanteau test proposed by Bouhaddioui and Dufour (2008).

A. Appendix: proofs

The following notations are adopted. The Euclidean scalar product of X_t and X_s is defined by $\langle X_t, X_s\rangle = X_t' X_s$ and the Euclidean norm of X_t by $\|X_t\| = \sqrt{\langle X_t, X_t\rangle}$. The scalar Δ denotes a generic positive bounded constant which may differ from place to place.

PROOF OF PROPOSITION 3.1 First, let

$$\Xi := [\Xi_1 \vdots \cdots \vdots \Xi_p \vdots \Xi_{p+1,1}] = [\Psi \vdots \Pi_1 \vdots \cdots \vdots \Pi_p]\mathbf{D}_p := \Pi\, \mathbf{D}_p \tag{A.1}$$

where $\mathbf{D}_p$ is a suitable nonsingular transformation matrix containing the unknown matrix $\mathbf{C}_1$. The ECM representation (2.10) can be written as

$$\Delta X_t = \Psi_0 X_{2,t-1} + \sum_{l=1}^{p} \Xi_l \varepsilon_{t-j} + \Xi_{p+1,1}\varepsilon_{1,t-p-1} + e_t(n). \tag{A.2}$$

The matrices Ξ and Ψ_0 are defined in Saikkonen (1992, equation (A.2)). Set

$$\Lambda := [\Xi \vdots \Psi_0], \quad W_t := W_t(p) := [Y_t', X_{2,t-1}']', \tag{A.3}$$

$$Y_t := Y_t(p) := [\varepsilon'_{t-1}, \ldots, \varepsilon'_{t-p}, \varepsilon'_{1,t-p-1}]'. \tag{A.4}$$

Consider the linear transformation

$$\mathbf{b}_t := \Sigma^{-1/2}\mathbf{a}_t, \quad \hat{\mathbf{b}}_t = \Sigma^{-1/2}\hat{\mathbf{a}}_t, \tag{A.5}$$

where Σ is defined in (2.18). Since $C^{(12)}_{\hat{b}}(j) = \Sigma_1^{-1/2} C^{(12)}_{\hat{a}}(j)\Sigma_2^{-1/2}$. Using the property $\mathrm{vec}(ABC) = (C' \otimes A)\mathrm{vec}(B)$, we have:

$$T(\hat{\mathbf{a}}, \Sigma) = n\sum_{j=1-n}^{n-1} k^2(j/M)\, \mathbf{c}_{\hat{a}}^{(12)}(j)' \left(\Sigma_2^{-1} \otimes \Sigma_1^{-1}\right) \mathbf{c}_{\hat{a}}^{(12)}(j)$$

$$= n\sum_{j=1-n}^{n-1} k^2(j/M)\, \mathbf{c}_{\hat{b}}^{(12)}(j)' \mathbf{c}_{\hat{b}}^{(12)}(j) = T_{\hat{\mathbf{b}}}^{(12)}. \tag{A.6}$$

Thus, to prove the result, it is sufficient to show that

$$T_{\mathbf{b}}^{(12)} - T_{\hat{\mathbf{b}}}^{(12)} = o_p(M^{1/2}). \tag{A.7}$$

The result follows by decomposing the latter difference in two parts:

$$T_B^{(12)} - T_{\hat{B}}^{(12)} = n\sum_{j=1-n}^{n-1} k^2(j/M)\left\{\left\| \mathbf{c}_{\hat{B}}^{(12)}(j) - \mathbf{c}_B^{(12)}(j)\right\|^2 + 2\left\langle \mathbf{c}_B^{(12)}(j), \mathbf{c}_{\hat{B}}^{(12)}(j) - \mathbf{c}_B^{(12)}(j)\right\rangle\right\}$$

$$= [T_n^{(1)} + T_{n-}^{(1)}] + T_n^{(2)} \tag{A.8}$$

$$T_n^{(1)} := n \sum_{j=0}^{n-1} k^2(j/M)\, \| \hat{\mathbf{c}}_{\mathbf{b}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j) \|^2 , \tag{A.9}$$

$$T_{n-}^{(1)} := n \sum_{j=1-n}^{-1} k^2(j/M) (\| \hat{\mathbf{c}}_{\mathbf{b}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j) \|^2 , \tag{A.10}$$

$$T_n^{(2)} := 2 \sum_{j=1-n}^{n-1} \langle \mathbf{c}_{\mathbf{b}}^{(12)}(j), \hat{\mathbf{c}}_{\mathbf{b}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j) \rangle , \tag{A.11}$$

and then showing that each part is $o_p(M^{1/2})$. Consider the positive lags $j \geq 0$, since for negative lags, the proof is similar by symmetry.

Define $\hat{\delta}_t = \mathbf{b}_t^{(1)} - \hat{\mathbf{b}}_t^{(1)}$ and $\hat{\eta}_t = \mathbf{b}_t^{(2)} - \hat{\mathbf{b}}_t^{(2)}$. From (3.2), we have

$$T_n^{(1)} = n \sum_{j=0}^{n-1} k^2(j/M)\, \| \hat{\mathbf{c}}_{\mathbf{b}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j) \|^2$$

$$= n \sum_{j=0}^{n-1} k^2(j/M) \left\| \frac{1}{n} \sum_{t=j+1}^{n} (\mathbf{b}_t^{(1)} \mathbf{b}_{t-j}^{(2)\prime} - \hat{\mathbf{b}}_t^{(1)} \hat{\mathbf{b}}_{t-j}^{(2)\prime}) \right\|^2 , \tag{A.12}$$

and using the Cauchy-Schwarz inequality, we obtain

$$T_n^{(1)} = n \sum_{j=0}^{n-1} k^2(j/M) \left\| \frac{1}{n} \sum_{t=j+1}^{n} (\mathbf{b}_t^{(1)} \hat{\eta}_{t-j}' + \hat{\delta}_t \mathbf{b}_{t-j}^{(2)\prime} - \hat{\delta}_t \hat{\eta}_{t-j}') \right\|^2 \leq 4n(T_{1n} + T_{2n} + T_{3n}) \tag{A.13}$$

with $T_{1n} = \sum_{j=0}^{n-1} k^2(j/M) \left\| \frac{1}{n} \sum_{t=j+1}^{n} \mathbf{b}_t^{(1)} \hat{\eta}_{t-j}' \right\|^2$, $T_{2n} = \sum_{j=0}^{n-1} k^2(j/M) \left\| \frac{1}{n} \sum_{t=j+1}^{n} \hat{\delta}_t \mathbf{b}_{t-j}^{(2)\prime} \right\|^2$ and $T_{3n} = \sum_{j=0}^{n-1} k^2(j/M) \left\| \frac{1}{n} \sum_{t=j+1}^{n} \hat{\delta}_t \hat{\eta}_{t-j}' \right\|^2$. It suffices to show that the terms T_{jn}, $j = 1, 2, 3$, are $o_p(M^{1/2}/n)$. We can then write:

$$\hat{\delta}_t = (\hat{\mathbf{b}}_t^{(1)} - \boldsymbol{\Sigma}_1^{-1/2} \mathbf{e}_t^{(1)}) + (\boldsymbol{\Sigma}_1^{-1/2} \mathbf{e}_t^{(1)} - \mathbf{b}_t^{(1)}) = \boldsymbol{\Sigma}_1^{-1/2} \{[\hat{\mathbf{a}}_t^{(1)} - \mathbf{e}_t^{(1)}] + [\mathbf{e}_t^{(1)} - \mathbf{a}_t^{(1)}]\}$$

$$= \boldsymbol{\Sigma}_1^{-1/2} \{(\boldsymbol{\Lambda}^{(1)} - \hat{\boldsymbol{\Lambda}}^{(1)}) \mathbf{W}_t^{(1)} + \xi_t(p_1)\} \tag{A.14}$$

where $\mathbf{e}_t^{(h)} := \mathbf{e}_t^{(h)}(n)$, $\boldsymbol{\Lambda}^{(h)}$ and $\mathbf{W}_t^{(h)}$, $h = 1, 2$ are defined as in (A.2) for each process, $\hat{\boldsymbol{\Lambda}}$ is the LS estimator of $\boldsymbol{\Lambda}$ and $\xi_t(p_1) = \sum_{l=p_1+1}^{\infty} \boldsymbol{\Phi}_l \mathbf{X}_{t-l}^{(1)}$ represents the bias of the VAR(p_1) approximation of $\{\mathbf{X}_t^{(1)}\}$. The second equality is from Saikkonen and Lütkepohl (1996, page 832). Also, using the result of Proposition 2.1, we deduce that

$$\| \hat{\boldsymbol{\Lambda}}^{(1)} - \boldsymbol{\Lambda}^{(1)} \|^2 = O_p\left(\frac{p_1}{n}\right) . \tag{A.15}$$

By equation (3.15) in Bouhaddioui and Roy (2006*b*), we have $\mathbf{E}(\| \xi_t(p_n^{(h)}) \|^2) = O\left(\sum_{l=p_n^{(h)}+1}^{\infty} \| \boldsymbol{\Phi}_l^{(h)} \| \right)^2$, $h = 1, 2$. Based on the result (3.17) in Bouhaddioui and Roy (2006*b*) and equation (2.15), we obtain:

$$T_{1n} = \sum_{j=0}^{n-1} k^2(j/M) \left\| \frac{1}{n}\sum_{t=j+1}^{n} \mathbf{b}_t^{(1)} \eta_{t-j}' \right\|^2 = O_p\left(\frac{p_n^{(2)}M}{n^2}\right)\left\{\frac{1}{M}\sum_{j=0}^{n-1} k^2(j/M)\right\}. \quad (A.16)$$

Since $p^{(2)}_n = o(n/M^{1/2})$, we have $T_{1n} = o_p(M^{1/2}/n)$. By symmetry, we can prove that $T_{2n} = o_p\left(\frac{M^{1/2}}{n}\right)$. For the third term T_{3n}, using the Cauchy-Schwarz inequality, we obtain

$$
\begin{aligned}
T_{3n} &= \sum_{j=0}^{n-1} k^2(j/M) \left\| n^{-1}\sum_{t=j+1}^{n} \hat{\eta}_t \hat{\text{o}}_{t-j}' \right\|^2 \\
&\leq \left\| \Lambda^{(1)} - \hat{\Lambda}^{(1)} \right\|^2 \left\| \Lambda^{(2)} - \hat{\Lambda}^{(2)} \right\|^2 \sum_{j=0}^{n-1} k^2(j/M) \left\| n^{-1}\sum_{t=j+1}^{n} \mathbf{w}_t^{(1)}(p_1)\mathbf{w}_{t-j}^{(2)}(p_2)' \right\|^2 \\
&+ \left\| \Lambda^{(1)} - \hat{\Lambda}^{(1)} \right\|^2 \sum_{j=0}^{n-1} k^2(j/M) \left\| n^{-1}\sum_{t=j+1}^{n} \mathbf{w}_t^{(1)}(p_1)\xi_{t-j}(p_2)' \right\|^2 \\
&+ \left\| \Lambda^{(2)} - \hat{\Lambda}^{(2)} \right\|^2 \sum_{j=0}^{n-1} k^2(j/M) \left\| n^{-1}\sum_{t=j+1}^{n} \xi_t(p_1)\mathbf{w}_{t-j}^{(2)}(p_2)' \right\|^2 \\
&+ \sum_{j=0}^{n-1} k^2(j/M) \left\| n^{-1}\sum_{t=j+1}^{n} \xi_t(p_1)\xi_{t-j}(p_2)' \right\|^2. \quad (A.17)
\end{aligned}
$$

Using the equations (3.19) - (3.22) in Bouhaddioui and Roy (2006a), the assumptions $p_n^{(h)} = o(n^{1/2}/M^{1/4})$, $n\sum_{l=p_n^{(h)}+1}^{\infty} \|\Phi_l^{(h)}\|^2 = o(n^{1/2}/M^{1/4})$ and the result (2.15), we conclude that $T_{3n} = o_p(M^{1/2}/n)$. Therefore, we obtain

$$T_n^{(1)} = n\sum_{j=0}^{n-1} k^2(j/M) \left\| \hat{\mathbf{c}}_{\mathbf{b}}^{(12)}(j) - \mathbf{c}^{(12)}_{\mathbf{b}}(j) \right\|^2 = o_p\left(M^{1/2}\right). \quad (A.18)$$

Finally, using Cauchy-Schwarz inequality once more, we have

$$|T_n^{(2)}| \leq n\sum_{j=1-n}^{n-1} k^2(j/M) | \langle \mathbf{c}^{(12)}_{\mathbf{b}}(j), \hat{\mathbf{c}}^{(12)}(j) - \mathbf{c}^{(12)}_{\mathbf{b}}(j) \rangle | \leq n\sum_{l=4}^{6} T_{ln}, \quad (A.19)$$

with

$$T_{4n} = \sum_{j=0}^{n-1} k^2(j/M) \left\| \mathbf{c}_{\mathbf{b}}^{(12)}(j) \right\| \left\| \frac{1}{n}\sum_{t=j+1}^{n} \hat{\delta}_t(\mathbf{b}_{t-j}^{(2)})' \right\|, \quad (A.20)$$

$$T_{5n} = \sum_{j=0}^{n-1} k^2(j/M) \left\| \mathbf{c}_{\mathbf{b}}^{(12)}(j) \right\| \left\| \frac{1}{n}\sum_{t=j+1}^{n} \mathbf{b}_t^{(1)} \hat{\eta}_{t-j}' \right\|, \quad (A.21)$$

$$T_{6n} = \sum_{j=0}^{n-1} k^2(j/M) \left\| \mathbf{c}_{\mathbf{b}}^{(12)}(j) \right\| \left\| \frac{1}{n}\sum_{t=j+1}^{n} \hat{\text{o}}_t \hat{\eta}_{t-j}' \right\|. \quad (A.22)$$

Thus, it is sufficient to show that the terms T_{jn}, j = 4, 5, 6, are $o_p(M^{1/2}/n)$. By conditioning on $(\mathbf{b}^{(2)}_s)^n_{s=-\infty}$ and using Jensen's inequality, we have

$$\mathbf{E}\left[T_{4n} \mid (\mathbf{b}_s^{(2)})_s^n{}_{=-\infty} \right] \leq \sum_{j=1-n}^{n-1} k^2(j/M)$$

$$\times \left[\mathbf{E}\left(\left\{\left(\frac{1}{n}\sum_{\tau=1}^{n} \mathbf{b}_\tau^{(1)} \mathbf{b}_{\tau-j}^{(2)T}\right)\left(\frac{1}{n}\sum_{t=j+1}^{n} \hat{\delta}_t \mathbf{b}_{t-j}^{(2)T}\right)\right\}^2 \Big| (\mathbf{b}_s^{(2)})_{s=-\infty}^{n}\right)\right]^{1/2}$$

$$\le \frac{M\Delta}{n^2}\left\{\frac{1}{M}\sum_{1-n}^{n-1} k^2(j/M)\right\}\left(\frac{1}{n}\sum_{\tau=1}^{n} \|\mathbf{b}_\tau^{(2)}\|^2\right)^{1/2}\left(\frac{1}{n}\sum_{t=1}^{n} \mathbf{E}\,\hat{\delta}_t^2\right)^{1/2}$$

$$= O_p\left(\frac{M\,(p_n^{(2)})^{1/2}}{n^{5/2}}\right) = o_p\left(\frac{M^{1/2}}{n^{3/2}}\right). \qquad (A.23)$$

The first equality is obtained by using the conditions on $p^{(2)}{}_n$, $\mathbf{\Phi}^{(2)}$, and the assumption of inde-pendence of the two innovation series. Then, $T_{4n} = o_p(M^{1/2}/n)$. By symmetry, we have also $T_{5n} = o_p(M^{1/2}/n)$. Finally, from Markov inequality, we have

$$\sum_{j=1}^{n-1} k^2(j/M)\,\|\mathbf{c_b}^{(12)}(j)\|^2 = O_p(M/n) \qquad (A.24)$$

hence, using the Cauchy-Schwarz inequality and the result for T_{3n}, we obtain that $T_{6n} = o_p(M/n)$. Thus, $T_n^{(2)} = o_p(M^{1/2})$ and the proof of Proposition 3.1 is completed. □

PROOF OF PROPOSITION 3.2 Since $D_n(k) = M\,D(k)\{1 + o(1)\}$, it is sufficient to show that

$$T(\hat{\mathbf{a}}, \hat{\mathbf{\Sigma}}) - T(\hat{\mathbf{a}}, \mathbf{\Sigma}) = O_p(M/n^{1/2}). \qquad (A.25)$$

Using the fact that $\mathbf{C}_{\hat{\mathbf{a}}}^{(hh)}(0) - \mathbf{\Sigma}_n^h = O_p(n^{-1/2})$ for h = 1, 2 [see Lütkepohl and Saikkonen (1997, p.133)], it follows that

$$\mathbf{C}_{\hat{\mathbf{a}}}^{(22)}(0)^{-1} \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{(11)}(0)^{-1} - \mathbf{\Sigma}_2^{-1} \otimes \mathbf{\Sigma}_1^{-1} = \mathbf{O}_p(n^{-1/2}). \qquad (A.26)$$

Thus,

$$T(\hat{\mathbf{a}}, \hat{\mathbf{\Sigma}}) - T(\hat{\mathbf{a}}, \mathbf{\Sigma}) = n\sum_{j=1-n}^{n-1} k^2(j/M)\,\mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)'\,\mathbf{O}_p(n^{-1/2})\,\mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)$$

$$= O_p(n^{1/2})\sum_{j=1-n}^{n-1} k^2(j/M)\,\mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)'\,\mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j). \qquad (A.27)$$

To complete the proof, it remains to prove that

$$B(n) = \sum_{j=1-n}^{n-1} k^2(j/M)\,\mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)'\,\mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j) = O_p(M/n). \qquad (A.28)$$

First, let us decompose B(n) in two parts

$$B(n) = \sum_{j=1-n}^{n-1} k^2(j/M)\{\mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)'\,\mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j) - \mathbf{c_a}^{(12)}(j)'\,\mathbf{c_a}^{(12)}(j)\} + \sum_{j=1-n}^{n-1} k^2(j/M)\,\mathbf{c_a}^{(12)}(j)'\,\mathbf{c_a}^{(12)}(j)$$

$$= B_1 + B_2. \qquad (A.29)$$

By an argument similar to the one used to prove (A.7) in Proposition 3.1, we have:

$$B_1(n) = \sum_{j=1-n}^{n-1} k^2(j/M)\{\mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)'\,\mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j) - \mathbf{c_a}^{(12)}(j)'\,\mathbf{c_a}^{(12)}(j)\} = o_p(M^{1/2}/n), \qquad (A.30)$$

and, by Markov inequality, it follows that

$$B_2(n) = \sum_{j=1-n}^{n-1} k^2(j/M)\,\mathbf{c}^{(12)}_{\mathbf{a}}(j)'\,\mathbf{c}^{(12)}_{\mathbf{a}}(j) = O_P(M/n)\,. \tag{A.31}$$

Combining the results for $B_1(n)$ and $B_2(n)$, we obtain that

$$T(\hat{\mathbf{a}}, \hat{\boldsymbol{\Sigma}}) - T(\tilde{\mathbf{a}}, \boldsymbol{\Sigma}) = O_p(n^{1/2})O_p(M/n) = O_p(M/n^{1/2}), \tag{A.32}$$

and the proof of Proposition 3.2 is completed. □

PROOF OF THEOREM 4.1 First, we note that the statistic Q_n is a normalized version of $T(\hat{\mathbf{a}}, \hat{\boldsymbol{\Sigma}})$ which can be viewed as the L_2-norm of a kernel-based estimator of the cross-coherency function between the two innovations processes. Thus, the statistic Q_n can be expressed as

$$Q_n = \frac{n\,\|\mathbf{s}^{(12)}_{\hat{\mathbf{a}}}\|_2^2 - m_1m_2S_n(k)}{\{2m_1m_2D_n(k)\}^{1/2}} \tag{A.33}$$

where $\mathbf{s}^{(12)}_{\hat{\mathbf{a}}}$ is the estimator of the cross-coherency function between the two innovations processes given by

$$\|\mathbf{s}^{(12)}_{\mathbf{a}}\|_2^2 = \sum_{j=-\infty}^{\infty} \gamma^{(12)}_{\mathbf{a}}(j)'\,(\boldsymbol{\Sigma}_2 \otimes \boldsymbol{\Sigma}_1)^{-1}\,\gamma^{(12)}_{\mathbf{a}}(j) \tag{A.34}$$

where $\gamma^{(12)}_{\mathbf{a}}(j) := \mathrm{vec}[\,\Gamma^{(12)}_{\mathbf{a}}(j)]$. For details, see Section 4 in Bouhaddioui and Roy (2006*a*). By definition of Q_n, we can write

$$\frac{M^{1/2}}{n}Q_n = \frac{M^{1/2}\|\mathbf{s}^{(12)}_{\hat{\mathbf{a}}}\|_2^2 - \frac{M^{1/2}}{n}m_1m_2S_n(k)}{\{2m_1m_2D(k)\}^{1/2}}$$

$$= \frac{\|\mathbf{s}^{(12)}_{\hat{\mathbf{a}}}\|_2^2}{\{2m_1m_2M^{-1}D_n(k)\}^{1/2}} - \frac{n^{-1}S_n(k)}{\{2M^{-1}D_n(k)\}^{1/2}}(m_1m_2)^{1/2}\,. \tag{A.35}$$

From (3.11), the last term of the previous equation goes to zero when $M/n \to 0$ as $n \to \infty$.

Using the linear transformation $\mathbf{b}_t = \boldsymbol{\Sigma}^{-1/2}\mathbf{a}_t$, as in Proposition 3.1, we have $\mathbf{s}^{(12)}_{\mathbf{a}} = \mathbf{s}^{(12)}_{\mathbf{b}}$. Also, since the processes $\mathbf{b}^{(1)}$ and $\mathbf{b}^{(2)}$ are stationary and by Lemma A.7 in Bouhaddioui and Roy (2006*a*), we have that

$$\|\mathbf{s}^{(12)}_{\tilde{\mathbf{b}}}\|_2^2 - \|\mathbf{s}^{(12)}_{\mathbf{b}}\|_2^2 \xrightarrow{p} 0 \tag{A.36}$$

where $\mathbf{s}^{(12)}_{\tilde{\mathbf{b}}}$ is defined as $\mathbf{s}^{(12)}_{\hat{\mathbf{b}}}$, the residual series $(\hat{\mathbf{b}}^{(1)}_t, \hat{\mathbf{b}}^{(2)}_t)_{t=1}^n$ being replaced by the innovation series $(\mathbf{b}^{(1)}_t, \mathbf{b}^{(2)}_t)_{t=1}^n$. Thus, to prove the consistency result (4.4), it is sufficient to verify that $\|\mathbf{s}^{(12)}_{\hat{\mathbf{b}}}\|_2^2 - \|\mathbf{s}^{(12)}_{\tilde{\mathbf{b}}}\|_2^2 \xrightarrow{p} 0$, which follows from the following lemma. □

Lemma A.1 *Under the assumptions of Theorem 4.1, we have*

$$\|\hat{s}^{(12)}_b\|_2^2 - \|\tilde{s}^{(12)}_b\|_2^2 \xrightarrow{p} 0 \tag{A.37}$$

PROOF OF LEMMA A.1 By definition of $\hat{s}^{(12)}_b$ and $\tilde{s}^{(12)}_b$, and by similar calculations to those for the proof in Proposition 3.1, we obtain

$$\begin{aligned}\|\hat{s}^{(12)}_b\|_2^2 - \|\tilde{s}^{(12)}_b\|_2^2 &= \sum_{j=1-n}^{n-1} k^2(j/M)\left(\|\hat{c}^{(12)}_b(j)\|^2 - \|c^{(12)}_b(j)\|^2\right)\\ &= \sum_{j=1-n}^{n-1} k^2(j/M)\,\|\hat{c}^{(12)}_b(j) - c^{(12)}_b(j)\|^2\\ &\quad + 2\sum_{j=1-n}^{n-1} k^2(j/M)\left\langle c^{(12)}_b(j), \hat{c}^{(12)}_b(j) - c^{(12)}_b(j)\right\rangle .\end{aligned} \tag{A.38}$$

It is sufficient to prove that the first term goes to zero in probability, because the second term can be bounded by a product of the first term and a finite quantity, using the Cauchy-Schwarz inequality. With the notations of Proposition 3.1, we can write

$$\sum_{j=1-n}^{n-1} k^2(j/M)\,\|\hat{c}_b^{(12)}(j) - c_b^{(12)}(j)\|^2 \le 4\sum_{l=1}^{3} T_{ln}, \tag{A.39}$$

where T_{ln}, $l = 1, 2, 3$, are defined in Proposition 3.1. We first prove that $T_{1n} \to 0$ in probability. By the Cauchy-Schwarz inequality, we obtain

$$T_{1n} \le M\left\{\frac{1}{M}\sum_{j=0}^{n-1} k^2(j/M)\right\}\left\{\frac{1}{n}\sum_{t=1}^{n}\|b_t^{(1)}\|^2\right\}\left\{\frac{1}{n}\sum_{t=1}^{n}\|\hat{\eta}_t\|^2\right\}. \tag{A.40}$$

By definition of $\hat{\eta}_t$, it follows that

$$\frac{1}{n}\sum_{t=j}^{n}\|\hat{\eta}_t\|^2 \le \frac{1}{n}\sum_{t=1}^{n}\left\{\|(\Lambda^{(2)} - \hat{\Lambda}^{(2)})W_t^{(2)}\|^2 + \|\xi_t(p_n^{(2)})\|^2\right\}. \tag{A.41}$$

Since $\sum\|a^{(11)}(l)\|$ is uniformly bounded by a positive constant Δ, and the parameters $\{\Phi_l\}$ are a linear function of the original parameters $\{G_{l,\infty}\}$, then the bias approximation can be bounded by

$$E\|\xi_t(p_n^{(2)})\|^2 \le \Delta\left(\sum_{l=p_2+1}^{\infty}\|\Phi_l^{(2)}\|\right)^2 = o(n^{-1}). \tag{A.42}$$

See also the result (A.12) in Saikkonen (1992). Under the assumptions on the process $\mathbf{b}$, on $p_n^{(2)}$ and on the parameters $(\Phi_l^{(2)})$, we have

$$T_{1n} = O_p\left(\frac{M(p_n^{(2)})^2}{n}\right) + O_p\left(M\sum_{l=p_2+1}^{\infty}\|\Phi_l^{(2)}\|^2\right) = o_p(1). \tag{A.43}$$

By symmetry, we can verify that $T_{2n} = o_p(1)$. For T_{3n}, we can write

$$T_{3n} = \sum_{j=0}^{n-1} k^2(j/M)\left\|\frac{1}{n}\sum_{t=j+1}^{n}\hat{o}_t\eta'_{t-j}\right\|^2$$

$$\leq M\{\frac{1}{M}\sum_{j=0}^{n-1} k^2(j/M)\}\{\frac{1}{n}\sum_{t=1}^{n}\hat{\delta}_t^{(1)2}\}^2\{\frac{1}{n}\sum_{t=1}^{n}\hat{\eta}_t^2\}. \quad (A.44)$$

By symmetry, we can prove that $\frac{1}{n}\sum_{t=1}^{n}\hat{\delta}_t = O_p(p_n)/n + O_p(1)\sum_{l=p_1+1}^{\infty}\|\Phi^{(1)}\|^2$, and using the same assumptions as those for T_{1n}, we obtain that $T_{3n} = o_p(1)$. Finally, we conclude that

$$\|\hat{\tilde{s}}_b^{(12)}\|^2 - \|\tilde{s}_b^{(12)}\|^2 = o_p(1). \quad (A.45)$$

This completes the proof of Lemma A.1 and then Theorem 4.1.

PROOF OF THEOREM 5.1 By the proof of Theorem 4.1,

$$Q_n = \frac{n\|\mathbf{s_b}^{(12)}\|_2^2 - m_1m_2S_n(k)}{\sqrt{2m_1m_2D_n(k)}} + o_p(1). \quad (A.46)$$

where

$$\begin{aligned}
n\|\mathbf{s_b}^{(12)}\|_2^2 &= \sum_{j=1-n}^{n-1} k^2(j/M)\mathbf{c_b}^{(12)}(j)\mathbf{c_b}'^{(12)}(j) \\
&= \sum_{j=1-n}^{n-1} nk^2(j/M)(\mathbf{c}^{(12)}\mathbf{b}(j) - \gamma^{(12)}\mathbf{b}(j))'(\mathbf{c}^{(12)}\mathbf{b}(j) - \gamma^{(12)}\mathbf{b}(j)) \\
&\quad +2n\sum_{j=1-n}^{n-1} k^2(j/M)\mathbf{cb}^{(12)}(j)'\gamma\mathbf{b}^{(12)}(j) - n\sum_{j=1-n}^{n-1} k^2(j/M)\gamma\mathbf{b}^{(12)}(j)'\gamma\mathbf{b}^{(12)}(j) \\
&= \sum_{j=1-n}^{n-1} nk^2(j/M)(\mathbf{c}^{(12)}\mathbf{b}(j) - \gamma^{(12)}\mathbf{b}(j))'(\mathbf{c}^{(12)}\mathbf{b}(j) - \gamma^{(12)}\mathbf{b}(j)) \\
&\quad +M^{1/2}\sum_{j=1-n}^{n-1} k^2(j/M)\lambda\mathbf{b}^{(12)}(j)'\lambda\mathbf{b}^{(12)}(j) + o_p(M^{1/4}).
\end{aligned}$$

Since there exists $j^* \in \mathbf{Z}$ such that $\mathbf{\Lambda}^{(12)}\mathbf{b}(j) = , : \forall j : (|j| > j^*)$, we have

$$\begin{aligned}
\sum_{j=-\infty}^{\infty} k^2(j/M)\lambda^{(12)}\mathbf{b}(j)'\lambda^{(12)}\mathbf{b}(j) &= \sum_{j=-j^*}^{j^*} k^2(j/M)\lambda^{(12)}\mathbf{b}(j)'\lambda^{(12)}\mathbf{b}(j) \\
&\to \sum_{j=-\infty}^{\infty} \lambda^{(12)}\mathbf{b}(j)'\lambda^{(12)}\mathbf{b}(j) := \beta(\mathbf{\Lambda}^{(12)}\mathbf{b}).
\end{aligned}$$

Thus,

$$\frac{n\|\mathbf{s_b}^{(12)}\|_2^2 - m_1m_2S_n(k)}{\sqrt{2m_1m_2D_n(k)}} \xrightarrow{L} Z + \frac{\beta(\mathbf{\Lambda_b}^{(12)})}{\sqrt{2m_1m_2D(k)}}$$

where $Z \sim N(0, 1)$. □

References

Ahn, S. k. and Reinsel, G. C. (1990), 'Estimation of partially nonstationary multivariate autore-gressive models', *Journal of the American Statistical Association* **85**, 813–823.

Akaike, H. (1969), 'Power spectrum estimation through autoregressive model fitting', *Annals of the Institute of Statistical Mathematics* **21**, 407–419.

Bahadur, R. R. (1960), 'Stochastic comparison of tests', *Annals of Mathematical Statistics* **31**, 276– 295.

Berk, K. N. (1974), 'Consistent autoregressive spectral estimates', *Annals of Statistics* **2**, 489–502.

Bhansali, R. J. (1996), 'Asymptotically efficient autoregressive model selection for multistep pre-diction', *Annals of the Institute of Statistical Mathematics* **48**, 577–602.

Bouhaddioui, C. (2002), Tests d'indépendance de deux séries multivariées autorégressives d'ordre infini, PhD thesis, Département de mathématiques et de statistique, Université de Montréal.

Bouhaddioui, C. and Dufour, J.-M. (2008), 'Tests for non-correlation of two infinite-order cointe-grated vector autoregressive series', *Journal of Applied Probability and Statistics* **3**(1), 78–94.

Bouhaddioui, C. and Roy, R. (2006*a*), 'A generalized portmanteau test for independence of two infinite order vector autoregressive series', *Journal of Time Series Analysis* **27**(4), 505–544.

Bouhaddioui, C. and Roy, R. (2006*b*), 'On the distribution of residual cross-correlations between two uncorrelated infinite order vector autoregressive series', *Statistics and Probability Letters* **76**(1), 58–68.

Burnham, K. P. and Anderson, D. R. (2002), *Model selection and multimodel inference : a practical information-theoretic approach*, Springer-Verlag, New York.

Duchesne, P. (2005), 'Testing for serial correlation of unknown form in cointegrated time series models?', *Annals of the Institute of Statistical Mathematics* **57**, 575–595.

Duchesne, P. and Roy, R. (2003), 'Robust tests for independence of two time series', *Statistica Sinica* **13**, 827–852.

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

Engle, R. and Granger, C. (1987), 'Co-integration and error correction: Representation, estimation and testing', *Econometrica* **55**, 251–276.

Ghosh, B. K. and Huang, W. M. (1991), 'The power and optimal kernel of the Bickel-Rosenblatt test for goodness of fit', *The Annals of Statistics* **17**, 999–1009.

Hallin, M., Jurecková,ˇ 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. (2005), 'Testing independance and causality between multivariate ARMA times series', *Journal of Time Series Analysis* **26**, 83–106.

Hallin, M. and Saidi, A. (2007), 'Optimal tests for non-correlation between multivariate time se-ries', *Journal of The American Statistical Association* **102**(479), 938–951.

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

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

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

Hong, Y. (1998), 'Testing for pairwise serial independence via the empirical distribution function', *Journal of the Royal Statistical Society, Series B* **60**(2), 271–496.

Hong, Y. (1999), 'Hypothesis testing in time series via the empirical characteristic function: A generalized spectral density approach', *Journal of the American Statistical Association* **94**(448), 1201–1220.

Lewis, R. and Reinsel, G. C. (1985), 'Prediction of multivariate time series by autoregressive model fitting', *Journal of Multivariate Analysis* **16**, 393–411.

Lütkepohl, H. (1985), 'The joint asymptotic distribution of multistep prediction errors of estimated vector autoregressions', *Economics Letters* **17**, 103–106.

Lütkepohl, H. (2001), Vector autoregressions, *in* B. Baltagi, ed., 'Companion to Theoretical Econo-metrics', Blackwell Companions to Contemporary Economics, Basil Blackwell, Oxford, U.K., chapter 32, pp. 678–699.

Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.

Lütkepohl, H. and Saikkonen, P. (1997), 'Impulse response analysis in infinite order cointegrated vector autoregressive processes', *Journal of Econometrics* **81**, 127–157.

Lütkepohl, H. and Saikkonen, P. (1999), Order selection in testing for the cointegrating rank of a VAR process, *in* 'Cointegration, Causality, and Forecasting. A Festschrift in Honour of Clive W.J. Granger, Engle, R.F., White, H.', Oxford University Press, Oxford, pp. 168–199.

Park, J. Y. (1990), ‘Testing for unit roots and cointegration by variable addition’, *Advances in Econometrics* **8**, 107–133.

Park, J. Y., Shin, K. and Wang, Y.-J. (2010), ‘A semiparametric cointegrating regression: Investi-gating the effects of age distributions on consumption and saving’, *Journal of Econometrics* **157**, 165–178.

Parzen, E. (1974), ‘Some recent advances in time series modeling’, *IEEE Transactions on Auto-matic Control* pp. 723–730.

Pham, D. T., Roy, R. and Cédras, L. (2003), ‘Tests for non-correlation of two cointegrated ARMA time series’, *Journal of Time Series Analysis* **24**, 553–577.

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

Reinsel, G. C. (1993), *Elements of Multivariate Time Series Analysis*, Springer-Verlag, New York.

Reinsel, G. C. (1997), *Elements of Multivariate Time Series Analysis*, second edn, Springer-Verlag, New York.

Saidi, A. (2007), ‘Consistent testing for non-correlation of two cointegrated ARMA time series’, *Canadian Journal of Statistics* **35**(1), 169–188.

Saikkonen, P. (1992), ‘Estimation and testing of cointegrated systems by an autoregressive approx-imation’, *Econometric Theory* **8**, 1–27.

Saikkonen, P. and Lütkepohl, H. (1994), Infinite order cointegrated autoregressive processes: Esti-mation and inference, Technical Report 5, Institut für Statistik und Ökonometrie, Humboldt-Universität zu Berlin.

Saikkonen, P. and Lütkepohl, H. (1996), ‘Infinite-order cointegrated vector autoregressive pro-cesses: Estimation and inference’, *Econometric Theory* **12**, 814–844.

Saikkonen, P. and Luukkonen, R. (1997), ‘Testing cointegration in infinite order vector autoregres-sive proceses’, *Journal of Econometrics* **81**, 93–126.

Shao, X. (2009), ‘A generalized portmanteau test for independence between two stationary time series’, *Econometric Theory* **25**, 195–210.

Yap, S. F. and Reinsel, G. C. (1995), ‘Estimation and testing for unit roots in a partially non-stationary vector autoregressive moving average model’, *Journal of the American Statistical Analysis* **90**, 253–267.