ANNALS OF ECONOMICS AND FINANCE **2**, 123–164 (2001)

Testing for Independence between Two stationary Time Series via the Empirical Characteristic Function

Yongmiao Hong

Department of Economics and
Department of Statistical Science
Cornell University, Ithaca, NY 14853, USA
E-mail: yh20@cornell.edu

This paper proposes an asymptotic one-sided $N(0,1)$ test for independence between two stationary time series using the empirical characteristic function. Unlike the tests based on the cross-correlation function (e.g. Haugh, 1976; Hong, 1996; Koch & Yang 1986), the proposed test has power against all pairwise cross-dependencies, including those with zero cross-correlation. By differentiating the empirical characteristic function at the origin, the present approach yields a modified version of Hong's (1996) test, which in turn generalizes Haugh's (1976) test. Other new tests can be derived by further differentiating the empirical characteristic function properly. A simulation study compares the new test with those of Haugh (1976), Hong (1996) and Koch & Yang (1986) in finite samples; the results show that the new test has reasonable sizes and good powers against linear and nonlinear cross-dependencies.

Key Words: Asymptotic normality; Cramer-von Mises statistic; Empirical characteristic function; Independence; Kernel function; Multivariate time series.
JEL Classification Numbers:

1. INTRODUCTION

Testing independence between time series is important in multivariate time series analysis. Haugh (1976) proposes a popular test by first prewhitening two time series and then testing whether the two residual series are independent via the residual cross-correlation function (see also Pierce, 1977; Pierce & Haugh, 1977). Haugh's statistic is based on the sum of squared residual cross-correlations of finitely many lags and has a null chi-square distribution. Koch & Yang (1986) extended Haugh's approach by accounting for a possible pattern among successive cross-correlation coefficients. The test places more weights on successful coefficients that fall

1529-7373/2001

on the same side of zero. This yields good power against alternatives in which two time series are related over a long distributed lag with relatively small lag coefficients of same sign. The test statistic is distributed as an infinite sum of weighted chi-square random variables, with the weights determined by the eigenvalues of some known matrix. Kock & Yang use Satterthwaite's (1941, 1946) chi-square approximation for the nonstandard distribution. Alternatively, Hong (1996) recently proposed an asymptotic one-sided $N(0,1)$ test by extending Haugh's approach via a frequency domain approach. Hong's test is based on the weighted sum of squared residual cross-correlations of various lags, where the weights depend on a kernel function used to estimate the residual coherency function. Typically, larger weights are assigned to lower order lags and smaller weights to higher order lags. Haugh's test can be viewed as a special case using the truncated kernel or uniform weighting. Both theory and simulation show that non-uniform weighting generally has better power than uniform weighting.

While the cross-correlation function provides a natural and convenient tool to investigate causal relationships between two time series, it only captures linear associations between two time series. There exist alternatives in which two time series exhibit zero cross-correlation but are not independent. For instance, some financial time series may have no or little associations with each other in mean but exhibit strong mutual dependence in variance or higher moments (e.g. Engle & Susmel, 1993). For such alternatives, the tests based on the cross-correlation function will have little power. The inconsistency of the cross-correlation based tests is unsatisfying from both theoretical and practical points of view.

In this paper, we extend Hong's (1996) approach and construct an asymptotic one-sided $N(0,1)$ test using the empirical characteristic function. The test is able to detect nonlinear dependencies missed by cross-correlation based tests. We compare the pairwise joint empirical characteristic function of the residuals of two time series with the product of their marginals of various lags, with different weights assigned to different lags. The test is consistent against all pairwise cross-dependencies between two time series, and is particularly powerful against alternatives in which the strength of cross-dependence between two time series decays to zero as the lag increases. Hong's (1996) cross-correlation test can be derived from the present approach by differentiating the empirical characteristic function at the origin. A variety of other new tests can also be obtained by differentiating the empirical characteristic function properly. We note that Csorgo (1985), de Silva & Griffiths (1980), and Feuerverger (1988) used the empirical characteristic function to test independence between two identically independently distributed random variables. In the time series context, Epps (1987, 1988) used the empirical characteristic function to test whether a stationary time series is Gaussian or whether a Gaussian time series is stationary. Our approach differs from these works in many aspects and in fact is a multivariate extension of Hong (1999), who considers hypothesis testing in univariate time series via the empirical characteristic function.

In Section 2 we introduce test statistics. Asymptotic distribution and consistency are established in Sections 3 & 4. In Section 5 a simulation study is conducted to investigate finite sample performance of the new test in comparison with those of Haugh (1976), Hong (1996), and Koch & Yang (1986). The mathematical proofs are collected in the appendix.

2. THE TEST STATISTICS

As in Hong (1996), we consider a bivariate linear process:

*Assumption A.1***:** Suppose (X_t, Y_t) is a bivariate stationary linear process such that

$$X_t = \sum_{j=0}^{\infty} a_j u_{t-j}, \text{ and } Y_t = \sum_{j=0}^{\infty} b_j v_{t-j}, \qquad (t = 1, ..., n)$$

where (u_t) and (v_t) are each an independent and identically distributed sequence respectively, with $E(u_t) = 0, E(v_t) = 0, E(u_t^4) < \infty, E(v_t^4) < \infty$; both (a_j) and (b_j) are sequences of real numbers such that $\sum_{j=0}^{\infty} |a_j| < \infty, \sum_{j=0}^{\infty} |b_j| < \infty$ with $a_0 = b_0 = 1$. Furthermore, both $A(z) = \sum_{j=0}^{\infty} a_j z^j$ and $B(z) = \sum_{j=0}^{\infty} b_j z^j$ are bounded and bounded away from zero for $|z| \leq 1$.

This assumption implies that X_t and Y_t are invertible:

$$X_t = \sum_{j=1}^{\infty} \alpha_j X_{t-j} + u_t, \; Y_t = \sum_{j=1}^{\infty} \beta_j Y_{t-j} + v_t,$$

where $1 - \sum_{j=1}^{\infty} \alpha_j L^j = (\sum_{j=0}^{\infty} a_j L^j)^{-1}$, $1 - \sum_{j=1}^{\infty} \beta_j L^j = (\sum_{j=0}^{\infty} b_j L^j)^{-1}$, and L is the lag operator. This includes as special cases autoregressive moving-average models of finite orders, as assumed in Haugh (1976) and Koch & Yang (1986). Given Assumption A.1, (X_t) and (Y_t) are independent if and only if innovations (u_t) and (v_t) are independent. Following Haugh (1976), Hong (1996), and Koch & Yang (1986), we first prewhiten (X_t) and (Y_t) and test independence between their residuals $(\hat{u}_t)$ and $(\hat{v}_t)$, say. As pointed out in Haugh (1976), this approach is much easier to handle and interpret, because it filters out autodependence within X_t and within Y_t. Otherwise, autodependence would complicate test statistics and affect their finite sample performances. Our approach permits use of the residuals from an autoregressive moving-average model, but misspecification of the autoregressive moving-average model would invalidate the asymptotic distribution of the test statistic. Following Hong (1996), we consider residuals from truncated autoregressions of orders p and q:

$$\hat{u}_t = X_t - \hat{\alpha}(p)'X_t(p),\ \hat{v}_t = Y_t - \hat{\beta}(q)'Y_t(q),$$

where $X_t(p) = (X_{t-1}, ..., X_{t-p})'$, $Y_t(p) = (Y_{t-1}, ..., Y_{t-q})'$, $\hat{\alpha}(p)$ and $\hat{\beta}(q)$ are the ordinary least squares estimators for the truncated autoregressions (cf. Berk, 1974).

Define the pairwise joint empirical characteristic function of $(\hat{u}_t, \hat{v}_{t-j})$

$$\hat{\varphi}_j(x, y) = \begin{cases} (n-j)^{-1}\sum_{t=1+j}^{n} e^{i(x\hat{u}_t + y\hat{v}_{t-j})} & (j \geq 0), \\ (n+j)^{-1}\sum_{t=1-j}^{n} e^{i(x\hat{u}_{t+j} + y\hat{v}_t)} & (j < 0), \end{cases}$$

where $i = \sqrt{-1}$ and $(x, y) \in R^2$. We consider the integrated measure

$$\hat{I}_n^{(m,l)}(j) = \int \left|\hat{\varphi}_j^{(m,l)}(x, y) - \hat{\varphi}_j^{(m,0)}(x, 0)\hat{\varphi}_j^{(0,l)}(0, y)\right|^2 w_1(x)w_2(y)dxdy,$$

where and hereafter the unspecified integral is taken over the entire Euclidean space (R^2 here),

$$\hat{\varphi}_j^{(m,l)}(x, y) = \partial^{m+l}\hat{\varphi}_j(x, y)/\partial x^m \partial y^l,$$

m, l are nonnegative integers, and w_1, w_2 are weighting functions satisfying:

Assumption A.2: For $j = 1, 2$, $w_j : R \to R^+$ is symmetric with $\int x^2 w_j(x)dx < \infty$.

The introduction of $w_1(x)$ and $w_2(y)$ ensures existence of $\hat{I}_n^{(m,l)}(j)$ because $\hat{\varphi}_j(\cdot, \cdot)$ is almost periodic (i.e. reaching the supremum value infinitely often as (x, y) tends to infinity). With proper choices of w_1, w_2, $\hat{I}_n^{(0,0)}(j)$ converges to zero in probability as $n \to \infty$ if and only if u_t and v_{t-j} are independent. Thus, $\hat{I}_n^{(0,0)}(j)$ is able to detect alternatives with zero cross-correlation. On the other hand, various choices of $m, l > 0$ will yield tests for cross-correlation in the (m, l)-moment between two time series; this directs the power of the test toward certain directions of interest.

Put $\hat{\varphi}_u(x) \equiv \hat{\varphi}_0(x, 0)$ and $\hat{\varphi}_v(y) \equiv \hat{\varphi}_0(0, y)$, the empirical marginal characteristic functions of $(\hat{u}_t)$ and $(\hat{v}_t)$. Let M be a positive integer such that $M \to \infty, M/n \to 0$. Our test statistic is a weighted sum of the $\hat{I}_n^{(m,l)}(j)$, namely,

$$C_n(m,l) = \frac{\sum_{j=1-n}^{n-1} k^2(j/M)\{(n-|j|)\hat{I}_n^{(m,l)}(j) - \hat{C}_o^{(m,l)}\}}{\left\{2\hat{D}_o^{(m,l)} \sum_{j=2-n}^{n-2} k^4(j/M)\right\}^{\frac{1}{2}}},$$

where

$$\begin{aligned} \hat{C}_o^{(m,l)} &= \int \{\hat{\varphi}_u^{(2m)}(0) - |\hat{\varphi}_u^{(m)}(x)|^2\} w_1(x)dx \int \{\hat{\varphi}_v^{(2l)}(0) - |\hat{\varphi}_v^{(l)}(y)|^2\} w_2(y)dy, \\ \hat{D}_o^{(m,l)} &= \int\int \left|\hat{\varphi}_u^{(2m)}(x+x') - \hat{\varphi}_u^{(m)}(x)\hat{\varphi}_u^{(m)}(x')\right|^2 w_1(x)w_1(x')dxdx' \\ &\quad \times \int\int \left|\hat{\varphi}_v^{(2l)}(y+y') - \hat{\varphi}_v^{(l)}(y)\hat{\varphi}_v^{(l)}(y')\right|^2 w_2(y)w_2(y')dydy', \end{aligned}$$

and the kernel function k of lag j satisfies:

Assumption A.3: $k : R \to [-1,1]$ is symmetric, and continuous at 0 and all except for a finite number of points, with $k(0) = 1, \int_{-\infty}^{\infty} k^2(z) < \infty$, and $|k(z)| \leq \Delta |z|^{-b}$ as $z \to \infty$ for some $0 < \Delta < \infty$ and $b > \frac{1}{2}$.

This includes the truncated, Bartlett, Daniell, Parzen, Turkey, and Quadratic-Spectral kernels. Cf. Priestley (1981). The truncated, Bartlett, Parzen and Turkey kernels are of compact support, because $k(z) = 0$ if $|z| > 1$. Here, only the $\hat{I}_n^{(m,l)}(j)$ for $|j| \leq M$ are used. In contrast, the Daniell and Quadratic-Spectral kernels are of unbounded support. For these kernels, all $(2n-1)$ $\hat{I}_n^{(m,l)}(j)$ are used. Typically, k gives large weights to small j and small weights to large j. This is expected to give good power against alternatives in which the strength of cross-dependence between X_t and Y_{t-j} decays to zero as the lag increases, as is typically the case for most stationary time series encountered in practice.

We now consider test statistics corresponding to various choices of (m,l).

Case 1 $((m,l) = (0,0))$: *Testing pairwise cross-dependence.*

Let $w_1(\cdot) = w_2(\cdot) = w(\cdot)$, where $w : R \to R^+$ satisfies Assumption A.2. Then

$$C_n(0,0) = \frac{\sum_{j=1-n}^{n-1} k^2(j/M)\{(n-|j|)\hat{I}_n^{(0,0)}(j) - \hat{C}_o^{(0,0)}\}}{\left\{2\hat{D}_o^{(0,0)} \sum_{j=2-n}^{n-2} k^4(j/M)\right\}^{\frac{1}{2}}},$$

where

$$\hat{C}_o^{(0,0)} = \int \{1 - |\hat{\varphi}_u(x)|^2\} w(x) dr \int \{1 - |\hat{\varphi}_v(y)|^2\} w(y) dy$$

$$\hat{D}_o^{(0,0)} = \int\int \left|\hat{\varphi}_u(x+x') - \hat{\varphi}_u(x)\hat{\varphi}_u(x')\right|^2 w(x)w(x') dx dx' \times \int\int \left|\hat{\varphi}_v(y+y') - \hat{\varphi}_v(y)\hat{\varphi}_v(y')\right|^2 w(y)w(y') dy dy'.$$

This is consistent for hypotheses

$$H_o : u_t \text{ and } v_{t-j} \text{ are independent for all } j \text{ versus}$$
$$H_A : u_t \text{ and } v_{t-j} \text{ are dependent for some } j.$$

Here, H_A characterizes all pairwise cross-dependencies. There exist alternatives in which (u_t) and (v_t) are pairwise independent but not mutually independent. The test $C_n(0,0)$ will miss them. As Feller (1957, p.117) points out, however, practical examples of pairwise independent events that are not mutually independent apparently do not exist.

Case 2 $((m,l) = (1,1))$: *Testing cross-correlation in mean.*

Let $w_1(\cdot) = w_2(\cdot) = \delta(\cdot)$, the Dirac delta function. Then $\hat{I}_n^{(1,1)}(j) = \hat{R}_{uv}^2(j)$, where $\hat{R}_{uv}(j)$ is the residual covariance function, namely,

$$\hat{R}_{uv}(j) = \begin{cases} (n-j)^{-1} \sum_{t=1+j}^{n} \{\hat{u}_t - \bar{u}(j)\}\{\hat{v}_{t-j} - \bar{v}(j)\}, & (j \geq 0), \\ (n+j)^{-1} \sum_{t=1-j}^{n} \{\hat{u}_{t+j} - \bar{u}(j)\}\{\hat{v}_t - \bar{v}(j)\}, & (j < 0), \end{cases}$$

with

$$\bar{u}(j) = \begin{cases} (n-j)^{-1} \sum_{t=1+j}^{n} \hat{u}_t, & (j \geq 0), \\ (n+j)^{-1} \sum_{t=1-j}^{n} \hat{u}_{t+j}, & (j < 0), \end{cases}$$

and $\bar{v}(j)$ similarly defined. It turns out that the test statistic

$$C_n(1,1) = \frac{\sum_{j=1-n}^{n-1} k^2(j/M)(n-|j|)\hat{\rho}_{uv}^2(j) - \sum_{j=1-n}^{n-1} k^2(j/M)}{\left\{2\sum_{j=2-n}^{n-2} k^4(j/M)\right\}^{\frac{1}{2}}},$$

where $\hat{\rho}_{uv}(j)$ is the sample cross-correlation function between $(\hat{u}_t)$ and $(\hat{v}_t)$, that is,

$$\hat{\rho}_{uv}(j) = \hat{R}_{uv}(j)/\{\hat{R}_{uu}(0)\hat{R}_{vv}(0)\}^{\frac{1}{2}},$$

with $\hat{R}_{uu}(0) = n^{-1}\sum_{t=1}^{n}\{\hat{u}_t - \bar{u}(0)\}^2$ and $\hat{R}_{vv}(0) = n^{-1}\sum_{t=1}^{n}\{\hat{v}_t - \bar{v}(0)\}^2$. This is a modified version of Hong's (1996) test, in the similar spirit of Ljung & Box's (1978) test being a modified version of Box & Pierce (1970). It is consistent for hypotheses

$$H_o : \rho_{uv}(j) = 0 \text{ for all } j \text{ versus } H_A : \rho_{uv}(j) \neq 0 \text{ for some } j,$$

where $\rho_{uv}(j)$ is the cross-correlation function between (u_t) and (v_t).

Case 3 $((m, l) = (2, 2))$: *Testing cross-correlation in variance.*

Let $w_1(\cdot) = w_2(\cdot) = \delta(\cdot)$, the Dirac delta function. Then

$$C_n(2,2) = \frac{\sum_{j=1-n}^{n-1} k^2(j/M)(n-|j|)\hat{\rho}^2_{u^2v^2}(j) - \sum_{j=1-n}^{n-1} k^2(j/M)}{\left\{2\sum_{j=2-n}^{n-2} k^4(j/M)\right\}^{\frac{1}{2}}},$$

where $\hat{\rho}_{u^2v^2}(j)$ is the sample cross-correlation function between $(\hat{u}_t^2)$ and $(\hat{v}_t^2)$, defined analogously as $\hat{\rho}_{uv}(j)$ with $\hat{u}_t$ and $\hat{v}_t$ replaced by $\hat{u}_t^2$ and $\hat{v}_t^2$ respectively. This test is consistent for hypotheses

$$H_o : \rho_{u^2v^2}(j) = 0 \text{ for all } j \text{ versus } H_A : \rho_{u^2v^2}(j) \neq 0 \text{ for some } j,$$

where $\rho_{u^2v^2}(j)$ is the cross-correlation function between (u_t^2) and (v_t^2). Thus, it is able to detect alternatives that exhibit zero cross-correlation in mean but have non-zero cross-correlation in variance, as occurs in some economic and high frequency financial time series. This test is similar in spirit to those of Granger & Anderson (1978) and McLeod & Li (1983) in testing autocorrelation in the second moment of a univariate time series. In fact, it is closely linked to the test for Granger-causality in variance recently developed in Hong (2001).

Other tests can be obtained by choosing proper (m, l). For example, $(m, l) = (2, 1)$ with $w_1(\cdot)$ and $w_2(\cdot)$ being the Dirac delta function yields a test for cross-correlation between (u_t) and (v_t^2). This can be used to test the causal relationship between (e.g.) the trading volume of a stock and its price volatility.

3. ASYMPTOTIC NULL DISTRIBUTION

To derive the null limit distribution, we impose an additional assumption.

Assumption A.4: $E(u_t^{4d}) < \infty$ and $E(v_t^{4d}) < \infty$, where $d = \max(m, l)$.

Among other things, this ensures that $\varphi_u(\cdot)$ and $\varphi_v(\cdot)$ are $4d$-th order continuously differentiable (cf. Chung, 1974, Theorem 6.4.1). If $m, l \leq 1$, this assumption is reductant given Assumption A.1.

THEOREM 3.1 (Asymptotic Normality). *Suppose that Assumptions A.1 - A.4 hold. Define* $B_u(p) = \sum_{j=p+1}^{\infty} j^{\frac{1}{2}}|\alpha_j|$ *and* $B_v(q) = \sum_{j=q+1}^{\infty} j^{\frac{1}{2}}|\beta_j|$. *Let* $M \to \infty, M/n \to 0$, *and let* $p + q = o(n^{\frac{1}{4}}/M^{\frac{1}{4}}), n^{\frac{1}{2}}\{B_u(p) + B_v(q)\} = o(n^{\frac{1}{4}}/M^{\frac{1}{4}})$. *If* (u_t) *and* (v_t) *are mutually independent, then for each pair* (m, l), $m, l \geq 0$, *we have* $C_n(m, l) \to N(0, 1)$ *in distribution as* $n \to \infty$.

The conditions on p, q are stronger than those of Hong (1996), because we consider a more general statistic here. For the implications of conditions on M, p, q, see an analogous discussion of Hong (1996). Although $(n - |j|)\hat{I}_n^{(m,l)}(j)$ is a Cramer von-Mises type statistic and therefore does not follow a standard chi-square distribution in general, Theorem 1 shows that a weighted sum of the $(n - |j|)\hat{I}_n^{(m,l)}(j)$ has a null asymptotic $N(0, 1)$ distribution for large M. This yields a simple procedure because no tabulation is needed. Our test is one-sided since negative values of $C_n(m, l)$ can occur only under the alternative (see Theorem 2 below). We note that Csorgo (1985), de Silva & Griffiths (1980), and Feuerverger (1988), who used empirical characteristic functions to test independence between two independent and identically distributed random variables, only consider a single or finitely many gridpoints. In contrast, we consider infinite points by integrating out the nuisance parameters (x, y). This ensures consistency of $C_n(0, 0)$ against all pairwise cross-dependencies.

4. CONSISTENCY

To establish consistency of our test $C_n(m, l)$, we first describe the dependence between (u_t) and (v_t).

Assumption A.5: The bivariate process (u_t, v_t) is a stationary mixing sequence with (a) either a uniform mixing coefficient $\phi(j)$ of size 2 or a strong mixing coefficient $\alpha(j)$ of size $\eta/(2 + 2\eta)$ for some $\eta > 0$; (b) $E|u_t|^{4d(1+\eta)} < \infty$ and $E|v_t|^{d(2+2\eta)} < \infty$, where $d = \max(m, l)$.

For mixing conditions and related concepts, see e.g. White (1984). Assumption A.5(a) ensures $\phi(j) \to 0$ or $\alpha(j) \to 0$ as $j \to \infty$, implying ergodicity. Assumption A.5(b) is stronger than Assumption A.4. Among other things, this ensures that the partial derivative of pairwise joint characteristic function $\varphi_j(x,y)$ of (u_t, v_{t-j})

$$\varphi_j^{(m,l)}(x,y) = \frac{\partial^{m+l}}{\partial x^m \partial y^l}\varphi_j(x,y)$$

exists, is bounded and continuous in $(x,y) \in R^2$.

THEOREM 4.1 (Consistency). *Suppose that Assumptions A.1-A.3 and A.5 hold. Let* $M \to \infty, M/n \to 0$, *and let* $p+q = o(n^{\frac{1}{2}}/M^{\frac{1}{4}}), B_u(p) + B_v(q) = o(M^{-\frac{1}{4}})$, *where* $B_u(p)$ *and* $B_v(q)$ *are as in Theorem 1. Then as* $n \to \infty$, *it*

$$(M^{\frac{1}{2}}/n)C_n(m,l) \to \sum_{|j|=0}^{\infty} I^{(m,l)}(j) / \left\{2D_o^{(m,l)} \int_{-\infty}^{\infty} k^4(z)dz\right\}^{\frac{1}{2}}$$

in probability, where $D_o^{(m,l)} = p\lim_{n\to\infty} \widehat{D}_o^{(m,l)}$, *and*

$$I^{(m,l)}(j) = \int\int |\varphi_j^{(m,l)}(x,y) - \varphi_j^{(m,l)}(x,0)\varphi_j^{(m,l)}(0,y)|^2 w_1(x)w_2(y)dxdy.$$

Thus, $\lim_{n\to\infty} P\{C_n(m,l) > c_n\} = 1$ *for any nonstochastic sequence* $\{c_n = o(n/M^{\frac{1}{2}})\}$.

We note that the asymptotic local and global powers of $C_n(m,l)$ can be analyzed using reasoning analogous to those of Hong (1996, Sections 4 & 5). In particular, non-uniform kernels have substantially better power than the truncated kernel or uniform weighting. Over a suitable class of kernel functions, the Daniell kernel, $k(z) = \sin(\pi z)/\pi z$ for $z \in (-\infty, \infty)$, maximizes the power of the test under both proper local and global alternatives. However, simulation results in Hong (1996) show that some commonly used non-uniform kernels have similar powers.

5. MONTE CARLO EVIDENCE

We now examine finite sample performance of $C_n(0,0)$ in comparison with those of Haugh (1976), Hong (1996) and Koch & Yang (1986), using Monte Carlo methods. Following Hong (1996), we consider two processes for X_t and Y_t : (a) $X_t = 0.5X_{t-1} + u_t$ and $Y_t = 0.5Y_{t-1} + v_t$; (b) $X_t = u_t + 0.5u_{t-1}$ and $Y_t = v_t + 0.5v_{t-1}$, where u_t and v_t are independent and identically distributed random variables. Three alternatives are considered:

$$\begin{aligned} &\textit{Alternative } 1 \ : \ \rho_{uv}(j) = \begin{cases} 0.3, & j = 0, \\ 0, & \text{otherwise;} \end{cases} \\ &\textit{Alternative } 2 \ : \ \rho_{uv}(j) = \begin{cases} 0.64^j/(\sum_{j=0}^{8} 0.64^j), & 0 \le j \le 8, \\ 0, & \text{otherwise.} \end{cases} \\ &\textit{Alternative } 3 \ : \ u_t = \epsilon_{1t}\epsilon_{2t},\ v_t = \epsilon_{2t}\epsilon_{3t}, \end{aligned}$$

where the ϵ_{jt} are mutually independent i.i.d. innovations. Under *Alternative 1*, (u_t) and (v_t) are correlated simultaneously but not otherwise. This pattern of very short cross-correlation is similar to those of many financial time series. Under *Alternative 2,* the cross-correlation function is positive and decreasing over $0 \le j \le 8$, with a maximum value of 0.233 at $j = 0$ and nearly 0 at $j = 8$. Similar patterns can be observed in some economic time series data. Under *Alternative 3*, (u_t) and (v_t) are uncorrelated but are not mutually independent.

The simulation experiment was carried out using a GAUSS random number generator on a personal computer. We consider the sample size of $n = 100$. To reduce the effects of initial values, we generate 150 observations and then discard the first 50. We use truncated autoregression models of order p and q to fit X_t and Y_t, with $p, q = 3$. For $C_n(0,0)$, we use the Bartlett kernel: $k(z) = 1 - |z|$ for $|z| \le 1$ and $k(z) = 0$ otherwise. This non-uniform kernel has a compact support, thus significantly reducing the computational cost involved. Three rates of $M = 5, 8, 12$ are used to examine effects of choosing different M. These parameters also apply to Hong's (1996) test

$$Q = \left\{ n \sum_{j=1-n}^{n-1} k^2(j/M)\hat{\sigma}_{uv}^2(j) - S_n(k) \right\} / \left\{2D_n(k)\right\}^{\frac{1}{2}},$$

where $\hat{\sigma}_{uv}(j)$ is the residual cross-correlation function defined as $\hat{\rho}_{uv}(j)$ but without substracting the sample means, $S_n(k) = \sum_{j=1-n}^{n-1}(1-|j|/n)k^2(j/M)$ and $D_n(k) = \sum_{j=2-n}^{n-2}(1-|j|/n)\{1-(|j|+1)/n\}k^4(j/M)$. For comparison, we include $C_n(1,1)$, the modified version of Q.

Haugh's (1976) two statistics are

$$S = n \sum_{j=-M}^{M} \hat{\sigma}_{uv}^2(j),\ S^* = n^2 \sum_{j=-M}^{M} (n-j)^{-1}\hat{\sigma}_{uv}^2(j).$$

TABLE 1.

Rejection Rates Under the Null Hypothesis of Independence Between X_t and Y_t

	$N(0,1)$						EXP					
M	5		8		12		5		8		12	
	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%
$C_n(0,0)$	11.6	6.7	10.9	6.6	10.7	6.4	11.1	7.3	12.3	7.4	12.4	7.0
$C_n(1,1)$	11.3	7.1	10.1	6.0	9.8	5.7	10.1	7.1	10.6	7.1	10.7	7.3
Q	10.8	6.8	9.7	5.8	8.4	5.3	9.7	7.2	10.7	6.7	10.1	6.5
S	7.8	3.6	6.5	2.2	3.9	1.1	8.7	4.2	7.1	3.4	7.1	3.8
S^*	8.7	4.5	7.8	3.3	7.7	3.2	10.2	4.9	9.1	4.5	10.3	6.1
r_1^*	8.6	3.2	6.8	2.1	5.0	2.1	10.1	4.0	7.8	3.8	6.6	3.5
r_{M-1}^*	7.2	4.0	7.2	3.2	6.8	2.5	9.0	3.9	8.9	3.9	7.1	3.2

Notes:
(1) $X_t = 0.5X_{t-1} + u_t, Y_t = 0.5Y_{t-1} + v_t$, where u_t and v_t are $N(0,1)$ or Exponential (0,1);
(2) 1000 Replications;
(3) Q, Hong's (1996) test; S, S^*, Haugh's (1976) tests; r_1^*, r_{M-1}^*, Koch & Yang's (1986) tests.

Both S and S^* are asymptotically χ^2_{2M+1} under the null hypothesis. Koch & Yang's (1986) statistic is

$$r_i^* = n \sum_{j=-M}^{M-i} \left\{ \sum_{l=0}^{i} \hat{\sigma}_{uv}(j+l) \right\}^2, i = 0, 1, ..., M-1.$$

Note $r_0^* = S$. For $i \geq 1$, r_i^* is asymptotically distributed as an infinite sum of chi-square random variables, with weights being the eigenvalues of some known matrix A_i. Koch & Yang (1986) use Satterthwaite's (1941,1946) χ^2 approximation: $\beta_i^{-1} r_i^* \sim \chi^2_{v_i}$, where $\beta_i = tr(A_i A_i')/tr(A_i)$ and $v_i = tr(A_i)^2/tr(A_i A_i')$. Koch & Yang (1986) give formula for β_i and v_i. An important issue with r_i^* is the choice of i given each M. As Koch & Yang (1986) point out, if there is no prior information about the alternative, then various choices of $i = 1, ..., M-1$ should be considered for each M. For brevity, we consider $i = 1$ and $M-1$ only.

Because the performances of each test are much the same whether X_t and Y_t follow process (a) or (b), we only report results when X_t and Y_t follow process (a). Table 1 reports size performances of all the tests at the 10% and 5% nominal significance levels, based on 1000 replications. Both $C_n(0,0)$ and $C_n(1,1)$ have reasonable sizes at the 10% level; they show a little overrejection at the 5% level, especially for exponential innovations. Both $C_n(1,1)$ and Q have similar sizes. The tests S, S^*, r_1^* and r_{M-1}^* all show underrejection at both the 10% and 5% levels in most cases. The S^* test has better size than S.

TABLE 2.

Rejection Rates Under Alternative 1

	$N(0,1)$						EXP					
M	5		8		12		5		8		12	
	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%
$C_n(0,0)$	58.6	53.2	50.2	41.8	40.6	37.8	74.0	72.0	69.8	65.6	63.0	57.8
$C_n(1,1)$	80.8	78.8	75.8	74.2	71.0	69.8	74.8	71.0	69.2	64.6	63.4	60.8
Q	80.8	78.0	75.4	74.2	71.2	70.6	74.6	72.0	69.8	67.4	65.0	61.2
S	41.6	49.8	30.2	41.6	20.6	35.4	42.8	45.8	30.8	38.8	20.6	25.4
S^*	43.6	48.8	33.0	39.6	26.4	32.8	44.2	44.4	34.4	35.4	26.8	22.4
r_1^*	33.2	38.4	26.0	35.0	17.0	31.0	35.0	37.8	27.0	31.8	19.0	23.4
r_{M-1}^*	27.4	34.8	16.6	25.8	10.8	17.6	29.8	30.8	17.4	22.6	10.4	17.4

Notes:
(1) $X_t = 0.5X_{t-1} + u_t, Y_t = 0.5Y_{t-1} + v_t$, where u_t and v_t are $N(0,1)$ or Exponential (0,1), $\rho_{uv}(j) = 0.3$ for $j = 0$ and $\rho_{uv}(j) = 0$ for $j \neq 0$;
(2) 1000 Replications;
(3) Q, Hong's (1996) test; S, S^*, Haugh's (1976) tests; r_1^*, r_{M-1}^*, Koch & Yang's (1986) tests.

Table 2 reports powers under *Alternative 1* at the 5% level, based on 500 replications. We use both asymptotic and empirical critical values, the latter obtained from the 1000 replications under the null hypothesis. The subsequent discussions are based on results using the empirical critical values (i.e. size-corrected powers). We first consider normal innovations. Both Q and $C_n(1,1)$ have virtually the same power and are most powerful, followed by, in their orders, $C_n(0,0), S, S^*, r_1^*$ and r_{M-1}^*. We note that $C_n(0,0)$ becomes more powerful under exponential errors than under normal errors, while all the cross-correlation tests becomes less powerful. For exponential errors, Q remains most powerful, and $C_n(1,1)$ has similar powers. Now, $C_n(0,0)$ has the same power as Q for $M = 5$, and is only slightly less powerful than Q for $M = 8, 12$. The $Q, C_n(1,1)$ and $C_n(0,0)$ tests have substantially better powers than the S, S^*, r_1^* and r_{M-1}^* tests. The r_{M-1}^* test remains the least powerful.

Table 3 reports powers under *Alternative 2*. Under normal errors, r_{M-1}^* is most powerful for $M = 5, 8$, but becomes less powerful than Q and $C_n(1,1)$ for $M = 12$. The Q test is slightly more powerful than $C_n(1,1)$, followed by γ_1^*, S, S^* and $C_n(0,0)$. The $C_n(0,0)$ test is least powerful. Again, $C_n(0,0)$ becomes more powerful under exponential errors than under normal errors; the cross-correlation tests have a mixed story, but their powers change only slightly. Under exponential errors, r_{M-1}^* remains most powerful for $M = 5$, is as powerful as Q for $M = 8$, and becomes less powerful than Q and $C_n(1,1)$ for $M = 12$. The Q test is slightly more powerful than $C_n(1,1)$, which in turn has better powers than r_1^*, S, S^* and $C_n(0,0)$. Now, $C_n(0,0)$

TABLE 3.

Rejection Rates Under Alternative 2

	$N(0,1)$						EXP					
M	5		8		12		5		8		12	
	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%
$C_n(0,0)$	40.6	37.4	37.8	30.8	33.0	29.2	53.2	46.4	46.4	42.6	43.8	39.0
$C_n(1,1)$	70.0	65.6	64.6	62.4	57.4	55.0	70.6	66.0	66.6	61.2	62.8	56.0
Q	71.8	69.8	66.8	64.2	60.6	59.8	73.4	68.6	69.6	66.0	64.6	58.8
S	36.6	44.0	25.6	38.4	18.0	32.6	42.2	47.4	32.0	39.4	23.2	28.0
S^*	39.0	41.6	31.8	36.8	24.0	31.0	44.6	44.8	35.8	36.6	30.0	25.2
r_1^*	55.6	62.2	43.4	53.8	32.0	45.8	59.6	61.6	47.6	51.8	35.6	41.2
r_{M-1}^*	67.8	75.4	58.4	69.0	40.0	52.8	70.2	72.6	61.4	66.0	42.6	51.8

Notes:
(1) $X_t = 0.5X_{t-1} + u_t, Y_t = 0.5Y_{t-1} + v_t$, where u_t and v_t are $N(0,1)$ or Exponential (0,1), $\rho_{uv}(j) = 0.233 * 0.8^{2j}$ for $0 \leq j \leq 8$ and $\rho_{uv}(j) = 0$ otherwise ;
(2) 1000 Replications;
(3) Q, Hong's (1996) test; S, S^*, Haugh's (1976) tests; r_1^*, r_{M-1}^*, Koch & Yang's (1986) tests.

TABLE 4.

Rejection Rates Under Alternative 3

	$N(0,1)$						EXP					
M	5		8		12		5		8		12	
	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%
$C_n(0,0)$	70.4	65.6	67.2	57.4	54.0	49.8	67.8	60.6	59.6	54.6	52.8	46.6
$C_n(1,1)$	21.4	20.4	19.6	17.8	16.2	14.8	31.0	26.2	29.2	25.0	25.8	22.8
Q	21.8	20.4	19.4	17.4	15.4	15.0	30.6	27.0	28.6	26.4	25.4	22.6
S	8.6	10.4	6.8	9.2	4.0	9.8	16.6	18.6	14.0	16.8	13.0	15.0
S^*	9.4	10.4	8.0	9.0	6.6	9.6	18.0	18.4	15.0	16.6	16.4	14.6
r_1^*	8.6	12.0	6.4	9.0	4.8	10.0	14.8	16.0	13.6	14.8	10.6	12.8
r_{M-1}^*	9.6	12.2	5.8	8.4	4.8	8.0	14.2	15.2	10.0	11.6	6.8	9.0

Notes:
(1) $X_t = 0.5X_{t-1} + u_t, Y_t = 0.5Y_{t-1} + v_t$, where $u_t = \varepsilon_{1t}\varepsilon_{3t}$, and $v_t = \varepsilon_{2t}\varepsilon_{3t}, \varepsilon_{1t}, \varepsilon_{2t}$ and ε_{3t} are serially and mutually independent $N(0,1)$ or Exponential (0,1);
(2) 1000 Replications;
(3) Q, Hong's (1996) test; S, S^*, Haugh's (1976) tests; r_1^*, r_{M-1}^*, Koch & Yang's (1986) tests.

remains least powerful for $M = 5$, but becomes more powerful than S and S^* for $M = 8, 12$.

Table 4 reports powers under *Alternative 3*. As expected, $C_n(0,0)$ is able to detect the nonlinear dependence, while all the cross-correlation tests have little or no powers.

In summary, the simulation study shows that the new test $C_n(0,0)$ has reasonable sizes and has good powers against nonlinear dependence which the cross-correlation tests will fail to detect. It also has good power against linear dependencies, especially in non-Gaussian situations.

MATHEMATICAL APPENDIX

Throughout, $0 < \Delta < \infty$ denotes a generic constant that may differ in different places; a^* and $Re(a)$ denote the complex conjugate and the real part of a; $||A|| = \{tr(AA')\}^{\frac{1}{2}}$ denotes the Euclidean norm of a real-valued vector or matrix A; "$\to^p$ " and "$\to^d$ " denote convergence in probability and in distribution. Theorems 1 and 2 are proved in Appendix A; Appendix B contains some lemmas used in Appendix A.

APPENDIX A

PROOF OF THEOREM 1: The proof of Theorem 1 consists of proving Theorems A.1-A.4 below, and application of Slutsky Theorem.

THEOREM A.1. *Let $\tilde{I}_n^{(m,l)}(j)$ be defined as $\hat{I}_n^{(m,l)}(j)$ with $(\hat{u}_t, \hat{v}_t)$ replaced with (u_t, v_t). Then*

$$M^{-\frac{1}{2}} \sum_{|j|=0}^{n-1} k^2(j/M)(n-|j|) \left\{ \hat{I}_n^{(m,l)}(j) - \tilde{I}_n^{(m,l)}(j) \right\} = o_P(1).$$

THEOREM A.2. *Define $\tilde{\gamma}_n^{(m,l)}(j) = \int |\tilde{\gamma}_j^{(m,l)}(x,y)|^2 w_1(x) w_2(y) dx dy$, where*

$$\tilde{\gamma}_j^{(m,l)}(x,y) = \begin{cases} (n-j)^{-1} \sum_{t=1+j}^{n} h_t^{(m)}(x) g_{t-j}^{(l)}(y), & (j \geq 0), \\ (n+j)^{-1} \sum_{t=1-j}^{n} h_{t+j}^{(m)}(x) g_t^{(l)}(y), & (j < 0), \end{cases}$$

$h_t^{(m)}(x) = \frac{\partial^m}{\partial x^m} e^{ixu_t} - \varphi_u^{(m)}(x)$ *and* $g_t^{(l)}(y) = \frac{\partial^l}{\partial y^l} e^{iyv_t} - \varphi_v^{(l)}(y)$. *Then*

$$M^{-\frac{1}{2}} \sum_{|j|=0}^{n-1} k^2(j/M)(n-|j|) \left\{ \tilde{I}_n^{(m,l)}(j) - \tilde{\gamma}_n^{(m,l)}(j) \right\} = o_P(1).$$

THEOREM A.3. *Let* $C_o^{(m,l)}$ *and* $D_o^{(m,l)}$ *be defined as* $\hat{C}_o^{(m,l)}$ *and* $\hat{D}_o^{(m,l)}$ *with* $\hat{\varphi}_u$ *and* $\hat{\varphi}_v$ *replaced with* φ_u *and* φ_v *respectively. Then*

$$\frac{\sum_{|j|=0}^{n-1} k^2(j/M)\left\{(n-|j|)\tilde{\gamma}_n^{(m,l)}(j) - C_o^{(m,l)}\right\}}{\left\{2D_o^{(m,l)}\sum_{|j|=0}^{n-2} k^4(j/M)\right\}^{\frac{1}{2}}} \to^d N(0,1).$$

THEOREM A.4. $\hat{C}_o^{(m,l)} - C_o^{(m,l)} = o_P(M^{-\frac{1}{2}})$, *and* $\hat{D}_o^{(m,l)} - D_o^{(m,l)} = o_P(1)$.

PROOF OF THEOREM A.1: Write

$$\begin{aligned}&|\hat{I}_j^{(m,l)}(x,y)|^2 - |\tilde{I}_j^{(m,l)}(x,y)|^2\\ =&|\hat{I}_j^{(m,l)}(x,y) - \tilde{I}_j^{(m,l)}(x,y)|^2\\ &+2\mathrm{Re}[\{\hat{I}_j^{(m,l)}(x,y) - \tilde{I}_j^{(m,l)}(x,y)\}\tilde{I}_j^{(m,l)}(x,y)^*]. \end{aligned} \tag{A1}$$

We now consider the first term. Put $\hat{h}_t(x) = e^{ix\hat{u}_t} - \varphi_u(x), \hat{g}_t(y) = e^{iy\hat{v}_t} - \varphi_v(y)$. Then by straightforward algebra, we have that for $j \geq 0$,

$$\begin{aligned}\hat{I}_j^{(m,l)}(x,y) =&(n-j)^{-1}\sum_{t=1+j}^{n} \hat{h}_t^{(m)}(x)\hat{g}_{t-j}^{(l)}(y)\\ &-\left\{(n-j)^{-1}\sum_{t=1+j}^{n}\hat{h}_t^{(m)}(x)\right\}\left\{(n-j)^{-1}\sum_{t=1+j}^{n}\hat{g}_{t-j}^{(l)}(y)\right\},\end{aligned} \tag{A2}$$

$$\begin{aligned}\tilde{I}_j(x,y) =&(n-j)^{-1}\sum_{t=1+j}^{n} h_t^{(m)}(x)g_{t-j}^{(l)}(y)\\ &-\left\{(n-j)^{-1}\sum_{t=1+j}^{n}h_t^{(m)}(x)\right\}\left\{(n-j)^{-1}\sum_{t=1+j}^{n}g_{t-j}^{(l)}(y)\right\},\end{aligned} \tag{A3}$$

where $h_t^{(m)}(x)$ and $g_t^{(l)}(y)$ are as in Theorem A.2. Put $\Delta\hat{h}_t(x)=\hat{h}_t(x)-h_t(x)$ and $\Delta\hat{g}_t(y)=\hat{g}_t(y)-g_t(y)$. Then for $j\geq 0$,

$$
\begin{aligned}
&\hat{I}_j^{(m,l)}(x,y)-\tilde{I}_j^{(m,l)}(x,y)\\
=&(n-j)^{-1}\sum_{t=1+j}^{n}\Delta\hat{h}_t^{(m)}(x)\Delta\hat{g}_{t-j}^{(l)}(y)\\
&+(n-j)^{-1}\sum_{t=1+j}^{n}\Delta\hat{h}_t^{(m)}(x)g_{t-j}^{(l)}(y)\\
&+(n-j)^{-1}\sum_{t=1+j}^{n}h_t^{(m)}(x)\Delta\hat{g}_{t-j}^{(l)}(y)\\
&-\left\{(n-j)^{-1}\sum_{t=1+j}^{n}\Delta\hat{h}_t^{(m)}(x)\right\}\left\{(n-j)^{-1}\sum_{t=1+j}^{n}\Delta\hat{g}_{t-j}^{(l)}(x)\right\}\\
&-\left\{(n-j)^{-1}\sum_{t=1+j}^{n}\Delta\hat{h}_t^{(m)}(x)\right\}\left\{(n-j)^{-1}\sum_{t=1+j}^{n}g_{t-j}^{(l)}(x)\right\}\\
&-\left\{(n-j)^{-1}\sum_{t=1+j}^{n}h_t^{(m)}(x)\right\}\left\{(n-j)^{-1}\sum_{t=1+j}^{n}\Delta\hat{g}_{t-j}^{(l)}(x)\right\}\\
=&\sum_{c=1}^{3}\hat{A}_{cnj}(x,y)-\sum_{c=4}^{6}\hat{A}_{cnj}(x,y),\ \text{say.}
\end{aligned}
$$

We can obtain a similar expression for $j<0$. Hence, we have

$$
\left|\hat{I}_j^{(m,l)}(x,y)-\tilde{I}_j^{(m,l)}(x,y)\right|^2\leq\Delta\sum_{c=1}^{6}\left|\hat{A}_{cnj}(x,y)\right|^2. \qquad \text{(A4)}
$$

We first consider $\hat{A}_{1nj}(x,y)$. Put $d_n(j) = k^2(j/M)(1-|j|/n)^{-1}$. By the Cauchy-Swartz inequality, we have

$$\begin{aligned}
&\sum_{|j|=0}^{n-1} k^2(j/M)(n-|j|)\int \left|\hat{A}_{1nj}(x,y)\right|^2 w_1(x)w_2(y)dxdy \\
\leq &\left\{n^{-1}\sum_{|j|=0}^{n-1} d_n(j)\right\}\left\{\sum_{t=1}^{n}\int\left|\Delta\hat{h}_t^{(m)}(x)\right|^2 w_1(x)dx\right\} \\
\times &\left\{\sum_{t=1}^{n}\int\left|\Delta\hat{g}_t^{(l)}(y)\right|^2 w_2(y)dy\right\} \\
= &O_P\left\{(M/n)(p^2+nB_u^2(p))(q^2+nB_v^2(q))\right\} \qquad \text{(A5)}
\end{aligned}$$

by Lemmas B.1-B.2 in Appendix B.

Next, we consider $\hat{A}_{2nj}(x,y)$. Conditional on $(u_t)_{t=1}^n$ and noting that $\{g_t^{(l)}(y)\}$ is an independent and identically distributed sequence with mean zero and $E|g_t^{(l)}(y)|^2 \leq \Delta$ given Assumption A.4, we have

$$\begin{aligned}
&\sum_{|j|=0}^{n-1} k^2(j/M)(n-|j|)\int E\left\{\left|\hat{A}_{2nj}(x,y)\right|^2 |(u_t)_{t=1}^n\right\} w_1(x)dx w_2(y)dy \\
= \; &n^{-1}\sum_{|j|=0}^{n-1} d_n(j)\sum_{t=j+1}^{n}\int|\Delta\hat{h}_t^{(m)}(x)|^2 w_1(x)dx\int E|g_{t-j}^{(l)}(y)|^2 w_2(y)dy \\
\leq \; &\Delta\left\{n^{-1}\sum_{|j|=0}^{n-1} d_n(j)\right\}\left\{\sum_{t=1}^{n}\int|\Delta\hat{h}_t^{(m)}(x)|^2 w_1(x)dx\right\} \\
= \; &O_P\{(M/n)(p^2+B_u^2(p)\}
\end{aligned}$$

by Lemmas B.1-B.2. It follows that

$$\begin{aligned}
&\sum_{|j|=0}^{n-1} k^2(j/M)(n-|j|)\int \left|\hat{A}_{2nj}(x,y)\right|^2 w_1(x)w_2(y)dxdy \\
= &O_P\{(M/n)(p^2+pB_u^2(p))\}. \qquad \text{(A6)}
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\sum_{|j|=0}^{n-1} k^2(j/M)(n-|j|)\int \left|\hat{A}_{3nj}(x,y)\right|^2 w_1(x)w_2(y)dxdy \\
= &O_P\{(M/n)(q^2+nB_v^2(q)\}. \qquad \text{(A7)}
\end{aligned}$$

We now turn to $\hat{A}_{4nj}(x,y)$. By the Cauchy-Schwarz inequality, we have

$$\begin{aligned} |\hat{A}_{4nj}(x,y)|^2 &= \left|(n-j)^{-1}\sum_{t=j+1}^{n}\Delta\hat{h}_t^{(m)}(x)\right|^2 \left|(n-j)^{-1}\sum_{t=j+1}^{n}\Delta\hat{g}_{t-j}^{(l)}(y)\right|^2 \\ &\leq (n-j)^{-2}\left\{\sum_{t=1}^{n}|\Delta\hat{h}_t^{(m)}(x)|^2\right\}\left\{\sum_{t=1}^{n}|\Delta\hat{g}_t^{(l)}(x)|^2\right\}, \end{aligned}$$

it follows that

$$\begin{aligned} &\sum_{|j|=0}^{n-1}k^2(j/M)(n-|j|)\int|\hat{A}_{4nj}(x,y)|^2w_1(x)w_2(y)dxdy \\ \leq&\left\{n^{-1}\sum_{|j|=0}^{n-1}d_n(j)\right\}\left\{\sum_{t=1}^{n}\int|\Delta\hat{h}_t^{(m)}(x)|^2w_1(x)dx\right\} \\ \times&\left\{\sum_{t=j+1}^{n}\int|\Delta\hat{g}_t^{(l)}(y)|^2w_2(y)dy\right\} \\ =&O_P\left\{(M/n)(p^2+nB_u^2(p))(q^2+nB_v^2(q))\right\} \qquad \text{(A8)} \end{aligned}$$

by Lemmas B.1-B.2. Similarly,

$$\begin{aligned} &\sum_{|j|=0}^{n-1}k^2(j/M)(n-|j|)\int|\hat{A}_{5nj}(x,y)|^2w_1(x)w_2(y)dxdy \\ =&O_P\{(M/n)(p^2+nB_u^2(p))\}, \qquad \text{(A9)} \\ &\sum_{|j|=0}^{n-1}k^2(j/M)(n-|j|)\int|\hat{A}_{6nj}(x,y)|^2w_1(x)w_2(y)dxdy \\ =&O_P\{(M/n)(q^2+nB_v^2(q))\}. \qquad \text{(A10)} \end{aligned}$$

Combining (A4)-(A10), we obtain

$$\begin{aligned} &\sum_{|j|=0}^{n-1}k^2(j/M)(n-|j|)\int\left|\hat{I}_j^{(m,l)}(x,y)-\tilde{I}_j^{(m,l)}(x,y)\right|^2 w_1(x)w_2(y)dxdy \\ =&O_P\left\{(M/n)(p^2+nB_u^2(p))(q^2+nB_v^2(q))\right\}. \qquad \text{(A11)} \end{aligned}$$

This term vanishes in probability given $p+q=o(n^{\frac{1}{4}}/M^{\frac{1}{4}}), n^{\frac{1}{2}}\{B_u(p)+B_v(q)\}=o(n^{\frac{1}{4}}/M^{\frac{1}{4}})$.

Next, by the Cauchy-Schwarz inequality, we have

$$\sum_{|j|=0}^{n-1} k^2(\frac{j}{M})(n-|j|) \int \left| \{\hat{I}_j^{(m,l)}(x,y) - \tilde{I}_j^{(m,l)}(x,y)\} \tilde{I}_j^{(m,l)}(x,y)^* \right| w_1(x) w_2(y) dx dy$$
$$\leq \left\{ \sum_{|j|=0}^{n-1} k^2(\frac{j}{M})(n-|j|) \int \left| \hat{I}_j^{(m,l)}(x,y) - \tilde{I}_j^{(m,l)}(x,y) \right|^2 w_1(x) w_2(y) dx dy \right\}^{\frac{1}{2}}$$
$$\times \left\{ \sum_{|j|=0}^{n-1} k^2(\frac{j}{M})(n-|j|) \tilde{I}^{(m,l)}(j) \right\}^{\frac{1}{2}}$$
$$= o_P(M^{\frac{1}{2}}) \quad \text{(A12)}$$

by (A11) and $\sum_{|j|=0}^{n-1} k^2(j/M)(n-|j|)\tilde{I}_n^{(m,l)}(j) = O_P(M)$, where the latter follows by Markov's inequality. Therefore, we have $M^{-\frac{1}{2}} \sum_{|j|=0}^{n-1} k^2(j/M)(n-|j|)\{\hat{I}_n^{(m,l)}(j) - \tilde{I}_n^{(m,l)}(j)\} = o_P(1)$ from (A1) and (A11)-(A12). This completes the proof. ▮

PROOF OF THEOREM A.2: Write

$$|\tilde{I}_j^{(m,l)}(x,y)|^2 - |\tilde{\gamma}_j^{(m,l)}(x,y)|^2$$
$$= |\tilde{I}_j^{(m,l)}(x,y) - \tilde{\gamma}_j^{(m,l)}(x,y)|^2 + 2\mathrm{Re}[\{\tilde{I}_j^{(m,l)}(x,y) - \tilde{\gamma}_j^{(m,l)}(x,y)\}\tilde{\gamma}_j^{(m,l)}(x,y)^*], \quad \text{(A13)}$$

By the Cauchy-Schwarz inequality and (A3), we have that for $j \geq 0$,

$$E\left|\tilde{I}_j^{(m,l)}(x,y) - \tilde{\gamma}_j^{(m,l)}(x,y)\right|^2$$
$$\leq \left\{ E\left|(n-j)^{-1} \sum_{t=1+j}^{n} h_t^{(m)}(x)\right|^4 E\left|(n-j)^{-1} \sum_{t=1+j}^{n} g_{t-j}^{(l)}(y)\right|^4 \right\}^{\frac{1}{2}}$$
$$\leq \Delta(n-j)^{-2}$$

given that $\{h_t^{(m)}(x)\}$ and $\{g_t^{(l)}(y)\}$ are each an independent and identically distributed sequence with fourth moments finite by Assumption A.4. A similar result holds for $j < 0$. Also, by the Cauchy-Schwarz inequality,

$$E\left|\{\tilde{I}_j^{(m,l)}(x,y) - \tilde{\gamma}_j^{(m,l)}(x,y)\}\tilde{\gamma}_j^{(m,l)}(x,y)^*\right|$$
$$\leq \left\{ E|\tilde{I}_j^{(m,l)}(x,y) - \tilde{\gamma}_j^{(m,l)}(x,y)|^2 E|\tilde{\gamma}_j^{(m,l)}(x,y)|^2 \right\}^{\frac{1}{2}}$$
$$\leq \Delta(n-|j|)^{-3/2},$$

where $E|\tilde{\gamma}_j^{(m,l)}(x,y)|^2 \leq \Delta(n-|j|)^{-1}$ because $\tilde{\gamma}_j^{(m,l)}(x,y)$ is a sum of an martingale difference sequence with the second moment finite. It follows that

$$\begin{aligned} &\sum_{|j|=0}^{n-1} k^2(j/M)(n-|j|)E\left|\tilde{I}_n^{(m,l)}(j) - \tilde{\gamma}_n^{(m,l)}(j)\right| \\ \leq\ & \Delta(M/n^{\frac{1}{2}})\left\{M^{-1}\sum_{|j|=0}^{n-1} d_n(j)\right\} \\ =\ & o(M^{\frac{1}{2}}) \end{aligned}$$

by Lemma B.1 and $M/n \to 0$. The desired result follows by Markov's inequality. ▌

PROOF OF THEOREM A.3: Put

$$H_{t,s}^{(m)} = \int h_t^{(m)}(x)h_s^{(m)}(x)^* w_1(x)dx,\ G_{t,s}^{(l)} = \int g_t^{(l)}(y)g_s^{(l)}(y)^* w_2(y)dy. \tag{A14}$$

Recall $\tilde{\gamma}_n^{(m,l)}(j) = \int |\tilde{\gamma}_j^{(m,l)}(x,y)|^2 w_1(x)w_2(y)dxdy$ as in Theorem A.2, we can write

$$\begin{aligned} &(n-|j|)\tilde{\gamma}_n^{(m,l)}(j) \\ =\ & \begin{cases} \dfrac{\sum_{t=1+j}^{n} H_{t,t}^{(m)}G_{t-j,t-j}^{(l)}}{n-j} + \dfrac{2\sum_{t=2+j}^{n}\sum_{s=1+j}^{t-1} Re(H_{t,s}^{(m)}G_{t-j,s-j}^{(l)})}{n-j} & (j \geq 0) \\ \dfrac{\sum_{t=1-j}^{n} H_{t+j,t+j}^{(m)}G_{t,t}^{(l)}}{n+j} + \dfrac{2\sum_{t=2-j}^{n}\sum_{s=1-j}^{t-1} Re(H_{t+j,s+j}^{(m)}G_{t,s}^{(l)})}{n+j} & (j < 0) \end{cases} \\ =\ & \bar{C}_n(j) + \bar{S}_n(j),\ \text{say.} \end{aligned}$$

It follows that

$$\sum_{|j|=0}^{n-1} k^2(j/M)(n-|j|)\tilde{\gamma}_n^{(m,l)}(j) = \sum_{|j|=0}^{n-1} k^2(j/M)\bar{C}_n(j) + \sum_{|j|=0}^{n-1} k^2(j/M)\bar{S}_n(j). \tag{A15}$$

Under the null hypothesis of independence between (u_t) and (v_t), $\bar{C}_n(j)$ is the sum of an independent and identically distributed sequence with

$$\begin{aligned} E\bar{C}_n(j) &= E\{H_{1,1}^{(m)}\}E\{G_{1,1}^{(l)}\} \\ &= \int E\left|h_1^{(m)}(x)\right|^2 w_1(x)dx \int E\left|g_1^{(l)}(y)\right|^2 w_2(y)dy \\ &= \int \{\varphi_u^{(2m)}(0) - |\varphi_u^{(m)}(x)|^2\}w_1(x)dx \int \{\varphi_v^{(2m)}(0) - |\varphi_v^{(m)}(y)|^2\}w_2(y)dy \\ &= C_o^{(m,l)}. \end{aligned}$$

Hence, we have, by the Cauchy-Schwarz inequality, that

$$\begin{aligned} & E\left\{\bar{C}_n(j) - C_o^{(m,l)}\right\}^2 \\ \le\; & (n-|j|)^{-1}E\{H_{1,1}^{(m)}\}^2E\{G_{1,1}^{(l)}\}^2 \\ \le\; & (n-|j|)^{-1}\left\{\int [E|h_1^{(m)}(x)|^4]^{\frac{1}{2}} w_1(x)dx\right\}^2 \left\{\int [E|g_1^{(l)}(y)|^4]^{\frac{1}{2}} w_2(y)dy\right\}^2 \\ \le\; & \Delta(n-|j|)^{-1} \end{aligned}$$

given Assumption A.4. This implies

$$\begin{aligned} \sum_{|j|=0}^{n-1} k^2(j/M)E\left|\bar{C}_n(j) - C_o^{(m,l)}\right| &\le \sum_{|j|=0}^{n-1} k^2(j/M)\left[E\{\bar{C}_n(j) - C_o^{(m,l)}\}^2\right]^{\frac{1}{2}} \\ &\le \Delta(M/n^{\frac{1}{2}})\left\{M^{-1}\sum_{|j|=0}^{n-1} d_n(j)\right\} \\ &= o(M^{\frac{1}{2}}) \end{aligned}$$

given Lemma B.1 and $M/n \to 0$. It follows that

$$\sum_{|j|=0}^{n-1} k^2(j/M)\bar{C}_n(j) = C_o^{(m,l)}\sum_{|j|=0}^{n-1} k^2(j/M) + o_P(M^{\frac{1}{2}}). \tag{A16}$$

Now, put $\bar{S}_n = \sum_{|j|=0}^{n-1} k^2(j/M)\bar{S}_n(j)$ and noting $d_n(j) = d_n(-j)$, we can write

$$\begin{aligned}
\bar{S}_n =& 2n^{-1}\sum_{j=1}^{n-1} d_n(j) \sum_{t=2+j}^{n}\sum_{s=1+j}^{t-1} Re(H_{t,s}^{(m)}G_{t-j,s-j}^{(l)}) \\
&+2n^{-1}\sum_{j=0}^{1-n} d_n(j) \sum_{t=2-j}^{n}\sum_{s=1-j}^{t-1} Re(H_{t+j,s+j}^{(m)}G_{t,s}^{(l)}) \\
=& 2n^{-1}\sum_{j=1}^{n-1} d_n(j) \sum_{t=2+j}^{n}\sum_{s=1+j}^{t-1} Re(H_{t,s}^{(m)}G_{t-j,s-j}^{(l)}) \\
&+2n^{-1}\sum_{j=0}^{n-1} d_n(j) \sum_{t=2+j}^{n}\sum_{s=1+j}^{t-1} Re(H_{t-j,s-j}^{(m)}G_{t,s}^{(l)}) \\
=& 2n^{-1}\sum_{t=3}^{n-1}\left\{\sum_{s=2}^{t-1}\sum_{j=1}^{s-1} d_n(j) Re(H_{t,s}^{(m)}G_{t-j,s-j}^{(l)})\right\} \\
&+2n^{-1}\sum_{t=2}^{n-1}\left\{\sum_{s=1}^{t-1}\sum_{j=0}^{s-1} d_n(j) Re(H_{t-j,s-j}^{(m)}G_{t,s}^{(l)})\right\} \\
=& n^{-1}\sum_{t=2}^{n-1}(S_{1nt}+S_{2nt}) \\
=& \bar{S}_{1n}+\bar{S}_{2n},\ \text{say,}
\end{aligned} \tag{A17}$$

where we set $S_{1n2} = 0$. Given (A14) and independence between (u_t) and (v_t), we have $E(S_{1nt}|F_{t-1}) = 0, E(S_{2nt}|F_{t-1}) = 0$ and $E(S_{1nt}S_{2nt}|F_{t-1}) = 0$, where $\{F_t\}$ is the sequence of sigma-fields consisting of (u_s, v_s), $s \le t$. It follows that $E(S_{nt}|F_{t-1}) = 0$, and so (S_{nt}, F_t) is an adapted martingale difference sequence. We can show asymptotic normality for $\bar{S}_n$ by Brown's (1971) martingale limit theorem, which implies $V(\bar{S}_n)^{-\frac{1}{2}}\bar{S}_n \to^d N(0,1)$ if

$$V(\bar{S}_n)^{-\frac{1}{2}} n^{-2}\sum_{t=2}^{n} E\left[S_{nt}^2 1\left\{|S_{nt}| > n\epsilon V(\bar{S}_n)^{\frac{1}{2}}\right\}\right] \to 0 \tag{A18}$$

for every constant $\epsilon > 0$, and

$$V(\bar{S}_n)^{-\frac{1}{2}} n^{-2}\sum_{t=2}^{n} E(S_{nt}^2|F_{t-1}) \to^p 1. \tag{A19}$$

Here, $1(\cdot)$ is the indicator function. We verify (A18) by showing

$$V(\bar{S}_n)^{-2} n^{-4} \sum_{t=2}^{n} E S_{nt}^4 \to 0.$$

From Lemma B.4 in Appendix B,

$$V(\bar{S}_n) = 2 D_o^{(m,l)} M \int_{-\infty}^{\infty} k^4(z) dz \{1 + o(1)\}$$

as $M \to \infty, M/n \to 0$. Hence, it suffices to show

$$M^{-2} n^{-4} \sum_{t=2}^{n} E|\tilde{S}_{cnt}|^4 \to 0$$

for $c = 1, 2$, where

$$\tilde{S}_{1nt} = 2 \sum_{s=2}^{t-1} \sum_{j=1}^{s-1} d_n(j) H_{t,s}^{(m)} G_{t-j,s-j}^{(l)}, \ \tilde{S}_{2nt} = 2 \sum_{s=1}^{t-1} \sum_{j=0}^{s-1} d_n(j) H_{t-j,s-j}^{(m)} G_{t,s}^{(l)}. \tag{A20}$$

We show for $\tilde{S}_{1nt}$ only; the proof for $\tilde{S}_{2nt}$ is the same. Rewrite

$$\begin{aligned} \tilde{S}_{1nt} =& 2 \int h_t^{(m)}(x) \sum_{s=2}^{t-1} h_s^{(m)}(x)^* \left\{ \sum_{j=1}^{s-1} d_n(j) G_{t-j,s-j}^{(l)} \right\} w_1(x) dx \\ =& 2 \int h_t^{(m)}(x) \sum_{s=2}^{t-1} h_s^{(m)}(x)^* Z_{ts}^{(l)} w_1(x) dx, \text{ say.} \end{aligned} \tag{A21}$$

Note that $h_s^{(m)}(x)$ and $Z_{ts}^{(l)}$ are independent, and given (t, s) where $t > s$, $Z_{ts}^{(l)}$ is a sum of an martingale difference sequence. Then by Minkowski's inequality, we obtain

$$
\begin{aligned}
& E\left|\tilde{S}_{1nt}\right|^4 \\
\leq\ & 16\left[\int\left\{E|h_t^{(m)}(x)|^4\right\}^{\frac{1}{4}}\left\{E\left|\sum_{s=2}^{t-1}h_s^{(m)}(x)^* Z_{ts}^{(l)}\right|^4\right\}^{\frac{1}{4}} w_1(x)dx\right]^4 \\
\leq\ & 16\left[\int\left\{E|h_t^{(m)}(x)|^4\right\}^{\frac{1}{4}}\left\{\sum_{s=2}^{t-1}\left\{E|h_s^{(m)}(x)|^4\right\}^{\frac{1}{2}}\left\{E|Z_{ts}^{(l)}|^4\right\}^{\frac{1}{2}}\right\}^{\frac{1}{2}} w_1(x)dx\right]^4 \\
\leq\ & 16t^2M^2\left\{M^{-1}\sum_{j=1}^{n-1}d_n(j)\right\}^2\left[\int\left\{E|h_1^{(m)}(x)|^4\right\}^{\frac{1}{4}} w_1(x)dx\right]^2 \\
& \times\left[\left\{E|g_1^{(l)}(y)|^4\right\}^{\frac{1}{4}} w_2(y)dy\right]^2 \\
=\ & O(t^2M^2),
\end{aligned}
$$

where, we have made use of the fact that

$$
E|Z_{ts}^{(l)}|^4 \leq \left\{\sum_{j=1}^{s-1}d_n(j)\left(E|G_{t-j,s-j}^{(l)}|^4\right)^{\frac{1}{2}}\right\}^2 \leq \Delta M^2\left\{M^{-1}\sum_{j=1}^{n-1}d_n(j)\right\}^2. \tag{A22}
$$

It follows that

$$
M^{-2}n^{-4}\sum_{t=2}^{n}E|\tilde{S}_{1nt}|^4 = O(n^{-1}).
$$

Similarly,

$$
M^{-2}n^{-4}\sum_{t=2}^{n}E|\tilde{S}_{2nt}|^4 = O(n^{-1}).
$$

Hence, condition (A18) holds.

Next, we verify (A19) by showing

$$
V(\bar{S}_n)^{-2}V\{n^{-2}\sum_{t=2}^{n}E(S_{nt}^2|F_{t-1})\} \to 0.
$$

Because $E(S_{nt}^2|F_{t-1}) = E(S_{1nt}^2|F_{t-1}) + E(S_{2nt}^2|F_{t-1})$, it suffices to show

$$
M^{-2}V\{n^{-2}\sum_{t=2}^{n}E(S_{cnt}^2|F_{t-1})\} \to 0
$$

for $c = 1, 2$. We show for $c = 1$; the proof for $c = 2$ is the same. Given

$$
\begin{aligned}
S_{1nt}^2 &= \left\{ 2\sum_{s=2}^{t-1}\sum_{j=1}^{s-1} d_n(j)[H_{t,s}^{(m)} G_{t-j,s-j}^{(l)} + H_{t,s}^{(m)*} G_{t-j,s-j}^{(l)*}] \right\}^2 \\
&= \left\{ 2\sum_{s=2}^{t-1}\sum_{j=1}^{s-1} d_n(j) H_{t,s}^{(m)} G_{t-j,s-j}^{(l)} \right\}^2 + \left\{ 2\sum_{s=2}^{t-1}\sum_{j=1}^{s-1} d_n(j) H_{t,s}^{(m)*} G_{t-j,s-j}^{(l)*} \right\}^2 \\
&\quad + 2\left| 2\sum_{s=2}^{t-1}\sum_{j=1}^{s-1} d_n(j) H_{t,s}^{(m)} G_{t-j,s-j}^{(l)} \right|^2 \\
&= \tilde{S}_{1nt}^2 + (\tilde{S}_{1nt}^*)^2 + 2|\tilde{S}_{1nt}|^2, \qquad \text{(A23)}
\end{aligned}
$$

it suffices to show $M^{-2}V\{n^{-2}\sum_{t=2}^{n} E(\tilde{S}_{1nt}^2|F_{t-1})\} \to 0$. Using (A21) and putting $D_u(x, x') \equiv E h_t^{(m)}(x) h_t^{(m)}(x')$, we have

$$
\begin{aligned}
&E(\tilde{S}_{1nt}^2|F_{t-1}) \\
=&E\left[\left\{ 2\sum_{s=2}^{t-1}\sum_{j=1}^{s-1} d_n(j) H_{t,s}^{(m)} G_{t-j,s-j}^{(l)} \right\}^2 |F_{t-1} \right] \\
=&4\sum_{s=2}^{t-1} \int D_u(x, x') h_s^{(m)}(x)^* h_s^{(m)}(x')^* (Z_{ts}^{(l)})^2 w_1(x) w_1(x') dx dx' \\
&+ 8\sum_{s_2=3}^{t-1} \int D_u(x, x') h_{s_2}^{(m)}(x)^* \sum_{s_1=2}^{s_2-1} h_{s_1}^{(m)}(x')^* Z_{ts_2}^{(l)} Z_{ts_1}^{(l)} w_1(x) w_2(x') dx dx' \\
=&4B_{1nt} + 8A_{1nt}, \text{ say.} \qquad \text{(A24)}
\end{aligned}
$$

We now consider the second term A_{1nt}, which is a sum of an martingale difference sequence over s_2. For $t_2 > t_1$,

$$\begin{aligned}
&E(A_{1nt_2}A_{1nt_1}^*)\\
=&\sum_{s_2=3}^{t_1-1} E\int D_u(x,x')D_u(\tilde{x},\tilde{x}')^*D_u(-x,\tilde{x})\left\{\sum_{s_1=2}^{s_2-1} h_{s_1}^{(m)}(x')^* Z_{t_2s_2}^{(l)}Z_{t_2s_1}^{(l)}\right\}\\
\times&\left\{\sum_{s_1=2}^{s_2-1} h_{s_1}^{(m)}(\tilde{x}')Z_{t_1s_2}^{(l)*}Z_{t_1s_1}^{(l)*}\right\} w_1(x)w_1(x')w_1(\tilde{x})w_1(\tilde{x}')dxdx'd\tilde{x}d\tilde{x}'\\
=&\int D_u(x,x')D_u(\tilde{x},\tilde{x}')^*D_u(-x,\tilde{x})D_u(-x',\tilde{x}')\\
\times&\sum_{s_2=3}^{t_1-1}\sum_{s_1=2}^{s_2-1} E\{Z_{t_2s_2}^{(l)}Z_{t_2s_1}^{(l)}Z_{t_1s_2}^{(l)*}Z_{t_1s_1}^{(l)*}\}w_1(x)w_1(x')w_1(\tilde{x})w_1(\tilde{x}')dxdx'd\tilde{x}d\tilde{x}',
\end{aligned}$$

where the second equality follows by independence between (u_t) and (v_t). Noting $Z_{ts}^{(l)} = \sum_{j=1}^{s-1} d_n(j)G_{t-j,s-j}^{(l)}$ and by straightforward but tedious algebra, we have that for $t_2 > s_2, t_1 > s_1$,

$$\left|E(Z_{t_2s_2}^{(l)}Z_{t_2s_1}^{(l)}Z_{t_1s_2}^{(l)*}Z_{t_1s_1}^{(l)*})\right| \le \begin{cases} \Delta M^2\left\{M^{-1}\sum_{j=1}^{n-1} d_n(j)\right\}^2, & (t_2=t_1)\\ \Delta M\left\{M^{-1}\sum_{j=1}^{n-1} d_n(j)\right\}, & (t_2>t_1).\end{cases}$$

Therefore, we have

$$\begin{aligned}
M^{-2}E\left|n^{-2}\sum_{t=2}^n A_{1nt}\right|^2 =&M^{-2}n^{-4}\sum_{t=2}^n E|A_{1nt}|^2\\
&+2M^{-2}n^{-4}\sum_{t_2=3}^n\sum_{t_1=2}^{t_2-1} E(A_{1nt_2}A_{1nt_1}^*)\\
=&O(n^{-1}+M^{-1}). \qquad \text{(A25)}
\end{aligned}$$

Next, we consider the first term in (A24). Put $\bar{D}_u \equiv \int |D_u(x,x')|^2 w_1(x)w_1(x')dxdx'$. Then

$$\begin{aligned}
&B_{1nt}\\
=&\bar{D}_u\sum_{s=2}^{t-1}(Z_{ts}^{(l)})^2\\
+&\sum_{s=2}^{t-1}\int D_u(x,x')\{h_s^{(m)}(x)^*h_s^{(m)}(x')^* - D_u(x,x')^*\}(Z_{ts}^{(l)})^2 w_1(x)w_1(x')dxdx'\\
=&B_{2nt}+A_{2nt},\ \text{say.} \qquad \text{(A26)}
\end{aligned}$$

Conditional on $(v_t)_{t=1}^n$, A_{2nt} is a sum of an independent and identically distributed sequence. It follows that

$$\begin{aligned} & E|A_{2nt}|^2 \\ = & \sum_{s=2}^{t-1} E\left|\int D_u(x,x')\{h_s^{(m)}(x)h_s^{(m)}(x') - D_u(x,x')\}^*(Z_{1ts}^{(l)})^2 w_1(x)w_1(x')dxdx'\right|^2 \\ \leq & \sum_{s=2}^{t-1}\left[\int |D_u(x,x')|\{E|h_1^{(m)}(x)h_1^{(m)}(x')|^2\}^{\frac{1}{2}}\{E|Z_{1ts}^{(l)}|^4\}^{\frac{1}{2}} w_1(x)w_1(x')dxdx'\right]^2 \\ \leq & \Delta t M^2\left\{M^{-1}\sum_{j=1}^{n-1} d_n(j)\right\}^2 \end{aligned}$$

given $E|Z_{ts}^{(l)}|^4 \leq \Delta M^2\left\{M^{-1}\sum_{j=1}^{n-1} d_n(j)\right\}^2$ as in (A22). Therefore,

$$M^{-2}E\left|n^{-2}\sum_{t=2}^{n} A_{2nt}\right|^2 \leq M^{-2}n^{-4}\left\{\sum_{t=2}^{n}(E|A_{2nt}|^2)^{\frac{1}{2}}\right\}^2 = O(n^{-1}). \quad \text{(A27)}$$

Now, we consider the first term in (A26). Recall $Z_{ts}^{(l)} = \sum_{j=1}^{s-1} d_n(j)G_{t-j,s-j}^{(l)}$ as in (A21), we decompose

$$\begin{aligned} B_{2nt} =& \bar{D}_u\left\{\sum_{s=2}^{t-1}\sum_{j=1}^{s-1} d_n^2(j)\{G_{t-j,s-j}^{(l)}\}^2\right\} \\ & + 2\bar{D}_u\left\{\sum_{s=3}^{t-1}\sum_{j_2=2}^{s-1}\sum_{j_1=1}^{j_2-1} d_n(j_1)d_n(j_2)G_{t-j_1,s-j_1}^{(l)}G_{t-j_2,s-j_2}^{(l)}\right\} \\ =& B_{3nt} + 2A_{3nt}, \text{ say.} \end{aligned} \quad \text{(A28)}$$

Recall $G_{t,s}^{(l)} = \int g_t^{(l)}(y)g_s^{(l)}(y)^* w_2(y)dy$ as in (A14), we have,

$$A_{3nt} = \bar{D}_u\sum_{j_1=1}^{t-3} d_n(j_1)\left\{\sum_{j_2=j_1+1}^{t-2} d_n(j_2)\sum_{s=j_2+1}^{t-1} G_{t-j_1,s-j_1}^{(l)}G_{t-j_2,s-j_2}^{(l)}\right\},$$

which is a sum of an martingale difference sequence over j_1. It follows that

$$\begin{aligned}
&E\left|A_{3nt}\right|^2\\
=\ &\bar{D}_u^2\sum_{j_1=1}^{t-3}d_n^2(j_1)E\left|\sum_{j_2=j_1+1}^{t-2}d_n(j_2)\sum_{s=j_2+1}^{t-1}G_{t-j_1,s-j_1}^{(l)}G_{t-j_2,s-j_2}^{(l)}\right|^2\\
\leq\ &\bar{D}_u^2\sum_{j_1=1}^{t-3}d_n^2(j_1)\left\{\sum_{j_2=j_1+1}^{t-2}d_n(j_2)\left(E\left|\sum_{s=j_2+1}^{t-1}G_{t-j_1,s-j_1}^{(l)}G_{t-j_2,s-j_2}^{(l)}\right|^2\right)^{\frac{1}{2}}\right\}^2\\
\leq\ &\Delta tM^3\left\{M^{-1}\sum_{j_1=1}^{n-1}d_n^2(j_1)\right\}\left\{M^{-1}\sum_{j_2=1}^{n-1}d_n(j_2)\right\}^2\\
=\ &O(tM^3),
\end{aligned}$$

where the first inequality follows by Minkowski's inequality and the second one by $E|\sum_{s=j_2+1}^{t-1}G_{t-j_1,s-j_1}^{(l)}G_{t-j_2,s-j_2}^{(l)}|^2\leq\Delta t$ for $t>s>j_2>j_1$. Therefore,

$$M^{-2}E\left|n^{-2}\sum_{t=2}^{n}A_{3nt}\right|^2\leq M^{-2}n^{-4}\left\{\sum_{t=2}^{n}(E|A_{3nt}|^2)^{\frac{1}{2}}\right\}^2=O(M/n). \tag{A29}$$

Finally, noting $E\{G_{t-j,s-j}^{(l)}\}^2=\int|D_v(y,y')|^2w_2(y)w_2(y')dydy'\equiv\bar{D}_v$, where $D_v(y,y')=Eg_t^{(l)}(y)g_t^{(l)}(y')$, we have

$$\begin{aligned}
&B_{3nt}\\
=&\bar{D}_u\sum_{s=2}^{t-1}\sum_{j=1}^{s-1}d_n^2(j)E\{G_{t-j,s-j}^{(l)}\}^2+\sum_{s=2}^{t-1}\sum_{j=1}^{s-1}d_n^2(j)\left[\{G_{t-j,s-j}^{(l)}\}^2-E\{G_{t-j,s-j}^{(l)}\}^2\right]\\
=&\bar{D}_u\bar{D}_v\sum_{s=2}^{t-1}\sum_{j=1}^{s-1}d_n^2(j)\\
+&\bar{D}_u\sum_{s=2}^{t-1}\sum_{j=1}^{s-1}d_n^2(j)\int\{g_{t-j}^{(l)}(y)g_{t-j}^{(l)}(y')-D_v(y,y')\}g_{s-j}^{(l)}(y)^*g_{s-j}^{(l)}(y')^*w_2(y)w_2(y')\\
+&\bar{D}_u\sum_{s=2}^{t-1}\sum_{j=1}^{t-2}d_n^2(j)\int D_v(y,y')\{g_{s-j}^{(l)}(y)^*g_{s-j}^{(l)}(y')^*-D_v(y,y')^*\}w_2(y)w_2(y')\\
=&E(\tilde{S}_{1nt}^2)+A_{4nt}+A_{5nt},\ \text{say.}
\end{aligned} \tag{A30}$$

By straightforward algebra, we obtain $E|A_{cnt}|^2 \leq \Delta t^2 M$ for $c = 4, 5$. It follows that

$$M^{-2}n^{-4}E\left|\sum_{t=j+1}^{n} A_{cnt}\right|^2 \leq M^{-2}n^{-4}\left\{\sum_{t=j+1}^{n}(E|A_{cnt}|^2)^{\frac{1}{2}}\right\}^2 = O(M^{-1}), \tag{A31}$$

$c = 4, 5$. Combining (A24)-(A31) and $E\tilde{S}_{1nt}^2 = 4\bar{D}_u\bar{D}_v\sum_{s=2}^{t-1}\sum_{j=1}^{s-1} d_n^2(j)$, we obtain

$$\begin{aligned} M^{-2}V\left\{n^{-2}\sum_{t=2}^{n} E(\tilde{S}_{1nt}^2|F_{t-1})\right\} &\leq \Delta M^{-2}\sum_{c=1}^{5} E\left|n^{-2}\sum_{t=2}^{n} A_{cnt}\right|^2 \\ &= O(M/n + M^{-1}) = o(1), \end{aligned}$$

thus ensuring (A19). This completes the proof. ∎

PROOF OF THEOREM A.4: Put $\hat{\delta}_u(x) = \hat{\varphi}_u^{(2m)}(0) - |\hat{\varphi}_u^{(m)}(x)|^2$ and $\hat{\delta}_v(y) = \hat{\varphi}_v^{(2l)}(0) - |\hat{\varphi}_v^{(l)}(y)|^2$. Then $\hat{C}_o^{(m,l)} = \int \hat{\delta}_u(x)w_1(x)dx \int \hat{\delta}_v(y)w_2(y)dy$. We shall show

$$\int \{\hat{\delta}_u(x) - \delta_u(x)\}w_1(x)dx = O_P(p/n^{\frac{1}{2}} + B_u(p)), \tag{A32}$$

$$\int \{\hat{\delta}_v(y) - \delta_y(y)\}w_2(y)dy = O_P(q/n^{\frac{1}{2}} + B_v(q)), \tag{A33}$$

where $\delta_u(x)$ and $\delta_v(y)$ are defined as $\hat{\delta}_u(x)$ and $\hat{\delta}_v(y)$ with $\hat{\varphi}_u$ and $\hat{\varphi}_v$ replaced by φ_u and φ_v. We show for (A32) only; the proof for (A33) is the same. Put $\tilde{\delta}_u(x) = \tilde{\varphi}_u^{(2m)}(0) - |\tilde{\varphi}_u^{(m)}(x)|^2$. Then $\hat{\delta}_u(x) = \hat{\delta}_u(x) - \tilde{\delta}_u(x) + \tilde{\delta}_u(x) - \delta_u(x)$. Noting $\hat{\varphi}_u(x) - \tilde{\varphi}_u(x) = n^{-1}\sum_{t=1}^{n}\Delta\hat{h}_t(x)$ where $\Delta\hat{h}_t(x) = e^{ix\hat{u}_t} - e^{ixu_t}$, we first decompose

$$\begin{aligned} &\hat{\delta}_u(x) - \tilde{\delta}_u(x) \\ =\ & n^{-1}\sum_{t=1}^{n}\Delta\hat{h}_t^{(2m)}(0) - \tilde{\varphi}_u^{(m)}(-x)\left\{n^{-1}\sum_{t=1}^{n}\Delta\hat{h}_t^{(m)}(x)\right\} \\ -\ & \tilde{\varphi}_u^{(m)}(x)\left\{n^{-1}\sum_{t=1}^{n}\Delta\hat{h}_t^{(m)}(-x)\right\} \\ -\ & \left\{n^{-1}\sum_{t=1}^{n}\Delta\hat{h}_t^{(m)}(x)\right\}\left\{n^{-1}\sum_{t=1}^{n}\Delta\hat{h}_t^{(m)}(-x)\right\}. \end{aligned}$$

For $0 \leq r \leq m$, we have

$$\begin{aligned}\int \left| n^{-1} \sum_{t=1}^{n} \Delta \hat{h}_t^{(r)}(x) \right|^2 w_1(x)dx &\leq n^{-1} \sum_{t=1}^{n} \int \left| \Delta \hat{h}_t^{(r)}(x) \right|^2 w_1(x)dx \\ &= O_P(p^2/n + B_u^2(p))\end{aligned}$$

by Lemma B.2. Also, by the Cauchy-Schwarz inequality and Lemma B.2, we have

$$\begin{aligned}& n^{-1} \sum_{t=1}^{n} \left| \Delta \hat{h}_t^{(2m)}(0) \right| \\ =\ & n^{-1} \sum_{t=1}^{n} \left| \hat{u}_t^{2m} - u_t^{2m} \right| \\ \leq\ & \left\{ n^{-1} \sum_{t=1}^{n} (\hat{u}_t^m - u_t^m)^2 \right\}^{\frac{1}{2}} \left\{ n^{-1} \sum_{t=1}^{n} (\hat{u}_t^m + u_t^m)^2 \right\}^{\frac{1}{2}} \\ \leq\ & 2 \left\{ n^{-1} \sum_{t=1}^{n} \left| \Delta \hat{h}_t^{(m)}(0) \right|^2 \right\}^{\frac{1}{2}} \left\{ n^{-1} \sum_{t=1}^{n} \left| \Delta \hat{h}_t^{(m)}(0) \right|^2 + n^{-1} \sum_{t=1}^{n} u_t^{2m} \right\}^{\frac{1}{2}} \\ =\ & O_P(p/n^{\frac{1}{2}} + B_u(p)).\end{aligned}$$

These, together with $E|\tilde{\varphi}_u^{(m)}(x)|^2 \leq \Delta,$ imply

$$\int \left| \hat{\delta}_u(x) - \tilde{\delta}_u(x) \right| w_1(x)dx = O_P(p/n^{\frac{1}{2}} + B_u(p)) \tag{A34}$$

by Markov's inequality. Similarly, noting

$$E|\tilde{\varphi}_u^{(r)}(x) - \varphi_u^{(r)}(x)|^2 = E|n^{-1} \sum_{t=1}^{n} h_t^{(r)}(x)|^2 \leq \Delta n^{-1}$$

for $0 \leq r \leq 2m$ given the independent and identical distribution assumption on (u_t) and Assumption A.4, we have

$$\int \left| \tilde{\delta}_u(x) - \delta_u(x) \right| w_1(x)dx = O_P(n^{-\frac{1}{2}}). \tag{A35}$$

Combining (A34)-(A35) yields (32). It follows from (A32)-(A33) that

$$\hat{C}_o^{(m,l)} - C_o^{(m,l)} = O_P \left\{ (p+q)/n^{\frac{1}{2}} + (B_u(p) + B_v(q)) \right\}. \tag{A36}$$

Hence, $\hat{C}_o^{(m,l)} - C_o^{(m,l)} = o_P(M^{-\frac{1}{2}})$ given the rate conditions on p, q. The proof for $\hat{D}_o^{(m,l)}$ is more tedious but similar; we omit it here. ▌

PROOF OF THEOREM 2: Because $\hat{C}_o^{(m,l)} \to^p C_o^{(m,l)}$ and $\hat{D}_o^{(m,l)} \to^p D_o^{(m,l)}$ by (A36) and the conditions on p and q,

$$\sum_{|j|=0}^{n-1} k^r(j/M) = M\int_{-\infty}^{\infty} k^r(z)dz\{1+o(1)\}$$

for $r = 2, 4$, and $M/n \to 0$, we have

$$\begin{aligned}&(M^{\frac{1}{2}}/n)C_n(m,l)\\ =& \left\{2D_o^{(m,l)}\int_{-\infty}^{\infty} k^4(z)dz\right\}^{-\frac{1}{2}}\left\{\sum_{|j|=0}^{n-1}\tilde{d}_n(j)\hat{I}_n^{(m,l)}(j)\right\}\{1+o(1)\}+o_P(1),\end{aligned}$$

where $\tilde{d}_n(j) = (1-|j|/n)k^2(j/p) \le d_n(j)$. Hence, it suffices to show

$$\sum_{|j|=0}^{n-1}\tilde{d}_n(j)\hat{I}_n^{(m,l)}(j) = \sum_{|j|=0}^{\infty} I^{(m,l)}(j) + o_P(1). \tag{A37}$$

Write

$$\begin{aligned}\sum_{|j|=0}^{n-1}\tilde{d}_n(j)\hat{I}_n^{(m,l)}(j) &= \sum_{|j|=0}^{n-1}\tilde{d}_n(j)I^{(m,l)}(j) + \sum_{|j|=0}^{n-1}\tilde{d}_n(j)\left\{\hat{I}_n^{(m,l)}(j) - I^{(m,l)}(j)\right\}\\ &= \sum_{|j|=0}^{n-1}\tilde{d}_n(j)I^{(m,l)}(j) + \hat{R}_n^{(m,l)}, \text{ say.}\end{aligned} \tag{A38}$$

Noting $I_j^{(m,l)}(x,y) = Eh_t^{(m)}(x)g_{t-j}^{(l)}(y)$, we have

$$|I_j^{(m,l)}(x,y)| \le 2\phi(j)^{\frac{1}{2}}\left\{E|h_t^{(m)}(x)|^2\right\}^{\frac{1}{2}}\left\{E|g_{t-j}^{(l)}(y)|^2\right\}^{\frac{1}{2}}$$

and

$$|I_j^{(m,l)}(x,y)| \le 2(2^{\frac{1}{2}}+1)\alpha(j)^{\frac{\eta}{2+2\eta}}\left\{E|h_t^{(m)}(u)|^2\right\}^{\frac{1}{2}}\left\{E|g_{t-j}^{(l)}(v)|^{2+2\eta}\right\}^{\frac{1}{2+2\eta}}$$

(cf. White, 1984, Corollary 6.16). It follows that $\sum_{j=1}^{\infty} I^{(m,l)}(j) < \infty$ given Assumption A.5. Therefore, for the first term of (A38),

$$\begin{aligned}\sum_{|j|=0}^{n-1} \tilde{d}_n(j) I^{(m,l)}(j) &= \sum_{|j|=0}^{\infty} I^{(m,l)}(j) + \sum_{|j|=0}^{n-1} \{\tilde{d}_n(j) - 1\} I^{(m,l)}(j) - \sum_{|j|=n}^{\infty} I^{(m,l)}(j) \\ &= \sum_{|j|=0}^{\infty} I^{(m,l)}(j) + o(1), \qquad \text{(A39)}\end{aligned}$$

where $\sum_{j=1}^{n-1}\{\tilde{d}_n(j) - 1\} I^{(m,l)}(j) \to 0$ by dominated convergence given $|\tilde{d}_n(j) - 1| \le 2$, $\tilde{d}_n(j) - 1 \to 0$ as $n \to \infty$ for any given j, and $\sum_{|j|=0}^{\infty} I^{(m,l)}(j) < \infty$; the latter also implies $\sum_{|j|=n}^{\infty} I^{(m,l)}(j) \to 0$.

We now show that $\hat{R}_n^{(m,l)}$ in (A38) vanishes in probability. Because

$$\begin{aligned}\hat{R}_n^{(m,l)} &= \sum_{j=1}^{n-1} \tilde{d}_n(j) \int \left| \hat{I}_j^{(m,l)}(x,y) - I_j^{(m,l)}(x,y) \right|^2 w_1(x) w_2(y) dx dy \\ &+ 2 \sum_{j=1}^{n-1} \tilde{d}_n(j) \mathrm{Re} \int \left[\{\hat{I}_j^{(m,l)}(x,y) - I_j^{(m,l)}(x,y)\} I_j^{(m,l)}(x,y)^* \right] \\ &\times w_1(x) w_2(y) dx dy,\end{aligned}$$

it suffices to show that the first term vanishes in probability; that the second term vanishes then follows by the Cauchy-Schwarz inequality. Write

$$\begin{aligned}\left| \hat{I}_j^{(m,l)}(x,y) - I_j^{(m,l)}(x,y) \right|^2 &\le 2 \left| \hat{I}_j^{(m,l)}(x,y) - \tilde{I}_j^{(m,l)}(x,y) \right|^2 \\ &+ 2 \left| \tilde{I}_j^{(m,l)}(x,y) - I_j^{(m,l)}(x,y) \right|^2,\end{aligned}$$

where $\tilde{I}_j^{(m,l)}(x,y)$ is defined as $\hat{I}_j^{(m,l)}(x,y)$ with $(\hat{u}_t, \hat{v}_t)$ replaced by (u_t, v_t). For the first term, we have

$$\begin{aligned}&\sum_{|j|=0}^{n-1} \tilde{d}_n(j) \int \left| \hat{I}_j^{(m,l)}(x,y) - \tilde{I}_j^{(m,l)}(x,y) \right|^2 w_1(x) w_2(y) dx dy \\ =& O_P\left\{ M(p^2/n + B_u^2(p))(q^2/n + B_v^2(q)) \right\} \\ =& o_P(1) \qquad \text{(A40)}\end{aligned}$$

by (A11) and $p + q = o(n^{\frac{1}{2}}/M^{\frac{1}{4}}), B_u(p) + B_v(q) = o(M^{-\frac{1}{4}})$. (Note that (A11) holds here as it does not assume independence between (u_t) and

(v_t).) For the second term, using (A3), we have

$$\begin{aligned}
& E\left|\tilde{I}_j^{(m,l)}(x,y) - I_j^{(m,l)}(x,y)\right|^2 \\
\leq\ & 2E\left|\tilde{\gamma}_j^{(m,l)}(x,y) - I_j^{(m,l)}(x,y)\right|^2 \\
& +2\left\{E\left|(n-j)^{-1}\sum_{t=j+1}^{n} h_t^{(m)}(x)\right|^4\right\}^{\frac{1}{2}}\left\{E\left|(n-j)^{-1}\sum_{t=j+1}^{n} g_{t-j}^{(l)}(y)\right|^4\right\}^{\frac{1}{2}} \\
\leq\ & \Delta(n-j)^{-1}
\end{aligned}$$

where $E|\tilde{\gamma}_j^{(m,l)}(x,y) - I_j^{(m,l)}(x,y)|^2 \leq \Delta(n-j)^{-1}$ by Lemma B.5 in Appendix B, $E|(n-j)^{-1}\sum_{t=j+1}^{n} h_t^{(m)}(x)|^4 \leq \Delta(n-j)^{-2}$ and $E|(n-j)^{-1}\sum_{t=j+1}^{n} g_{t-j}^{(l)}(y)|^4 \leq \Delta(n-j)^{-2}$ given Assumption A.4 and that (u_t) and (v_t) are each an independent and identically distributed sequence. It follows that

$$\begin{aligned}
\sum_{j=1}^{n-1} \tilde{d}_n(j)E\left|\tilde{\gamma}_j^{(m,l)}(x,y) - I_j^{(m,l)}(x,y)\right|^2 \leq & \Delta(M/n)\left\{M^{-1}\sum_{j=1}^{n-1} d_n(j)\right\} \\
= & O(M/n). \qquad \text{(A41)}
\end{aligned}$$

Combining (A40)-(A41) yields $\hat{R}_n^{(m,l)} \to^p 0$. This, together with (A38)-(A39), implies (A37). The proof is thus completed. ▌

APPENDIX B

LEMMA B.1. *Suppose that Assumption A.3 holds. Let $M \to \infty, M/n \to 0$. Then $M^{-1}\sum_{|j|=0}^{n-1} k^2(j/M)(1-|j/n|)^{-1} = O(1)$.*

Proof. Put $q = M^{1-\frac{1}{2b}} n^{\frac{1}{2b}}$. Then $q/M \to \infty, q/n \to 0$. We have

$$\begin{aligned}
& M^{-1} \sum_{|j|=0}^{n-1} k^2(j/M)(1-|j/n|)^{-1} \\
= & \ M^{-1} \sum_{|j|=0}^{q} k^2(j/M)(1-|j/n|)^{-1} + M^{-1} \sum_{|j|=q+1}^{n-1} k^2(j/M)(1-|j/n|)^{-1} \\
\leq & \ (1-q/n)^{-1} M^{-1} \sum_{|j|=0}^{q} k^2(j/M) + C^2 n M^{2b-1} q^{-2b} \sum_{|j|=q+1}^{n-1} (j/q)^{-2b} \\
= & \ O(1)
\end{aligned}$$

for all n sufficiently large, where $M^{-1}\sum_{|j|=0}^{q} k^2(j/M) \to \int_{-\infty}^{\infty} k^2(z)dz$ and $\sum_{|j|=q+1}^{n-1}(j/q)^{-2b} \leq \Delta < \infty$ given Assumption A.3. ▌

Lemma B.2. *Suppose that Assumptions A.1-A.2 and A.4 hold. Let $B_u(p)$ and $B_v(q)$ be defined as in Theorem 1. Then for $0 \leq r \leq m$,*

$$\begin{aligned}
n^{-1} \sum_{t=1}^{n} \int \left|(i\hat{u}_t)^r e^{ix\hat{u}_t} - (iu_t)^r e^{ixu_t}\right|^2 w_1(x)dx &= O_P\left\{p^2/n + B_u^2(p)\right\}, \\
n^{-1} \sum_{t=1}^{n} \int \left|(i\hat{v}_t)^r e^{iy\hat{v}_t} - (iv_t)^r e^{iyv_t}\right|^2 w_2(y)dx &= O_P\left\{q^2/n + B_v^2(q)\right\}.
\end{aligned}$$

Proof. We show for $(\hat{u}_t)$ only; the proof for $(\hat{v}_t)$ is the same. Because $(i\hat{u}_t)^r e^{ix\hat{u}_t} - (iu_t)^r e^{ixu_t} = \{(i\hat{u}_t)^r - (iu_t)^r\}e^{ix\hat{u}_t} + (iu_t)^r(e^{ix\hat{u}_t} - e^{ixu_t})$, and $|e^{ix\hat{u}_t} - e^{ixu_t}| \leq |x\hat{u}_t - xu_t|$ (cf. Chung, 1974, p.154), we have

$$\left|(i\hat{u}_t)^r e^{ix\hat{u}_t} - (iu_t)^r e^{ixu_t}\right|^2 \leq 2\left(\hat{u}_t^r - u_t^r\right)^2 + 2x^2 u_t^{2r}(\hat{u}_t - u_t)^2. \quad \text{(A42)}$$

By the binomial formula, we obtain

$$\begin{aligned}
n^{-1}\sum_{t=1}^{n}\left(\hat{u}_t^r-u_t^r\right)^2 =& n^{-1}\sum_{t=1}^{n}\{(\hat{u}_t-u_t+u_t)^r-u_t^r\}^2\\
=& n^{-1}\sum_{t=1}^{n}\left\{\sum_{c=1}^{r}\binom{r}{c}(\hat{u}_t-u_t)^c u_t^{r-c}\right\}^2\\
\le& \Delta(r)\sum_{c=1}^{r}\binom{r}{c}^2 n^{-1}\sum_{t=1}^{n}(\hat{u}_t-u_t)^{2c}u_t^{2(r-c)}\\
\le& \Delta(r)\sum_{c=1}^{r}\binom{r}{c}^2\left\{n^{-1}\sum_{t=1}^{n}(\hat{u}_t-u_t)^{4c}\right\}^{\frac{1}{2}}\left\{n^{-1}\sum_{t=1}^{n}u_t^{4(r-c)}\right\}^{\frac{1}{2}},
\end{aligned} \tag{A43}$$

where $\Delta(r)$ is a finite constant depending on r only. Here, $n^{-1}\sum_{t=1}^{n}u_t^{4(r-c)}=O_P(1)$ by Markov's inequality and Assumption A.4.

We now consider $n^{-1}\sum_{t=1}^{n}(\hat{u}_t-u_t)^{4c}$. Put $b_t(p)=\sum_{j=p+1}^{\infty}\alpha_j X_{t-j}$. Because $X_t=X_t'(p)\hat{\alpha}(p)+\hat{u}_t=X_t'(p)\alpha(p)+b_t(p)+u_t$, we have

$$\begin{aligned}
n^{-1}\sum_{t=1}^{n}(\hat{u}_t-u_t)^{4c} =& n^{-1}\sum_{t=1}^{n}\{X_t'(p)(\alpha(p)-\hat{\alpha}(p))+b_t(p)\}^{4c}\\
\le& \Delta(c)\left\{\|\hat{\alpha}(p)-\alpha(p)\|^{4c}n^{-1}\sum_{t=1}^{n}\|X_t(p)\|^{4c}+n^{-1}\sum_{t=1}^{n}b_t^{4c}(p)\right\}\\
=& O_P\left\{p^2/n+p\left(\sum_{j=p+1}^{\infty}|\alpha_j|\right)^2\right\}^{2c}\\
=& O_P\left\{p^2/n+B_u^2(p)\right\}^{2c}
\end{aligned} \tag{A44}$$

by Lemma B.3 below, $n^{-1}\sum_{t=1}^{n}||X_t(p)||^{4c}=O_P(p^{2c})$ and $n^{-1}\sum_{t=1}^{n}b_t^{4c}(p)=O_P\{\sum_{j=p+1}^{\infty}|\alpha_j|\}^{4c}$. Because r is a fixed integer and $p^2/n+B_u^2(p)\to 0$, the order of $n^{-1}\sum_{t=1}^{n}\left(\hat{u}_t^r-u_t^r\right)^2$ is determined by the slowest vanishing term ($c=1$) in (A43). It follows that we have

$$n^{-1}\sum_{t=1}^{n}(\hat{u}_t^r-u_t^r)^2=O_P\left\{p^2/n+B_u^2(p)\right\}. \tag{A45}$$

Also, we have by the Cauchy-Schwarz inequality and (A44) with $c = 1$ that

$$
\begin{aligned}
n^{-1}\sum_{t=1}^{n} u_t^{2r}(\hat{u}_t - u_t)^2 \leq & \left\{n^{-1}\sum_{t=1}^{n}(\hat{u}_t - u_t)^4\right\}^{\frac{1}{2}} \left\{n^{-1}\sum_{t=1}^{n} u_t^{4r}\right\}^{\frac{1}{2}} \\
= & O_P\left\{p^2/n + B_u^2(p)\right\}. \qquad \text{(A46)}
\end{aligned}
$$

Combining (A42), (A45)-(A46) and Assumption A.2 yields the desired result. ▮

LEMMA B.3. *Suppose that Assumptions A.1 and A.4 hold. Let $\hat{\alpha}(p)$ and $\hat{\beta}(q)$ be the ordinary least squares estimators for the truncated autoregressions of X_t and Y_t. Then*

$$
\begin{aligned}
\|\hat{\alpha}(p) - \alpha(p)\|^2 &= O_P\left\{p/n + \left(\sum_{j=p+q}^{\infty}|\alpha_j|\right)^2\right\}, \\
\left\|\hat{\beta}(q) - \beta(q)\right\|^2 &= O_P\left\{q/n + \left(\sum_{j=p+q}^{\infty}|\beta_j|\right)^2\right\}.
\end{aligned}
$$

Proof. We show for $\hat{\alpha}(p)$ only. Put

$$
b_t(p) = \sum_{j=p+1}^{\infty}\alpha_j X_{t-j}, \hat{R}(p) = (n-p)^{-1}\sum_{t=p+1}^{n} X_t(p)X_t'(p)
$$

and $R(p) = EX_t(p)X_t'(p)$. Then $\hat{\alpha}(p) = \hat{R}(p)^{-1}(n-p)^{-1}\sum_{t=p+1}^{n} X_t(p)X_t$. Hence,

$$
\begin{aligned}
\alpha(p) - \hat{\alpha}(p) = & \hat{R}(p)^{-1}\left\{(n-p)^{-1}\sum_{t=p+1}^{n} X_t(p)(u_t + b_t(p))\right\} \\
= & \hat{R}(p)^{-\frac{1}{2}}\left\{\hat{R}(p)^{-\frac{1}{2}}\hat{A}(p) + \hat{R}(p)^{-\frac{1}{2}}\hat{B}(p)\right\} \\
= & R(p)^{-\frac{1}{2}}\left\{\hat{R}(p)^{-\frac{1}{2}}\hat{A}(p) + \hat{R}(p)^{-\frac{1}{2}}\hat{B}(p)\right\} \\
& + \left\{\hat{R}(p)^{-\frac{1}{2}} - R(p)^{-\frac{1}{2}}\right\}\left\{\hat{R}(p)^{-\frac{1}{2}}\hat{A}(p) + \hat{R}(p)^{-\frac{1}{2}}\hat{B}(p)\right\} \\
= & T_{1n} + T_{2n}, \text{ say}, \qquad \text{(A47)}
\end{aligned}
$$

where $\hat{A}(p) = (n-p)^{-1} \sum_{t=p+1}^{n} X_t(p)u_t$ and

$$\hat{B}(p) = (n-p)^{-1} \sum_{t=p+1}^{n} X_t(p)b_t(p).$$

For the first term,

$$\begin{aligned} \|T_{1n}\|^2 \leq & 2\lambda_{\min}^{-1}\{R(p)\} \left\{ \left\| \hat{R}(p)^{-\frac{1}{2}} \hat{A}(p) \right\|^2 + \left\| \hat{R}(p)^{-\frac{1}{2}} \hat{B}(p) \right\|^2 \right\} \\ = & O_P \left\{ p/n + \left(\sum_{j=p+1}^{\infty} |\alpha_j| \right)^2 \right\}, \end{aligned} \tag{A48}$$

where

$$\begin{aligned} \left\| \hat{R}(p)^{-\frac{1}{2}} \hat{A}(p) \right\|^2 &= \left\| \hat{A}'(p)R(p)^{-1}\hat{A}(p) + \hat{A}'(p) \left\{ \hat{R}(p)^{-1} - R(p)^{-1} \right\} \hat{A}(p) \right\|^2 \\ &\leq 2\lambda_{\min}^{-1} \{R(p)\} \left\| \hat{A}(p) \right\|^2 + 2 \left\| \hat{R}(p)^{-1} - R(p)^{-1} \right\| \left\| \hat{A}(p) \right\|^2 \\ &= O_P(p/n) \end{aligned}$$

given $\lambda_{\min} \{R(p)\} \geq \Delta > 0$ (cf. Berk, 1974, p. 491), $\left\| \hat{A}(p) \right\|^2 = O_P(p/n)$ by Chebyshev's inequality and $\left\| \hat{R}(p)^{-1} - R(p)^{-1} \right\| = O_P(p/n^{\frac{1}{2}}) \to 0$, as can be obtained from the proof of Lemma 3 of Berk (1974, p.493). Also, by projection, we have

$$\begin{aligned} \left\| \hat{R}(p)^{-\frac{1}{2}} \hat{B}(p) \right\|^2 &= \hat{B}(p)'\hat{R}(p)^{-1}\hat{B}(p) \leq (n-p)^{-1} \sum_{t=p+1}^{n} b_t^2(p) \\ &= O_P \left(\sum_{j=p+1}^{\infty} |\alpha_j| \right)^2, \end{aligned}$$

where the last equality follows by Markov's inequality and $Eb_t^2(p) \leq \Delta(\sum_{j=p+1}^{\infty} |\alpha_j|)^2$. Next, for the second term,

$$\begin{aligned} \|T_{2n}\|^2 \leq & \left\| \hat{R}(p)^{-\frac{1}{2}} - R(p)^{-\frac{1}{2}} \right\|^2 \left\| \hat{R}(p)^{-\frac{1}{2}} \hat{A}(p) + \hat{R}(p)^{-\frac{1}{2}} \hat{B}(p) \right\|^2 \\ = & o_P(\|T_{1n}\|^2), \end{aligned} \tag{A49}$$

where $||\hat{R}(p)^{-\frac{1}{2}} - R(p)^{-\frac{1}{2}}||^2 = o_P(1)$ given $||\hat{R}(p)^{-1} - R(p)^{-1}||^2 = o_P(1)$ and $\lambda_{\min}\{R(p)\} \geq \Delta > 0$. Combining (A47)-(A48) yields the desired result. ▌

LEMMA B.4. *Suppose that Assumptions A.1-A.4 hold. Let $M \to \infty, M/n \to 0$. If (u_t) and (v_t) are mutually independent, then*

$$V(\bar{S}_n) = 2D_o^{(m,l)} \sum_{|j|=0}^{n-2} k^4(j/M) + O(M/n),$$

where $D_o^{(m,l)}$ is as in Theorem A.3 .

Proof. Because $E(S_{nt}|F_{t-1}) = 0$ and $E(S_{1nt}S_{2nt}|F_{t-1}) = 0$, we have $E\bar{S}_n^2 = n^{-2}\sum_{t=2}^{n-1} ES_{1nt}^2 + ES_{2nt}^2$,where S_{1nt} and S_{2nt} are as in (A17). We first consider ES_{1nt}^2. Rewrite $S_{1nt} = 2\sum_{s=2}^{t-1} Re(H_{t,s}^{(m)} \sum_{j=1}^{s-1} d_n(j)G_{t-j,s-j}^{(l)})$. Conditional on u_t and $(v_s)_{s=1}^n$, S_{1nt} is a sum of an independent and identically distributed sequence. We thus have

$$\begin{aligned} ES_{1nt}^2 &= 4\sum_{s=2}^{t-1} E\left\{\sum_{j=1}^{s-1} d_n(j) Re(H_{t,s}^{(m)} G_{t-j,s-j}^{(l)})\right\}^2 \\ &= 4\sum_{s=2}^{t-1}\sum_{j=1}^{s-1} d_n^2(j) E\left\{Re(H_{t,s}^{(m)} G_{t-j,s-j}^{(l)})\right\}^2 \\ &= 4D_o^{(m,l)} \sum_{s=2}^{t-1}\sum_{j=1}^{s-1} d_n^2(j), \end{aligned}$$

where the last two equalities follow because for $t > s > j_1 > j_2$

$$E\left\{G_{t-j_1,s-j_1}^{(l)} G_{t-j_2,s-j_2}^{(l)})\right\} = 0$$

and for $t > s > j$,

$$
\begin{aligned}
& 4E\left\{Re(H_{t,s}^{(m)}G_{t-j,s-j}^{(l)})\right\}^2 \\
= \;& E\left\{H_{t,s}^{(m)}G_{t-j,s-j}^{(l)} + (H_{t,s}^{(m)})^*(G_{t-j,s-j}^{(l)})^*\right\}^2 \\
= \;& \Big\{E(H_{t,s}^{(m)})^2E(G_{t-j,s-j}^{(l)})^2 \\
+ \;& E(H_{t,s}^{(m)*})^2E(G_{t-j,s-j}^{(l)*})^2 + 2E|H_{t,s}^{(m)}|^2E|G_{t-j,s-j}^{(l)}|^2\Big\} \\
= \;& 2\int |Eh_1^{(m)}(x)h_1^{(m)}(x')|^2 w_1(x)w_2(x')dxdx' \\
\times \;& \int |Eg_1^{(l)}(y)g_1^{(l)}(y')|^2 w_1(y)w_2(y')dydy' \\
+ \;& 2\int |Eh_1^{(m)}(x)h_1^{(m)}(-x')|^2 w_1(x)w_2(x')dxdx' \\
\times \;& \int |Eg_1^{(l)}(y)g_1^{(l)}(-y')|^2 w_1(y)w_2(y')dydy' \\
= \;& 4D_o^{(m,l)},
\end{aligned}
$$

where $Eh_1^{(m)}(x)h_1^{(m)}(x') = \varphi_u^{(2m)}(x+x') - \varphi_u^{(m)}(x)\varphi_u^{(m)}(x')$ and

$$Eg_1^{(l)}(y)g_1^{(l)}(y') = \varphi_u^{(2l)}(x+x') - \varphi_u^{(l)}(x)\varphi_u^{(l)}(x').$$

Note that the last equality follows by symmetry of $w_1(\cdot)$ and $w_2(\cdot)$. Similarly, $ES_{2nt}^2 = 4D_o^{(m,l)}\sum_{s=2}^{t-1}\sum_{j=0}^{s-1} d_n^2(j)$. Hence, we have

$$
\begin{aligned}
E\bar{S}_n^2 &= 4D_o^{(m,l)}n^{-2}\sum_{t=2}^{n-1}\sum_{s=1}^{t-1}\sum_{|j|=0}^{s-1} d_n^2(j) \\
&= 2D_o^{(m,l)}\sum_{|j|=0}^{n-2}\{1-(n-|j|)^{-1}\}k^4(j/M) \\
&= 2D_o^{(m,l)}\sum_{|j|=0}^{n-2}k^4(j/M) + O(M/n),
\end{aligned}
$$

where $\sum_{|j|=0}^{n-1}(n-|j|)^{-1}k^4(j/M) \le n^{-1}\sum_{|j|=0}^{n-1} d_n(j) = O(M/n)$ by Lemma B.1. ▮

Lemma B.5. *Suppose that Assumptions A.1-A.2 and A.4-A.5 hold. Let* $\tilde{\gamma}_j^{(m,l)}$ *be defined as in Theorem A.2. Then* $E\left|\tilde{\gamma}_j^{(m,l)}(x,y) - I_j^{(m,l)}(x,y)\right|^2 \le \Delta(n-j)^{-1}$.

Proof. Put $C_{ut}(x) = \cos(xu_t) - E\cos(xu_t)$, $S_{ut}(x) = \sin(xu_t) - E\sin(xu_t)$, $C_{vt}(y) = \cos(yv_t) - E\cos(yv_t)$, and $S_{vt}(y) = \sin(yv_t) - E\sin(yv_t)$. Define

$$\begin{aligned} W_{1tj}^{(m,l)}(u,v) &= C_{ut}^{(m)}(x)C_{vt-j}^{(l)}(y) - E\{C_{ut}^{(m)}(x)C_{vt-j}^{(l)}(y)\}, \\ W_{2tj}^{(m,l)}(u,v) &= S_{ut}^{(m)}(x)S_{vt-j}^{(l)}(y) - E\{S_{ut}^{(m)}(x)S_{vt-j}^{(l)}(y)\}, \\ W_{3tj}^{(m,l)}(u,v) &= C_{ut}^{(m)}(x)S_{vt-j}^{(l)}(y) - E\{C_{ut}^{(m)}(x)S_{vt-j}^{(l)}(y)\}, \\ W_{4tj}^{(m,l)}(u,v) &= S_{ut}^{(m)}(x)C_{vt-j}^{(l)}(y) - E\{S_{ut}^{(m)}(x)C_{vt-j}^{(l)}(y)\} \end{aligned}$$

for $j \geq 0$. Noting that $I_j^{(m,l)}(x,y) = E\tilde{\gamma}_j^{(m,l)}(x,y)$, we have

$$\begin{aligned} &(n-j)\left\{\tilde{\gamma}_j^{(m,l)}(x,y) - I_j^{(m,l)}(x,y)\right\} \\ &= \sum_{t=j+1}^{n}\left\{W_{1tj}^{(m,l)}(x,y) - W_{2tj}^{(m,l)}(x,y)\right\} \\ &+ i\sum_{t=j+1}^{n}\left\{W_{3tj}^{(m,l)}(x,y) + W_{4tj}^{(m,l)}(x,y)\right\}. \end{aligned}$$

It follows that

$$(n-j)^2 E\left|\tilde{\gamma}_j^{(m,l)}(x,y) - I_j^{(m,l)}(x,y)\right|^2 \leq 2\sum_{c=1}^{4} E\left\{\sum_{t=j+1}^{n} W_{ctj}^{(m,l)}(x,y)\right\}^2.$$

We now show $E\{\sum_{t=j+1}^{n} W_{ctj}^{(m,l)}(x,y)\}^2 \leq \Delta(n-j)$ for $c = 1,2,3,4$. We first consider $c = 1$ by applying White's (1984) Lemma 6.19. Given (x,y), $W_{1tj}^{(m,l)}(x,y)$ is a measurable function of (u_t, v_t) and (u_{t-j}, v_{t-j}) for each t and j, with $EW_{1tj}(x,y) = 0$ for all t, j. Furthermore, by Jensen and C_r inequalities and Assumption A.5, we have

$$\begin{aligned} &E\left|W_{1tj}^{(m,l)}(x,y)\right|^{2+2\eta} \\ &\leq \Delta(\eta)\left\{E\left|C_{ut}^{(m)}(x)C_{vt-j}^{(l)}(y)\right|^{2+2\eta} + \left|EC_{ut}^{(m)}(x)C_{vt-j}^{(l)}(y)\right|^{2+2\eta}\right\} \\ &\leq 2\Delta(\eta)E\left|C_{ut}^{(m)}(x)C_{vt-j}^{(l)}(y)\right|^{2+2\eta} \\ &\leq 2\Delta(\eta)\left(E\left|u_t\right|^{4m(1+\eta)} E|v_t|^{4l(1+\eta)}\right)^{\frac{1}{2}} \\ &\leq \Delta \end{aligned}$$

In addition, (u_t, v_t) is a stationary mixing process with either $\phi(j)$ of size 2 or $\alpha(j)$ of size $\eta/(2+2\eta)$, $\eta > 0$. It follows that all the conditions of Lemma 6.19 of White (1984) are satisfied. Therefore, $E\{\sum_{t=j+1}^{n} W_{1tj}^{(m,l)}(x,y)\}^2 \leq \Delta(n-j)$. Similar results hold for $c = 2, 3, 4$. Hence, $E|\tilde{\gamma}_j^{(m,l)}(x,y) - I_j^{(m,l)}(x,y)|^2 \leq \Delta(n-j)^{-1}$. This completes the proof. ▌

REFERENCES

Berk, K. N., 1974, Consistent autoregressive spectral estimates. *Ann. Statist.* **2**, 489-502.

Box, G. and Pierce, D., 1971, Distribution of residual autocorrelations in autoregressive integrated moving average time series models. *J. Am. Statist. Assoc.* **65**, 1509-1527.

Brown, B. M., 1971, Martingale central limit theorems. *Ann. Math. Statist.* 42, 59-66.

Chung, K.L., 1974, *A Course in Probability Theory, 2nd Ed.* New York: Academic Press.

Csorgo, S., 1985, Testing for independence by the empirical characteristic function. *J. Multivar. Anal.* **16**, 290-299.

De Sliva, B. M. and C. R. Griffiths, 1980, A test of independence for bivariate symmetric stable distributions. *Austral. J. Statist.* **22**, 172-177.

Engle, R. F. and R. Susmel, 1993, Common volatility in international equity markets. *J. Busi. Econ. Statist.* **11**, 167-176.

Epps, T. W., 1987, Testing that a stationary time series is Gaussian. *Ann. Statist.* **15**, 177-203.

Epps, T. W., 1988, Testing that a Gaussian time series is stationary. *Ann. Statist.* **16**, 1667-1683.

Feller, W., 1957, *An Introduction to Probability Theory and Its Applications, 2nd Ed.* New York: John Wiley.

Feuerverger, A., 1987, On some ecf procedures for testing independence. In: *Time Series and Econometric Modelling*, I. B. MacNeill & G. J. Umphrey (eds.), pp.189-206. Boston: D. Reidel.

Granger, C. W. J. and A. P. Anderson, 1978, *An Introduction to Bilinear Time Series Models.* Gottingen: Bandenhoek and Ruprecht.

Haugh, L. D., 1976, Checking the independence of two covariance-stationary time series: a univariate residual cross correlation approach. *J. Am. Statist. Assoc.* **71**, 378-385.

Hong, Y., 1996, Testing for independence between two covariance stationary time series. *Biometrika* **83**, 615-625.

Hong, Y., 1999, Hypothesis testing in time series via the empirical chracteristic function: a generalized spectral density approach. *J. Am. Statist. Assoc.* **94**, 1201-1220.

Hong, Y., 2001, A test for volatility spillover with application to foreign exchange rates. Forthcoming in *J. Econometric.*

Kock, P. D. and S.-S. Yang, 1986, A method for testing the independence of two time series that accounts for a potential pattern in the cross-correlation function. *J. Am. Statist. Assoc.* **81**, 533-544.

Ljung, G. and G. Box, 1978, A measure of lack of fit in time series models. *Biometrika* 65, 297-303.

Mcleod, A. I. and W. K. LI, 1983, Diagnostic checking ARMA time series models using squared residual autocorrelations. *J. Time Ser. Anal.* **4**, 269-273.

Pierce, A., 1977, Lack of dependence among economic variables. *J. Am. Statist. Assoc.* **72**, 11-22.

Pierce, A. and L. D. Haugh, 1977, Causality in temporal systems. *J. Econometric.* **5**, 265-293.

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

Satterthwaite, F. E., 1941, Synthesis of variance. *Psychometrica* **6**, 309-316.

Satterthwaite, F. E., 1946, An approximate distribution of estimates of variance components. *Biometrics Bulletin* **2**, 110-114.

White, H., 1984, *Asymptotic Theory for Econometricians.* New York: Academic Press.