

# A Generalized Portmanteau Goodness-of-fit Test for Time Series Models

Willa W. Chen and Rohit S. Deo

*New York University*

## **Abstract**

We present a goodness of fit test for time series models based on the discrete spectral average estimator. Unlike current tests of goodness of fit, the asymptotic distribution of our test statistic allows the null hypothesis to be either a short or long range dependence model. Our test is in the frequency domain, is easy to compute and does not require the calculation of residuals from the fitted model. This is especially advantageous when the fitted model is not a finite order autoregressive model. The test statistic is a frequency domain analogue of the test by Hong (1996) which is a generalization of the Box-Pierce (1970) test statistic. A simulation study shows that our test has power comparable to that of Hong's test and superior to that of another frequency domain test by Milhoj (1981).

*Keywords:* Portmanteau test, long memory, goodness-of-fit.

## **1 Introduction**

Most conventional goodness-of-fit tests for time series models are based on the autocorrelations of residuals from the fitted model. Examples of such tests include the portmanteau statistic of Box and Pierce (1970) and its generalization, based on arbitrary kernel functions, by Hong (1996). The Box-Pierce statistic is obtained as a particular case of the Hong statistic by using the truncated uniform kernel. Simulations by Hong show that his statistic computed using kernels

other than the truncated uniform kernel gives better power than the Box-Pierce statistic against autoregressive (AR) processes and fractionally integrated processes.

Box and Pierce (1970) derived the null distribution of their test for autoregressive moving average (ARMA) models and Hong derived the his null distribution only for finite order autoregressive models. Both these results require assumptions that rule out long memory processes which have hyperbolically decaying correlation functions and spectral densities unbounded at the origin. Furthermore, both tests requires the computation of residuals from the fitted model, which can be quite tedious when the model does not have a finite order autoregressive representation. Also, in such cases, the residuals are not uniquely defined.

A test statistic which circumvents the computation of residuals from the fitted model was proposed by Milhoj (1981). To test the hypothesis that the observations  $x_t$ ,  $t = 1, \dots, n$ , are from a process with spectral density  $f(\lambda)$ , he suggested the test statistic,  $M_n^d = \left\{ \sum_{j=1}^{n-1} V_j \right\}^{-2} \sum_{j=1}^{n-1} V_j^2$  where  $V_j = I(\lambda_j)/f(\lambda_j)$ ,  $I(\lambda) = (2\pi n)^{-1} \left| \sum_{t=1}^n x_t e^{-i\lambda t} \right|^2$  is the periodogram of the observations and  $\lambda_j = 2\pi j/n$  is the  $j$ th Fourier frequency. Though Milhoj's test statistic is easily computed, his theoretical results are restricted to short memory time series models with bounded spectral densities. Assuming Gaussianity, Beran (1992) extended Milhoj's results to long memory time series models which have unbounded spectral densities at the origin. Examples of long memory processes are the autoregressive fractionally integrated moving average (ARFIMA) process (see Hosking (1981)). Beran stated that the null distribution of  $M_n^d$  in the presence of long memory is the same as that derived by Milhoj (1981) in the case of short memory. Beran obtained his results by claiming that  $M_n^d$  is asymptotically equivalent to its integral version  $M_n = \left\{ \int_0^{2\pi} V(\lambda) d\lambda \right\}^{-2} \int_0^{2\pi} V^2(\lambda) d\lambda$  where  $V(\lambda) = I(\lambda)/f(\lambda)$ .

However, Deo and Chen (2000) showed that even in the case of Gaussian white noise,  $M_n^d$  and  $M_n$  do not have the same asymptotic distribution and that the variance of  $M_n$  is two thirds that of  $M_n^d$ . Thus, the asymptotic distribution of  $M_n^d$  in the long memory case is still an open question.

In this paper, we introduce a test statistic which is a frequency domain analogue of Hong's statistic. We derive the asymptotic null distribution for both short memory models and long memory models. Since our test does not require the calculation of residuals, it can be easily

applied to long memory processes such as the ARFIMA models which do not possess finite order AR representations. Our test delivers uniformly better power than the periodogram-based test  $M_n^d$  of Milhoj.

In the next section, we define our test statistic and provide the theoretical results on its asymptotic null distribution for short and long memory models. The power properties of our test are studied in section 3 through simulations. The proofs are relegated to the Appendix at the end.

## 2 The test statistic

To motivate our test statistic, it is instructive to consider Hong's statistic to test the null hypothesis that the observations,  $x_t$ ,  $t = 1, 2, \dots, n$ , are from an AR( $p$ ) process,  $x_t = \alpha_0 + \alpha_1 x_{t-1} + \dots + \alpha_p x_{t-p} + \varepsilon_t$ , where  $\varepsilon_t$  are zero mean white noise. Let  $e_t$  be the residuals from the fitted model,  $e_t = x_t - \hat{\alpha}_0 - \hat{\alpha}_1 x_{t-1} - \dots - \hat{\alpha}_p x_{t-p}$ , where  $\hat{\alpha}_0, \hat{\alpha}_1, \dots, \hat{\alpha}_p$  are the estimates of the parameters  $\alpha_0, \alpha_1, \dots, \alpha_p$ . Hong's (1996) test statistic is

$$H_n = \sum_{j=1}^{n-1} k^2(j/p_n) \hat{\rho}_{e,j}^2,$$

where  $k(\cdot)$  is a suitable kernel function,  $\hat{\rho}_{e,j} = \hat{\gamma}_{e,j} / \hat{\gamma}_{e,0}$  are the sample autocorrelations of the residuals and  $\hat{\gamma}_{e,j}$  are their sample autocovariances,

$$\hat{\gamma}_{e,j} = \frac{1}{n} \sum_{t=|j|+1}^n (e_t - \bar{e})(e_{t-|j|} - \bar{e}), \quad j = 0, \pm 1, \dots, \pm(n-1).$$

By Parseval's identity,  $H_n$  can be written as

$$H_n = \frac{1}{2} \left( \sum_{j=-(n-1)}^{n-1} k^2(j/p_n) \hat{\rho}_{e,j}^2 - 1 \right) = \frac{1}{2} \left\{ \left( \int_0^{2\pi} \hat{f}_e(\lambda) d\lambda \right)^{-2} \left( 2\pi \int_0^{2\pi} \hat{f}_e^2(\lambda) d\lambda \right) - 1 \right\} \quad (1)$$

where

$$\hat{f}_e(\lambda) = \frac{1}{2\pi} \sum_{|j|<n} k(j/p_n) \hat{\gamma}_{e,j} e^{-i\lambda j}. \quad (2)$$

The kernel function  $k$  here is also called the lag window, and  $\hat{f}_e(\lambda)$  the lag-weights spectral density estimator. Let  $I_{n,e}$  be the mean corrected periodogram of the residuals given by

$$I_{n,e}(\lambda) = \frac{1}{2\pi n} \left| \sum_{t=1}^n (e_t - \bar{e}) e^{-i\lambda t} \right|^2.$$

Using the relation

$$\widehat{\gamma}_{e,j} = \int_0^{2\pi} I_{n,e}(\omega) e^{i\omega j} d\omega,$$

we have an equivalent form of  $\widehat{f}_e(\lambda)$  in the frequency domain,

$$\widehat{f}_e(\lambda) = \int_0^{2\pi} W(\lambda - \omega) I_{n,e}(\omega) d\omega, \quad (3)$$

where  $W$ , the spectral window corresponding to the lag window  $k$  is its Fourier transform

$$W(\lambda) = \frac{1}{2\pi} \sum_{|h| < n} k(h/p_n) e^{-ih\lambda}. \quad (4)$$

Expressions (1) and (3) provide the motivation for our test statistic. To test a general null hypothesis that the observations  $x_t$  are from a process with spectral density  $f(\cdot)$ , we propose the following test statistic

$$T_n = \left\{ \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \widehat{f}_e(\lambda_\ell) \right\}^{-2} \left\{ \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \widehat{f}_e^2(\lambda_\ell) \right\}, \quad (5)$$

where

$$\widetilde{f}_e(\lambda) = \frac{2\pi}{n} \sum_{j=1}^{n-1} \frac{W(\lambda - \lambda_j) I(\lambda_j)}{f(\lambda_j)},$$

and  $I$  is the periodogram of the observations  $x_1, \dots, x_n$ . Note that  $\widetilde{f}_e$  is a discrete version of  $\widehat{f}_e$  in (3) with  $I_{n,e}$  replaced by  $I/f$ . Thus, we whiten the process in the frequency domain instead of in the time domain. This not only avoids the computation of residuals but also allows one to easily test for arbitrary spectral densities. Furthermore,  $T_n$  is obtained by discretizing the integral in (1) with  $\widehat{f}_e$  replaced by  $\widetilde{f}_e$ . Also note that  $T_n$  is mean invariant because  $\widetilde{f}_e$  is evaluated only at Fourier frequencies. This is especially favourable in the presence of long memory, since the sample mean is not fully efficient in that case. (See Beran, 1994, p. 6)

Hong (1996) established the asymptotic normality of  $H_n$  for AR models. We show that  $T_n$  is asymptotically normal under a null hypothesis which can be either short memory or long memory if the process is Gaussian. The properties of a long memory process differ substantially from those of a short memory process and hence the proof of the asymptotic results for long memory models requires a more delicate approach than that for short memory models. We now state the assumptions we make and our main results.

Throughout the rest of this paper, we assume that  $\{x_t\}$  is a stationary linear process of the form,

$$x_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad (6)$$

where the innovations  $\varepsilon_t$  satisfy the following assumption.

**Assumption 1** *The series  $\{\varepsilon_t\}$  is independently and identically distributed with mean zero, variance  $\sigma^2$  and  $E(\varepsilon_t^8) < \infty$ .*

We also make the following assumptions about the kernel  $k(\cdot)$  and the bandwidth  $p_n$ .

**Assumption 2a** *The kernel function  $k : R \rightarrow [-1, 1]$  is a symmetric function that is continuous at zero and at all but a finite number of points, with  $k(0) = 1$ . If the kernel function  $k$  is of unbounded support, then for some  $\delta \geq 1$ ,  $z^\delta |k(z)| < \infty$  as  $z \rightarrow \infty$ .*

**Assumption 3** *The bandwidth  $p_n$  satisfies  $\log^6 n/p_n \rightarrow 0$  and  $p_n/n \rightarrow 0$ . If the kernel function  $k$  is of unbounded support, the bandwidth  $p_n$  also satisfies  $p_n^{\delta+1/2} \log n/n^\delta \rightarrow 0$ .*

More restrictive assumptions are made on the bandwidth when the kernel is of unbounded support with the choice of bandwidth depending on the rate of decay of the kernel. The faster the kernel decays, the less restrictive the condition on  $p_n$ . It is worth noting that all the kernels used in practice satisfy Assumption 2a. The next theorem states the asymptotic distribution of  $T_n$  when  $\{x_t\}$  is a short memory process.

**Theorem 1** *Let  $x_1, \dots, x_n$  be  $n$  observations from a stationary linear process defined by (6) with coefficients  $\psi_j$  such that  $\sum_{j=0}^{\infty} |\psi_j| j^{1/2} < \infty$  and innovations  $\varepsilon_t$  satisfying Assumption 1. Let  $f(\cdot)$  be the spectral density of the process such that  $\inf_{\lambda} f(\lambda) > 0$ . Let  $T_n$  be as in (5) and  $W$  be defined by (4) with kernel function  $k$  satisfying Assumption 2a and bandwidth  $p_n$  satisfying Assumption 3. Then*

$$\frac{n(T_n - C_n(k))}{D_n(k)^{1/2}} \rightarrow N(0, 1)$$

in distribution as  $n \rightarrow \infty$ , where

$$C_n(k) = \frac{1}{n\pi} \sum_{j=1}^{n-1} (1 - j/n) k^2(j/p_n) + \frac{1}{2\pi}$$

and

$$D_n(k) = \frac{2}{\pi^2} \sum_{j=1}^{n-2} \{(1 - j/n)(1 - (j + 1)/n)\} k^4(j/p_n).$$

It can be shown that a process satisfying the assumptions in Theorem 1 has bounded spectral density and autocovariances that are absolutely summable (Brockwell and Davis, 1991, ex 3.9). Such a process is a short memory process, an example of which is the autoregressive moving average (ARMA) model. If the chosen kernel is of compact support, assumptions 2a and 3a on the kernel function  $k$  and bandwidth  $p_n$  are identical to the assumptions made by Hong (1996) for the test statistic  $H_n$ . However, the assumptions on the process  $\{x_t\}$  of Theorem 1 are satisfied by a broad range of short memory models while the asymptotic theory of  $H_n$  is established only for autoregressive (AR) models.

To establish the asymptotic normality of  $T_n$  when the process is a long memory process, we restrict the process  $\{x_t\}$  to be Gaussian. We also require additional assumptions on  $k$  and  $p_n$ , which we state next.

**Assumption 2b** *In addition to Assumption 2a, the kernel function  $k$  is differentiable almost everywhere and satisfies  $\int |k'(z)k(z)| dz < \infty$ .*

All the kernels used in practice satisfy Assumption 2b. We now state the asymptotic distribution of  $T_n$  when  $\{x_t\}$  is a long memory process. For the long memory case, we make the extra assumption that the process  $x_t$  is Gaussian. We feel that this assumption can be relaxed just as in the short memory case in Theorem 1, though at the expense of much greater complexity in the proof.

**Theorem 2** *Let  $x_1, \dots, x_n$  be  $n$  observations from a stationary Gaussian linear process defined by (6) that has a spectral density  $f(\lambda) \sim a\lambda^{-2d}$ , as  $\lambda \rightarrow 0$ , where  $a$  is a constant and  $d \in (0, 0.5)$ . Also let the spectral density satisfy  $\inf_{\lambda} f(\lambda) > 0$ . Let  $T_n$  be defined as in Theorem 1 with kernel function  $k$  satisfying Assumption 2b and bandwidth  $p_n$  satisfying Assumption 3b. Then*

$$\frac{n(T_n - C_n(k))}{D_n(k)^{1/2}} \rightarrow N(0, 1)$$

where  $C_n(k)$  and  $D_n(k)$  are as in Theorem 1.

A stationary linear process which has a spectral density satisfying the assumption of Theorem 2 is a long memory process. It can be shown that the autocovariances decay to zero hyperbolically and are not summable for such a process (Theorem 2.24, Zygmund, 1959) Examples of long memory processes satisfying Assumption 2 are autoregressive fractionally integrated moving-average (ARFIMA) models (Granger & Joyeux, 1980 and Hosking, 1981) and fractional Gaussian noise (Mandelbrot and Van Ness, 1968).

In applications, the null hypothesis of interest is the composite hypothesis that the process has spectral density  $f(\boldsymbol{\theta}, \cdot)$  for some unknown  $\boldsymbol{\theta}$  in the parameter space  $\Theta$ . Under this composite null, the test statistic becomes

$$T_n(\hat{\boldsymbol{\theta}}) = \left\{ \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \tilde{f}_e(\hat{\boldsymbol{\theta}}, \lambda_\ell) \right\}^{-2} \left\{ \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \tilde{f}_e^2(\hat{\boldsymbol{\theta}}, \lambda_\ell) \right\}, \quad (7)$$

where

$$\tilde{f}_e(\hat{\boldsymbol{\theta}}, \lambda) = \frac{2\pi}{n} \sum_{j=1}^{n-1} \frac{W(\lambda - \lambda_j) I(\lambda_j)}{f(\hat{\boldsymbol{\theta}}, \lambda_j)}$$

and  $\hat{\boldsymbol{\theta}}$  is some estimator of  $\boldsymbol{\theta}$  based on the sample  $x_1, \dots, x_n$ . Under certain additional assumptions, we show in the next two theorems that the asymptotic null distribution of  $T_n(\hat{\boldsymbol{\theta}})$  remains the same as that of  $T_n$  in Theorem 1 and in Theorem 2. We first state the additional assumptions we need.

**Assumption 4** *Let  $\Theta_0$  be a compact subset of  $\Theta$ , where  $\Theta$  is a finite dimensional parameter space. Let the spectral density of the process  $\{x_t\}$  be  $f(\boldsymbol{\theta}_0, \cdot)$ , where  $\boldsymbol{\theta}_0$  is the true parameter vector that lies in the interior of  $\Theta_0$ . Assume that the estimator  $\hat{\boldsymbol{\theta}} \in \Theta$  satisfies  $\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\| = O_p(n^{-1/2})$ .*

The following is an assumption on the spectral density for short memory process.

**Assumption 5** *The spectral density  $f(\boldsymbol{\theta}, \lambda)$  satisfies the following conditions for  $(\boldsymbol{\theta}, \lambda) \in \Theta \times [0, 2\pi]$ :*

- (i)  $f(\boldsymbol{\theta}, \lambda)$  and  $f^{-1}(\boldsymbol{\theta}, \lambda)$  are continuous at all  $(\boldsymbol{\theta}, \lambda)$ .
- (ii)  $\frac{\partial}{\partial \theta_j} f^{-1}(\boldsymbol{\theta}, \lambda)$  and  $\frac{\partial^2}{\partial \theta_j \partial \theta_k} f^{-1}(\boldsymbol{\theta}, \lambda)$  are continuous and finite at all  $(\boldsymbol{\theta}, \lambda)$ .

It is very easy to establish that Assumptions 4 and 5 are satisfied by all ARMA models. The next theorem states the asymptotic distribution of  $T_n(\hat{\boldsymbol{\theta}})$  when  $\{x_t\}$  is a short memory process.

**Theorem 3** Let  $x_1, \dots, x_n$  be  $n$  observations from a stationary linear process satisfying the same assumptions as those of Theorem 1. Let the estimated parameter vector  $\hat{\boldsymbol{\theta}}$  satisfy Assumption 4 and the spectral density of the process  $\{x_t\}$  satisfy Assumption 5. Also let  $T_n(\hat{\boldsymbol{\theta}})$  be defined by (7) with kernel function  $k$  and bandwidth  $p_n$  satisfying the same assumptions as those of Theorem 1. Then

$$\frac{n \left( T_n(\hat{\boldsymbol{\theta}}) - C_n(k) \right)}{D_n(k)^{1/2}} \rightarrow N(0, 1)$$

in distribution as  $n \rightarrow \infty$ , where  $C_n(k)$  and  $D_n(k)$  are defined as in Theorem 1.

To establish the asymptotic distribution of  $T_n(\hat{\boldsymbol{\theta}})$  when  $\{x_t\}$  is a long memory process, we need the following assumption on  $\hat{\boldsymbol{\theta}}$  and the spectral density  $f(\boldsymbol{\theta}, \cdot)$ .

**Assumption 6** Let  $\Theta_0$  be a compact subset of  $\Theta$ , where  $\Theta$  is a finite dimensional parameter space. Let the spectral density of the process  $\{x_t\}$  be  $f(\boldsymbol{\theta}_0, \cdot)$ , where  $\boldsymbol{\theta}_0 = (\boldsymbol{\beta}, d_0)'$  is the true parameter vector that lies in the interior of  $\Theta_0 \times [\delta_1, 0.5 - \delta_1]$  for some  $0 < \delta_1 < 0.5$ . Assume that the estimator  $\hat{\boldsymbol{\theta}} \in \Theta$  satisfies  $\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\| = O_p(n^{-1/2})$ .

**Assumption 7** Let  $\boldsymbol{\theta} = (\boldsymbol{\beta}, d)'$ , where  $(\boldsymbol{\beta}, d) \in \Theta \times (0, 2\pi)$ ,  $d \in (0, 0.5)$ . For any  $\delta > 0$ , the spectral density  $f(\boldsymbol{\theta}, \lambda)$  satisfies the following conditions:

(i)  $f(\boldsymbol{\theta}, \lambda)$  is continuous at all  $(\boldsymbol{\theta}, \lambda)$ ,  $f^{-1}(\boldsymbol{\theta}, \lambda)$  is continuous at all  $(\boldsymbol{\theta}, \lambda)$  and

$$f(\boldsymbol{\theta}, \lambda) = O(|\lambda|^{-2d}) \text{ as } \lambda \rightarrow 0.$$

(ii)  $\frac{\partial}{\partial \theta_j} f^{-1}(\boldsymbol{\theta}, \lambda)$  and  $\frac{\partial^2}{\partial \theta_j \partial \theta_k} f^{-1}(\boldsymbol{\theta}, \lambda)$  are continuous at all  $(\boldsymbol{\theta}, \lambda)$  and

$$\frac{\partial}{\partial \theta_j} f^{-1}(\boldsymbol{\theta}, \lambda) = O(|\lambda|^{2d-\delta}) \text{ as } \lambda \rightarrow 0,$$

$$\frac{\partial^2}{\partial \theta_j \partial \theta_k} f^{-1}(\boldsymbol{\theta}, \lambda) = O(|\lambda|^{2d-\delta}) \text{ as } \lambda \rightarrow 0.$$

(iii) There exists a constant  $C$  with

$$|f(\boldsymbol{\theta}_1, \lambda) - f(\boldsymbol{\theta}_2, \lambda)| \leq C \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\| f(\boldsymbol{\theta}_2, \lambda)$$



uniformly for all  $\lambda$  and all  $\boldsymbol{\theta}_1 = (\boldsymbol{\beta}_1, d_1)'$  and  $\boldsymbol{\theta}_2 = (\boldsymbol{\beta}_2, d_2)'$  such that  $d_1 < d_2$ .

All the conditions of Assumptions 6 and 7 are satisfied by fractional Gaussian noise and ARFIMA processes (See Dahlhaus (1989)). We now state the asymptotic distribution of  $T_n(\hat{\boldsymbol{\theta}})$  when  $\{x_t\}$  is a long memory process.

**Theorem 4** *Let  $x_1, \dots, x_n$  be  $n$  observations from a stationary Gaussian linear process satisfying the same assumptions as those of Theorem 2. Let the estimated parameter vector  $\hat{\boldsymbol{\theta}}$  satisfy Assumption 6 and the spectral density of  $\{x_t\}$  satisfy Assumption 7. Also let  $T_n(\hat{\boldsymbol{\theta}})$  be defined by (7) with kernel function  $k$  and bandwidth  $p_n$  satisfying the same assumptions as those of Theorem 2. Then*

$$\frac{n(T_n(\hat{\boldsymbol{\theta}}) - C_n(k))}{D_n(k)^{1/2}} \rightarrow N(0, 1)$$

in distribution as  $n \rightarrow \infty$ , where  $C_n(k)$  and  $D_n(k)$  are defined as in Theorem 1.

### 3 Simulation Studies

We generated 5000 replications of Gaussian series of length  $n = 128$  and  $512$  from a variety of AR and ARFIMA processes. The algorithm of Davies and Harte (1987) was used in the data generation of ARFIMA models. For each series, we computed the three test statistics: (i) Our statistic  $T_n$ . (ii) Hong's statistic  $H_n$ . (iii) The Milhoj statistic  $M_n$ . The statistics were suitably normalized so that they would have an asymptotic standard normal distribution under the null.

For  $T_n$  and  $H_n$ , we used the following three kernels:

(i) Bartlett  $k(z) = 1 - |z|, |z| \leq 1,$

$$= 0 \quad \text{otherwise,}$$

(ii) Tukey  $k(z) = \frac{1}{2}(\cos(z\pi) + 1), |z| \leq 1,$

$$= 0 \quad \text{otherwise,}$$

(iii) Quadratic Spectral (QS),  $k(z) = \frac{25}{12z^2} \left( \frac{\sin(6\pi z/5)}{6\pi z/5} - \cos(6\pi z/5) \right), z \in (-\infty, \infty).$

For computing  $T_n$  and  $H_n$ , we used three bandwidths,  $p_n = [3n^{0.2}]$ ,  $[3n^{0.3}]$  and  $[3n^{0.4}]$ . Note that there is no bandwidth involved in computing  $M_n$ .

In Tables I and II, we report the sizes of the three tests under the null hypothesis of an AR(1) and an ARFIMA(0,d,0) respectively. The AR(1) parameter was set to 0.8 while the long memory parameter  $d$  in the ARFIMA(0,  $d$ , 0) was set at 0.4. It can be seen that for both the models, all three statistics are undersized at both the 5% and 10% level. The amount by which they are undersized decreases as the bandwidth  $p_n$  increases. The  $M_n$  statistic is least undersized, while the sizes of  $T_n$  are comparable to those of  $H_n$ .

To compare the power of the tests, we considered the following four cases: (a) fitting an AR(1) to data generated by an AR(2),  $x_t = 0.8x_{t-1} - 0.1x_{t-2} + \varepsilon_t$ . (b) fitting an ARFIMA(1,  $d$ , 0) to data generated by an ARM(1,1),  $x_t = 0.8x_{t-1} + \varepsilon_t + 0.2\varepsilon_{t-1}$ . (c) fitting an ARMA(1, 1) to data generated by an ARFIMA(0,  $d$ , 0),  $(1 - B)^{0.4} x_t = \varepsilon_t$  where  $B$  denotes the backshift operator (d) fitting an ARFIMA(0,  $d$ , 0) to data generated by an ARFIMA(1,  $d$ , 0),  $(1 - B)^{0.4} (1 - 0.1B) x_t = \varepsilon_t$ . The results are reported in tables III, IV, V and VI respectively.

It is seen that both the tests  $T_n$  and  $H_n$  have significantly higher power than  $M_n$  in all the alternatives considered. This is not surprising, since the tests  $T_n$  and  $H_n$  give decreasing weights to higher lag sample correlations, while  $M_n$  gives uniform weight at all lags. It might be tempting to believe that this property of  $M_n$  may be useful in detecting long memory alternatives. This belief is however belied by Table V, where we fit a short memory model to a long memory series and yet  $M_n$  is outperformed by a wide margin by both the other tests. On the other hand, it is seen that the power of  $T_n$  is very similar to the power of  $H_n$ , with neither test outperforming the other significantly in any situation considered.

## 4 Appendix: Proofs

We will only provide the proofs for long memory models. The proofs for short memory models are similar though much simpler and are available from the authors. In this appendix, we will often use the following decomposition of  $I(\lambda)$ ,

$$I(\lambda) = |\psi(\lambda)|^2 I_\varepsilon(\lambda) + I(\lambda) - |\psi(\lambda)|^2 I_\varepsilon(\lambda),$$

where  $\psi(\lambda) = \sum_{k=0}^{\infty} \psi_k e^{-i\lambda k}$  and  $I_\varepsilon(\lambda)$  is the periodogram of the innovations  $\varepsilon_t$  in (6). Then

$$\frac{I(\lambda)}{f(\lambda)} = \frac{2\pi}{\sigma^2} I_\varepsilon(\lambda) + R(\lambda), \quad (8)$$

where

$$R(\lambda) = \frac{I(\lambda)}{f(\lambda)} - \frac{2\pi}{\sigma^2} I_\varepsilon(\lambda). \quad (9)$$

Let  $\hat{\gamma}_{\varepsilon,j}$  be the  $j^{\text{th}}$  sample covariance of the  $\varepsilon_t$  given by  $\hat{\gamma}_{\varepsilon,j} = n^{-1} \sum_{t=|j|+1}^n (\varepsilon_t - \bar{\varepsilon})(\varepsilon_{t-|j|} - \bar{\varepsilon})$ , for  $|j| \leq n-1$ .

### Proof of Theorem 2

Let  $I_\varepsilon(\lambda) = (2\pi n)^{-1} \left| \sum_{t=1}^n \varepsilon_t e^{i\lambda t} \right|^2$  be the periodogram of the innovations  $\varepsilon_t$  without mean correction. For the Fourier frequencies,  $\lambda_k$ ,  $k = 1, \dots, (n-1)$ , we have  $I_\varepsilon(\lambda_k) = I_{n,\varepsilon}(\lambda_k)$ , where  $I_{n,\varepsilon}$  is the periodogram of the mean corrected innovations  $\varepsilon_t - \bar{\varepsilon}$ . Also define

$$\hat{f}_{\varepsilon,d}(\lambda) = \frac{2\pi}{n} \sum_{j=1}^{n-1} W(\lambda - \lambda_j) I_\varepsilon(\lambda_j).$$

In Lemmas (1), (2) and (3) below, we show that

$$\frac{n}{p_n^{1/2}} \left\{ \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \left( \tilde{f}_e^2(\lambda_\ell) - \frac{4\pi^2}{\sigma^4} \hat{f}_{\varepsilon,d}^2(\lambda_\ell) \right) \right\} = o_p(1),$$

$$\frac{n}{p_n^{1/2}} \left\{ \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \frac{4\pi^2}{\sigma^4} \hat{f}_{\varepsilon,d}^2(\lambda_\ell) - \frac{2\pi}{\sigma^4} \sum_{j=-(n-1)}^{n-1} k^2(j/p_n) \hat{\gamma}_{\varepsilon,j}^2 \right\} = o_p(1),$$

and

$$\frac{n}{p_n^{1/2}} \left[ \left\{ \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \tilde{f}_e(\lambda_\ell) \right\}^2 - \left( \frac{2\pi}{\sigma^2} \hat{\gamma}_{\varepsilon,0} \right)^2 \right] = o_p(1).$$

Also, by Lemma 3,  $\left\{ (2\pi/n) \sum_{\ell=0}^{n-1} (2\pi/\sigma^2) \tilde{f}_e(\lambda_\ell) \right\}^2 = (4\pi^2/\sigma^4) \hat{\gamma}_{\varepsilon,0}^2$  and  $\sqrt{n}(\hat{\gamma}_{\varepsilon,0} - \sigma^2) = O_p(1)$ . The Theorem now follows by Theorem 1 of Hong (1996) and the fact that  $p_n^{-1} D_n(k) \rightarrow D(k) \equiv \int_0^\infty k^4(z) dz < \infty$  as  $n \rightarrow \infty$  by assumption 2a.

### Proof of Theorem 4

By Theorem 1 it suffices to show that

$$\frac{n}{p_n^{1/2}} \left( T_n(\boldsymbol{\theta}_0) - T_n(\hat{\boldsymbol{\theta}}) \right) = o_p(1), \quad (10)$$

which we do by establishing that

$$\frac{n}{p_n^{1/2}} \left\{ \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \left( \tilde{f}_e^2(\boldsymbol{\theta}_0, \lambda_\ell) - \tilde{f}_e^2(\hat{\boldsymbol{\theta}}, \lambda_\ell) \right) \right\} = o_p(1) \quad (11)$$

and

$$\frac{n}{p_n^{1/2}} \left[ \left\{ \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \tilde{f}_e(\boldsymbol{\theta}_0, \lambda_\ell) \right\}^2 - \left\{ \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \tilde{f}_e(\hat{\boldsymbol{\theta}}, \lambda_\ell) \right\}^2 \right] = o_p(1). \quad (12)$$

We will prove only (11) since the proof of (12) is similar. Let

$$\mathcal{G}_{\boldsymbol{\theta}}(\lambda_j, \lambda_h) = \frac{1}{f_{\boldsymbol{\theta}}(\lambda_j)} \frac{1}{f_{\boldsymbol{\theta}}(\lambda_h)}.$$

Then the LHS of (11) is

$$\begin{aligned} & \frac{n}{p_n^{1/2}} \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \left( \frac{2\pi}{n} \sum_{j=1}^{n-1} W(\lambda_{\ell-j}) \frac{I(\lambda_j)}{f_{\boldsymbol{\theta}_0}(\lambda_j)} \right)^2 - \left( \frac{2\pi}{n} \sum_{j=1}^{n-1} W(\lambda_{\ell-j}) \frac{I(\lambda_j)}{f_{\hat{\boldsymbol{\theta}}}(\lambda_j)} \right)^2 \\ &= \frac{(2\pi)^2}{n^2 p_n^{1/2}} \sum_{\ell=0}^{n-1} \sum_{j,h=1}^{n-1} W(\lambda_{\ell-j}) W(\lambda_{\ell-h}) I(\lambda_j) I(\lambda_h) \left( \mathcal{G}_{\boldsymbol{\theta}_0}(\lambda_j, \lambda_h) - \mathcal{G}_{\hat{\boldsymbol{\theta}}}(\lambda_j, \lambda_h) \right). \end{aligned}$$

By a similar argument of deriving (28), the RHS of the above equation is

$$\frac{2\pi}{n p_n^{1/2}} \sum_{j,h=1}^{n-1} I(\lambda_j) I(\lambda_h) \left( \mathcal{G}_{\boldsymbol{\theta}_0}(\lambda_j, \lambda_h) - \mathcal{G}_{\hat{\boldsymbol{\theta}}}(\lambda_j, \lambda_h) \right) \mathcal{K}_n(\lambda_{j-h}), \quad (13)$$

where  $\mathcal{K}_n(\lambda_{j-h})$  is defined as (29). For every  $\lambda_j$  and  $\lambda_h$ , we have by a Taylor series expansion,

$$\begin{aligned} & \mathcal{G}(\lambda_j, \lambda_h, \boldsymbol{\theta}_0) - \mathcal{G}(\lambda_j, \lambda_h, \hat{\boldsymbol{\theta}}) \\ &= \sum_u \left( \frac{1}{f_{\boldsymbol{\theta}_0}(\lambda_h)} \frac{\partial f^{-1}(\lambda_j, \boldsymbol{\theta}_0)}{\partial \theta_u} + \frac{1}{f_{\boldsymbol{\theta}_0}(\lambda_j)} \frac{\partial f^{-1}(\lambda_h, \boldsymbol{\theta}_0)}{\partial \theta_u} \right) (\hat{\theta}_u - \theta_{0u}) \\ & \quad + \frac{1}{2} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)' \frac{\partial^2 \mathcal{G}(\lambda_j, \lambda_h, \tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}^2} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0), \end{aligned}$$

where  $\tilde{\boldsymbol{\theta}}_{jh} = \boldsymbol{\theta}_0 + \alpha_{jh} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$  for some  $0 < \alpha_{jh} < 1$  and

$$\begin{aligned} \frac{\partial^2 \mathcal{G}(\lambda_j, \lambda_h, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} &= \frac{1}{f_{\boldsymbol{\theta}}(\lambda_h)} \frac{\partial^2 f^{-1}(\lambda_j, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} + \frac{\partial f^{-1}(\lambda_j, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial f^{-1}(\lambda_h, \boldsymbol{\theta})'}{\partial \boldsymbol{\theta}} \\ & \quad + \frac{1}{f_{\boldsymbol{\theta}}(\lambda_j)} \frac{\partial^2 f^{-1}(\lambda_h, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} + \frac{\partial f^{-1}(\lambda_h, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial f^{-1}(\lambda_j, \boldsymbol{\theta})'}{\partial \boldsymbol{\theta}} \end{aligned} \quad (14)$$

To prove (11), we will show that (13) is  $o_p(1)$  by verifying, for each  $u$ ,

$$\frac{2\pi}{np_n^{1/2}} \sum_{j,h=1}^{n-1} I(\lambda_j)I(\lambda_h) \left( \frac{1}{f_{\theta_0}(\lambda_j)} \frac{\partial f^{-1}(\lambda_h, \theta_0)}{\partial \theta_u} \right) (\hat{\theta}_u - \theta_{0u}) \mathcal{K}_n(\lambda_{j-h}) = o_p(1), \quad (15)$$

and

$$\frac{2\pi}{np_n^{1/2}} \sum_{j,h=1}^{n-1} I(\lambda_j)I(\lambda_h) (\hat{\theta} - \theta_0)' \frac{\partial^2 \mathcal{G}(\lambda_j, \lambda_h, \hat{\theta})}{\partial \theta^2} (\hat{\theta} - \theta_0) \mathcal{K}_n(\lambda_{j-h}) = o_p(1). \quad (16)$$

We first show (15). Let

$$g(\lambda) = \frac{\partial \ln f(\lambda, \theta_0)}{\partial \theta_u}, \quad (17)$$

then

$$\frac{1}{f_{\theta_0}(\lambda_j)} \frac{\partial f^{-1}(\lambda_h, \theta_0)}{\partial \theta_u} = - \frac{1}{f_{\theta_0}(\lambda_j)} \frac{1}{f_{\theta_0}(\lambda_h)} g(\lambda_h).$$

Since  $\hat{\theta} - \theta_0 = O_p(n^{-1/2})$ , (15) is true if

$$\sum_{j,h=1}^{n-1} \frac{I(\lambda_j)}{f_{\theta_0}(\lambda_j)} \frac{I(\lambda_h)}{f_{\theta_0}(\lambda_h)} g(\lambda_h) \mathcal{K}_n(\lambda_{j-h}) = o_p(n^{3/2} p_n^{1/2}).$$

By (8), it is thus enough to show that

$$\sum_{j,h=1}^{n-1} I_\varepsilon(\lambda_j) I_\varepsilon(\lambda_h) g(\lambda_h) \mathcal{K}_n(\lambda_{j-h}) = o_p(n^{3/2} p_n^{1/2}), \quad (18)$$

$$\sum_{j,h=1}^{n-1} I_\varepsilon(\lambda_j) R_{\theta_0}(\lambda_h) g(\lambda_h) \mathcal{K}_n(\lambda_{j-h}) = o_p(n^{3/2} p_n^{1/2}), \quad (19)$$

and

$$\sum_{j,h=1}^{n-1} R_{\theta_0}(\lambda_j) R_{\theta_0}(\lambda_h) g(\lambda_h) \mathcal{K}_n(\lambda_{j-h}) = o_p(n^{3/2} p_n^{1/2}). \quad (20)$$

Since  $g(\lambda) = O(\lambda^{-\delta})$  by assumption 7, (19) and (20) can be shown by an argument similar to that used to establish (30) and (31). To show (18), we let

$$a_m = \sum_{h=1}^{n-1} g(\lambda_h) e^{-i\lambda_h m}.$$

The LHS of (18) is

$$\sum_{j,h=1}^{n-1} I_\varepsilon(\lambda_j) I_\varepsilon(\lambda_h) g(\lambda_h) \left( \sum_{p=-(n-1)}^{n-1} k_p^2 e^{i\lambda_j - hp} + 2 \sum_{p=1}^{n-1} k_p k_{n-p} e^{i\lambda_j - hp} \right)$$

$$\begin{aligned}
&= \frac{1}{(2\pi n)^2} \sum_{p=-(n-1)}^{n-1} k_p^2 \sum_{s,t,u,v} \varepsilon_s \varepsilon_t \varepsilon_u \varepsilon_v \sum_{h=1}^{n-1} g(\lambda_h) e^{-i\lambda_h(u-v-p)} \sum_{j=1}^{n-1} e^{-i\lambda_j(s-t+p)} \\
&\quad + \frac{2}{(2\pi n)^2} \sum_{p=1}^{n-1} k_p k_{n-p} \sum_{s,t,u,v} \varepsilon_s \varepsilon_t \varepsilon_u \varepsilon_v \sum_{h=1}^{n-1} g(\lambda_h) e^{-i\lambda_h(u-v-p)} \sum_{j=1}^{n-1} e^{-i\lambda_j(s-t+p)} \\
&= \frac{1}{4\pi^2 n} \left( \sum_{p=-(n-1)}^{n-1} k_p^2 \sum_{t,u,v} a_{u-v-p} \varepsilon_t \varepsilon_{t-p} \varepsilon_u \varepsilon_v + 2 \sum_{p=1}^{n-1} k_p k_{n-p} \sum_{t,u,v} a_{u-v-p} \varepsilon_t \varepsilon_{t-p} \varepsilon_u \varepsilon_v \right) \\
&\quad - \frac{1}{4\pi^2 n^2} \left( \sum_{p=-(n-1)}^{n-1} k_p^2 \sum_{s,t,u,v} a_{u-v-p} \varepsilon_s \varepsilon_t \varepsilon_u \varepsilon_v + 2 \sum_{p=1}^{n-1} k_p k_{n-p} \sum_{s,t,u,v} a_{u-v-p} \varepsilon_s \varepsilon_t \varepsilon_u \varepsilon_v \right). \quad (21)
\end{aligned}$$

We will show that both terms of the last expression in (21) have second moments of order  $o(n^3 p_n)$ .

By the Cauchy Schwarz inequality, we have

$$\begin{aligned}
&E \left( \sum_{p=-(n-1)}^{n-1} k_p^2 \sum_{t,u,v} a_{u-v-p} \varepsilon_t \varepsilon_{t-p} \varepsilon_u \varepsilon_v + 2 \sum_{p=1}^{n-1} k_p k_{n-p} \sum_{t,u,v} a_{u-v-p} \varepsilon_t \varepsilon_{t-p} \varepsilon_u \varepsilon_v \right)^2 \\
&= O\left(\frac{1}{16\pi^4 n^2} \sum_{p_1, p_2=-(n-1)}^{n-1} k_{p_1}^2 k_{p_2}^2 \sum_{t_1, t_2, u_1, u_2, v_1, v_2} a_{u_1-v_1-p_1} a_{u_2-v_2-p_2} \right. \\
&\quad \left. \times E(\varepsilon_{t_1} \varepsilon_{t_2} \varepsilon_{t_1-p_1} \varepsilon_{t_2-p_2} \varepsilon_{u_1} \varepsilon_{u_2} \varepsilon_{v_1} \varepsilon_{v_2})\right) \\
&\quad + O\left(\frac{1}{16\pi^4 n^2} \sum_{p_1, p_2=-(n-1)}^{n-1} k_{n-p_1}^2 k_{n-p_2}^2 \sum_{t_1, t_2, u_1, u_2, v_1, v_2} a_{u_1-v_1-p_1} a_{u_2-v_2-p_2} \right. \\
&\quad \left. \times E(\varepsilon_{t_1} \varepsilon_{t_2} \varepsilon_{t_1-p_1} \varepsilon_{t_2-p_2} \varepsilon_{u_1} \varepsilon_{u_2} \varepsilon_{v_1} \varepsilon_{v_2})\right)
\end{aligned}$$

Since  $\varepsilon_t$  are independent with zero mean, the above expectation is positive only when the random variables inside the parentheses consist of products of even powers of the  $\varepsilon_t$ . Thus, the above expression is dominated by two cases: one is when  $p_1 = p_2 = 0$ ,  $u_1 = u_2$  and  $v_1 = v_2$  while the other is when  $p_1 = p_2 = 0$ ,  $u_1 = v_1$  and  $u_2 = v_2$ . Using lemma 6, the order of these two cases is

$$\begin{aligned}
&O \left( n^{-2} \sum_{t_1, t_2, u_1, u_2} a_{u_1-u_2}^2 + n^{-2} \sum_{t_1, t_2, u_1, u_2} a_0^2 \right) \\
&= O \left( n^{-2} n^{5+2\delta} + n^{-2} n^{4+2\delta} \right) \\
&= o \left( n^3 p_n \right).
\end{aligned}$$

It can be shown that the second moment of the second term in (21) is also of order  $o(n^3 p_n)$  by similar arguments. We have thus established (15).

Next, we establish (16). Let  $\mathcal{A}_{uv}(\lambda_j, \lambda_h, \tilde{\boldsymbol{\theta}}_{jh})$  denote the  $(u, v)$ <sup>th</sup> element of the matrix

$$\frac{\partial^2 \mathcal{G}(\lambda_j, \lambda_h, \tilde{\boldsymbol{\theta}}_{jh})}{\partial \boldsymbol{\theta}^2}.$$

Then, by (14) and Assumption 7,

$$\mathcal{A}_{uv}(\lambda_j, \lambda_h, \tilde{\boldsymbol{\theta}}_{jh}) = O\left(\lambda_j^{2\tilde{d}_{jh}} \lambda_h^{2\tilde{d}_{jh}} \lambda_j^{-\delta} \lambda_h^{-\delta}\right), \quad (22)$$

where  $\tilde{\boldsymbol{\theta}}_{jh} = (\tilde{\boldsymbol{\beta}}_{jh}, \tilde{d}_{jh})'$ . Since  $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 = O_p(n^{-1/2})$ , (16) will follow if for every  $(u, v)$

$$\sum_{j,h=1}^{n-1} I(\lambda_j) I(\lambda_h) \mathcal{A}_{uv}(\lambda_j, \lambda_h, \tilde{\boldsymbol{\theta}}_{jh}) \mathcal{K}_n(\lambda_{j-h}) = o_p(n^2 p_n^{1/2}).$$

To show this, it suffices, by (8), to prove that

$$\sum_{j,h=1}^{n-1} f(\boldsymbol{\theta}_0, \lambda_j) f(\boldsymbol{\theta}_0, \lambda_h) I_\varepsilon(\lambda_j) I_\varepsilon(\lambda_h) \mathcal{A}_{uv}(\lambda_j, \lambda_h, \tilde{\boldsymbol{\theta}}_{jh}) \mathcal{K}_n(\lambda_{j-h}) = o_p(n^2 p_n^{1/2}), \quad (23)$$

$$\sum_{j,h=1}^{n-1} f(\boldsymbol{\theta}_0, \lambda_j) f(\boldsymbol{\theta}_0, \lambda_h) I_\varepsilon(\lambda_j) R_{\boldsymbol{\theta}_0}(\lambda_h) \mathcal{A}_{uv}(\lambda_j, \lambda_h, \tilde{\boldsymbol{\theta}}_{jh}) \mathcal{K}_n(\lambda_{j-h}) = o_p(n^2 p_n^{1/2}),$$

and

$$\sum_{j,h=1}^{n-1} f(\boldsymbol{\theta}_0, \lambda_j) f(\boldsymbol{\theta}_0, \lambda_h) R_{\boldsymbol{\theta}_0}(\lambda_j) R_{\boldsymbol{\theta}_0}(\lambda_h) \mathcal{A}_{uv}(\lambda_j, \lambda_h, \tilde{\boldsymbol{\theta}}_{jh}) \mathcal{K}_n(\lambda_{j-h}) = o_p(n^2 p_n^{1/2}).$$

We will prove only the first of these, since the proof for the other two is similar. Letting

$$Y_n = \sum_{j,h=1}^{n-1} f(\boldsymbol{\theta}_0, \lambda_j) f(\boldsymbol{\theta}_0, \lambda_h) I_\varepsilon(\lambda_j) I_\varepsilon(\lambda_h) \mathcal{A}_{uv}(\lambda_j, \lambda_h, \tilde{\boldsymbol{\theta}}_{jh}) \mathcal{K}_n(\lambda_{j-h}),$$

we have

$$Y_n = I(\hat{d} \geq d_0) Y_n + I(\hat{d} < d_0) Y_n. \quad (24)$$

First consider  $\hat{d} \geq d_0$ . Then  $\tilde{d}_{jh} \geq d_0$  for all  $j, h$ . Hence, by Assumption 7 and (22), we have  $f(\boldsymbol{\theta}_0, \lambda_j) f(\boldsymbol{\theta}_0, \lambda_h) \mathcal{A}_{uv}(\lambda_j, \lambda_h, \tilde{\boldsymbol{\theta}}_{jh}) = O(\lambda_j^{-\delta} \lambda_h^{-\delta})$  for all  $j, h$ . Also, by the Cauchy Schwarz inequality,  $\sup_{j,h} E |I_\varepsilon(\lambda_j) I_\varepsilon(\lambda_h)| < K < \infty$  and it follows from (32) that

$$E \left| I(\hat{d} \geq d_0) Y_n \right| = O\left(\sum_{j,h=1}^{n-1} \lambda_j^{-\delta} \lambda_h^{-\delta} n(j-h)^{-1}\right) = O(n^2 \log n) = o(n^2 p_n^{1/2}). \quad (25)$$

Now consider  $\hat{d} < d_0$ . Then  $0 < \tilde{d}_{jh} < d_0$  for all  $j, h$ . By part (iii) of Assumption 7 we get that

$$f(\boldsymbol{\theta}_0, \lambda_j) f(\boldsymbol{\theta}_0, \lambda_h) f^{-1}(\tilde{\boldsymbol{\theta}}_{jh}, \lambda_j) f^{-1}(\tilde{\boldsymbol{\theta}}_{jh}, \lambda_h) = (1 + \Delta_j)(1 + \Delta_h),$$

where  $|\Delta_j| \leq K \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\| \lambda_j^{-2d_0}$  for all  $j$ . Furthermore,  $f(\tilde{\boldsymbol{\theta}}_{jh}, \lambda_j) f(\tilde{\boldsymbol{\theta}}_{jh}, \lambda_h) \mathcal{A}_{uv}(\lambda_j, \lambda_h, \tilde{\boldsymbol{\theta}}_{jh}) = O(\lambda_j^{-\delta} \lambda_h^{-\delta})$  by (22). Using these bounds and the fact that  $\sup_{j,h} E |I_\varepsilon(\lambda_j) I_\varepsilon(\lambda_h)| < K < \infty$ , we get

$$\begin{aligned} E |I(\hat{d} < d_0) Y_n| &= O(n^2 \log n) + O(n^{3/2} \log n) \\ &\quad + I(d < 1/4) O(n \log n) + I(d \geq 1/4) O(n^{4d+2\delta} \log n) \\ &= o(n^2 p_n^{1/2}). \end{aligned} \quad (26)$$

Thus, (23) follows from (24), (25) and (26).

**Lemma 1** *Under the assumptions in Theorem 2,*

$$\frac{n}{p_n^{1/2}} \left\{ \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \left( \hat{f}_e^2(\lambda_\ell) - \frac{4\pi^2}{\sigma^4} \hat{f}_{\varepsilon,d}^2(\lambda_\ell) \right) \right\} = o_p(1). \quad (27)$$

**Proof.** The LHS of (27) is

$$\begin{aligned} &\frac{n}{p_n^{1/2}} \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \left[ \left( \frac{2\pi}{n} \sum_{j=1}^{n-1} W(\lambda_{\ell-j}) \frac{I(\lambda_j)}{f(\lambda_j)} \right)^2 - \left( \frac{2\pi}{\sigma^2} \frac{2\pi}{n} \sum_{j=1}^{n-1} W(\lambda_{\ell-j}) I_\varepsilon(\lambda_\ell) \right)^2 \right] \\ &= \frac{(2\pi)^3}{n^2 p_n^{1/2}} \sum_{\ell=0}^{n-1} \sum_{j,h=1}^{n-1} W(\lambda_{\ell-j}) W(\lambda_{\ell-h}) \{ I_\varepsilon(\lambda_j) R(\lambda_h) + I_\varepsilon(\lambda_h) R(\lambda_j) + R(\lambda_j) R(\lambda_h) \}. \end{aligned}$$

Letting  $k_s = k(s/p_n)$  and  $\Phi(\lambda_j, \lambda_h) = I_\varepsilon(\lambda_j) R(\lambda_h) + I_\varepsilon(\lambda_h) R(\lambda_j) + R(\lambda_j) R(\lambda_h)$ , the last line of the above equation becomes

$$\begin{aligned} &\frac{2\pi}{n^2 p_n^{1/2}} \sum_{\ell=0}^{n-1} \sum_{j,h=1}^{n-1} \sum_{p,q=-(n-1)}^{n-1} k_p k_q e^{-i(\lambda_{\ell-j} p + \lambda_{\ell-h} q)} \Phi(\lambda_j, \lambda_h) \\ &= \frac{2\pi}{n^2 p_n^{1/2}} \sum_{j,h=1}^{n-1} \sum_{p,q=-(n-1)}^{n-1} k_p k_q e^{i(\lambda_j p + \lambda_h q)} \Phi(\lambda_j, \lambda_h) \sum_{\ell=0}^{n-1} e^{-i\lambda_\ell (p+q)} \\ &= \frac{2\pi}{n p_n^{1/2}} \left( \sum_{j,h=1}^{n-1} \Phi(\lambda_j, \lambda_h) \left( \sum_{p=-(n-1)}^{n-1} k_p^2 e^{i\lambda_j - h p} + 2 \sum_{p=1}^{n-1} k_p k_{n-p} e^{i\lambda_j - h p} \right) \right). \end{aligned} \quad (28)$$



Let

$$\mathcal{K}_n(\lambda_s) = \sum_{p=-(n-1)}^{n-1} k_p^2 e^{i\lambda_j - hp} + 2 \sum_{p=1}^{n-1} k_p k_{n-p} e^{i\lambda_j - hp}. \quad (29)$$

We will show that (28) is  $o_p(1)$  by verifying

$$E \left| \sum_{j,h=1}^{n-1} R(\lambda_j) I_\varepsilon(\lambda_h) \mathcal{K}_n(\lambda_{j-h}) \right| = o\left(np_n^{1/2}\right), \quad (30)$$

and

$$E \left| \sum_{j,h=1}^{n-1} R(\lambda_j) R(\lambda_h) \mathcal{K}_n(\lambda_{j-h}) \right| = o\left(np_n^{1/2}\right). \quad (31)$$

To prove the above two equations, we will need a bound for  $\mathcal{K}_n(\lambda_s)$ . Using the facts that  $p_n^{-1} \sum |k_p| = O(1)$  and  $\sum_{\ell=a}^b e^{i\lambda \ell} = O(\lambda^{-1})$  uniformly in  $a$  and  $b$  for  $0 < \lambda < \pi$  (see Zygmund, 1977, p. 2), by applying summation by parts and by Assumption 2b, for  $s \neq 0$  we obtain

$$\begin{aligned} \sum_{p=-(n-1)}^{n-1} k_p^2 e^{i\lambda_s p} &= 2 \sum_{p=1}^{n-1} k_p^2 \cos(\lambda_s p) + 1 \\ &= 2 \sum_{p=1}^{n-2} (k_p^2 - k_{p+1}^2) \sum_{u=1}^p \cos(\lambda_s u) + 2k_{n-1}^2 \sum_{u=1}^{n-1} \cos(\lambda_s u) + 1 \\ &= 2 \sum_{p=1}^{n-2} (k_p - k_{p+1})(k_p + k_{p+1}) \sum_{u=1}^p \cos(\lambda_s u) + O(1) \\ &= O\left(\sum_{m=1}^{n-2} \frac{1}{p_n} k'_p (k_p + k_{p+1}) \sum_{u=1}^p \cos(\lambda_s u)\right) \\ &= O(\lambda_s^{-1}), \end{aligned}$$

where  $p < \tilde{p} < p + 1$ . Similarly,

$$\sum_{p=1}^{n-1} k_p k_{n-p} e^{i\lambda_s p} = O(\lambda_s^{-1}),$$

and hence

$$\mathcal{K}_n(\lambda_s) = O(\lambda_s^{-1}). \quad (32)$$

We shall only derive (30) and (31) when  $j \neq h$ , since the proofs for  $j = h$  are similar and simpler.

To prove (30) we note that the LHS of (30) is bounded by

$$E \left| \sum_{j=1}^{\log^2 n} \sum_{h \neq j} R(\lambda_j) I_\varepsilon(\lambda_h) \mathcal{K}_n(\lambda_{j-h}) \right| + 2E \left| \sum_{j=\log^2 n+1}^{n-1} \sum_{h=1}^{\log^2 n} R(\lambda_j) I_\varepsilon(\lambda_h) \mathcal{K}_n(\lambda_{j-h}) \right|$$

$$+E \left| \sum_{j=\log^2 n+1}^{n-1} \sum_{h=\log^2 n+1}^{n-1} R(\lambda_j) I_\varepsilon(\lambda_h) \mathcal{K}_n(\lambda_{j-h}) \right|.$$

Using the Cauchy Schwarz inequality, lemma 5, equation (32) and the fact that  $\max_j E(I_\varepsilon^2(\lambda_j)) < \infty$ , the first term and second term of the equation above are of the order

$$\begin{aligned} & O \left( \sum_{j=1}^{\log^2 n} \sum_{h \neq j} |\mathcal{K}_n(\lambda_{j-h})| + \sum_{j=\log^2 n+1}^{n-1} \sum_{h=1}^{\log^2 n} |\mathcal{K}_n(\lambda_{j-h})| \right) \\ &= O \left( \sum_{j=1}^{\log^2 n} \sum_{h \neq j} \lambda_{j-h}^{-1} + \sum_{j=\log^2 n+1}^{n-1} \sum_{h=1}^{\log^2 n} \lambda_{j-h}^{-1} \right) \\ &= O(n \log^3 n). \end{aligned}$$

To verify the third term is  $o(np_n^{1/2})$ , we will show that

$$E \left| \sum_{j=\log^2 n+1}^{n-1} \sum_{j \neq h=\log^2 n+1}^{n-1} R(\lambda_j) I_\varepsilon(\lambda_h) \mathcal{K}_n(\lambda_{j-h}) \right|^2 = o(n^2 p_n).$$

By Assumption 3, lemma 4 and (32),

$$\begin{aligned} & E \left| \sum_{j=\log^2 n+1}^{n-1} \sum_{j \neq h=\log^2 n+1}^{n-1} R(\lambda_j) I_\varepsilon(\lambda_h) \mathcal{K}_n(\lambda_{j-h}) \right|^2 \\ &= 2 \sum_{h_1=\log^2 n+1}^{n-1} \sum_{j_1=\log^2 n+1}^{h_1-1} \sum_{h_2=\log^2 n+1}^{n-1} \sum_{j_2=\log^2 n+1}^{h_2-1} E(R(\lambda_{j_1}) I_\varepsilon(\lambda_{h_1}) R(\lambda_{j_2}) I_\varepsilon(\lambda_{h_2})) \\ & \quad \times \mathcal{K}_n(\lambda_{j_1-h_1}) \mathcal{K}_n(-\lambda_{j_2-h_2}) \\ & + 2 \sum_{h_1=\log^2 n+1}^{n-1} \sum_{j_1=h_1+1}^{n-1} \sum_{h_2=\log^2 n+1}^{n-1} \sum_{j_2=h_2+1}^{n-1} E(R(\lambda_{j_1}) I_\varepsilon(\lambda_{h_1}) R(\lambda_{j_2}) I_\varepsilon(\lambda_{h_2})) \\ & \quad \times \mathcal{K}_n(\lambda_{j_1-h_1}) \mathcal{K}_n(-\lambda_{j_2-h_2}) \\ &= O \left( \sum_{h_1=\log^2 n+1}^{n-1} \sum_{j_1=\log^2 n+1}^{h_1} \sum_{h_2=\log^2 n+1}^{n-1} \sum_{j_2=\log^2 n+1}^{h_2} j_1^{-d} h_1^{d-1} j_2^{-d} h_2^{d-1} \log h_1 \log h_2 \lambda_{j_1-h_1}^{-1} \lambda_{j_2-h_2}^{-1} \right) \\ &= O(n^2 \log^6 n). \end{aligned}$$

Thus (30) is proved. The proof of (31) is similar to that of (30).

**Lemma 2** Under Assumptions 1a, 2a and 3a,

$$\frac{n}{p_n^{1/2}} \left\{ \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \widehat{f}_{\varepsilon,d}^2(\lambda_\ell) - \frac{2\pi}{\sigma^4} \sum_{j=-(n-1)}^{n-1} k^2(j/p_n) \widehat{\gamma}_{\varepsilon,j}^2 \right\} = o_p(1).$$

**Proof.** Since  $I_\varepsilon(\lambda_j) = I_{n,\varepsilon}(\lambda_j)$  and  $I_{n,\varepsilon}(0) = 0$ , we have

$$I_\varepsilon(\lambda_j) = \frac{1}{2\pi} \sum_{h=-(n-1)}^{n-1} \widehat{\gamma}_{\varepsilon,h} e^{-i\lambda_j h}, \text{ for } j = 1, \dots, (n-1).$$

Now

$$\begin{aligned} & \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \widehat{f}_{\varepsilon,d}^2(\lambda_\ell) \\ &= \left(\frac{2\pi}{n}\right)^3 \frac{1}{4\pi^2 \sigma^4} \sum_{\ell, j_1, j_2=0}^{n-1} \sum_{p_1, p_2, h_1, h_2=-(n-1)}^{n-1} k_{p_1} k_{p_2} \widehat{\gamma}_{\varepsilon, h_1} \widehat{\gamma}_{\varepsilon, h_2} e^{-i\lambda_{\ell-j_1} p_1} e^{i\lambda_{\ell-j_2} p_2} e^{-i\lambda_{j_1} h_1} e^{i\lambda_{j_2} h_2} \\ &= \frac{2\pi}{n^3 \sigma^4} \sum_{p_1, p_2, h_1, h_2=-(n-1)}^{n-1} k_{p_1} k_{p_2} \widehat{\gamma}_{\varepsilon, h_1} \widehat{\gamma}_{\varepsilon, h_2} \sum_{\ell=0}^{n-1} e^{-i\lambda_{p_1-p_2} \ell} \sum_{j_1=0}^{n-1} e^{-i\lambda_{h_1-p_1} j_1} \sum_{j_2=0}^{n-1} e^{-i\lambda_{h_2-p_2} j_2} \\ &= \frac{2\pi}{\sigma^4} \sum_{p=-(n-1)}^{n-1} k_p^2 \left( \widehat{\gamma}_{\varepsilon,p}^2 + \widehat{\gamma}_{\varepsilon,p} \widehat{\gamma}_{\varepsilon, n-|p|} \right) + \frac{4\pi}{\sigma^4} \sum_{p=1}^{n-1} k_p k_{n-p} \left( \widehat{\gamma}_{\varepsilon,p}^2 + \widehat{\gamma}_{\varepsilon,p} \widehat{\gamma}_{\varepsilon, n-p} \right). \end{aligned}$$

Hence, to show Lemma 2, it is sufficient to prove that

$$\sum_{p=-(n-1)}^{n-1} k_p k_{n-|p|} \widehat{\gamma}_{\varepsilon,p}^2 = o_p(n^{-1} p_n^{1/2}) \quad (33)$$

and

$$\sum_{p=-(n-1)}^{n-1} \left( k_p^2 + k_p k_{n-|p|} \right) \widehat{\gamma}_{\varepsilon,p} \widehat{\gamma}_{\varepsilon, n-|p|} = o_p(n^{-1} p_n^{1/2}). \quad (34)$$

In the steps which follow next, we will assume that  $k$  has unbounded support. If  $k$  has bounded support, all terms involving  $k_p k_{n-|p|}$  are zero in both (33) and (34) and the proof is extremely simple. By Assumptions 2a and 3,

$$\begin{aligned} E \left| \frac{n}{p_n^{1/2}} \sum_{p=1}^{n-1} k_p k_{n-p} \widehat{\gamma}_{\varepsilon,p}^2 \right| &\leq \frac{n}{p_n^{1/2}} \sum_{p=1}^{n-1} |k_p k_{n-p}| \frac{n-p}{n^2} \\ &= \frac{n}{p_n^{1/2}} \sum_{p=1}^{n/2} |k_p k_{n-p}| \left( \frac{n-p}{n^2} + \frac{p}{n^2} \right) = \frac{1}{p_n^{1/2}} \sum_{p=1}^{n/2} |k_p| \frac{p_n^\delta}{(n-p)^\delta} \\ &= O \left( \frac{p_n^{\delta+1}}{n^\delta p_n^{1/2} p_n} \sum_{p=1}^{n/2} |k_p| \right) = o(1) \end{aligned}$$

since  $p_n^{-1} \sum_{p=1}^{n/2} |k_p| = O(1)$ . We now verify equation (34).

$$E \left( \sum_{p=1}^{n-1} \left( k_p^2 + k_p k_{n-|p|} \right) \widehat{\gamma}_{\varepsilon,p} \widehat{\gamma}_{\varepsilon, n-p} \right)^2$$

$$\begin{aligned}
&= \sum_p \left( k_p^2 + k_p k_{n-|p|} \right)^2 E \left( \hat{\gamma}_{\varepsilon,p}^2 \hat{\gamma}_{\varepsilon,n-p}^2 \right) \\
&\quad + \sum_{p \neq q} \left( k_p^2 + k_p k_{n-|p|} \right) \left( k_q^2 + k_q k_{n-|q|} \right) E \left( \hat{\gamma}_{\varepsilon,p} \hat{\gamma}_{\varepsilon,n-p} \hat{\gamma}_{\varepsilon,q} \hat{\gamma}_{\varepsilon,n-q} \right). \quad (35)
\end{aligned}$$

By Lemma 1 on page 186 of Grenander and Rosenblatt (1957),  $E \left( \hat{\gamma}_{\varepsilon,p}^2 \hat{\gamma}_{\varepsilon,n-p}^2 \right) = O(n^{-2})$  while  $E \left( \hat{\gamma}_{\varepsilon,p} \hat{\gamma}_{\varepsilon,n-p} \hat{\gamma}_{\varepsilon,q} \hat{\gamma}_{\varepsilon,n-q} \right) = O(n^{-3})$ . Hence, by Assumption 2a, the first term of (35) is

$$O \left( \sum_p \left( k_p^2 + k_p k_{n-|p|} \right)^2 \frac{1}{n^2} \right) = O \left( \frac{p_n}{n^2} \right),$$

while the second term of (35) is

$$O_p \left( \frac{1}{n^3} \sum_{p \neq q} \left( k_p^2 + k_p k_{n-|p|} \right) \left( k_q^2 + k_q k_{n-|q|} \right) \right) = O_p \left( \frac{p_n^2}{n^3} \right).$$

and the lemma is established.

**Lemma 3** *Under the assumptions in Theorem 2,*

$$\left\{ \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \tilde{f}_\varepsilon(\lambda_\ell) \right\}^2 - \left( \frac{2\pi}{\sigma^2} \right)^2 \hat{\gamma}_{\varepsilon,0}^2 = O_p(n^{-1} \log^2 n) \quad (36)$$

and

$$\left\{ \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \hat{f}_{\varepsilon,d}(\lambda_\ell) \right\}^2 = \hat{\gamma}_{\varepsilon,0}^2.$$

**Proof.** The proof of the second claim of the lemma is contained in the proof of the first claim, which we show below. By (8),

$$\begin{aligned}
\frac{2\pi}{n} \sum_{\ell=0}^{n-1} \tilde{f}_\varepsilon(\lambda_\ell) &= \frac{2\pi}{n} \sum_{\ell=0}^{n-1} \frac{2\pi}{n} \sum_{j=1}^{n-1} W(\lambda_{\ell-j}) \left( \frac{2\pi}{\sigma^2} I_\varepsilon(\lambda_j) + R(\lambda_j) \right) \\
&= \frac{2\pi}{n^2} \sum_{j=1}^{n-1} \sum_{q=-(n-1)}^{n-1} k_q e^{-i\lambda_j q} \left( \frac{2\pi}{\sigma^2} I_\varepsilon(\lambda_j) + R(\lambda_j) \right) \sum_{\ell=0}^{n-1} e^{i\lambda_\ell q} \\
&= \frac{2\pi}{n} \sum_{j=1}^{n-1} \frac{2\pi}{\sigma^2} I_\varepsilon(\lambda_j) + \frac{2\pi}{n} \sum_{j=1}^{n-1} R(\lambda_j).
\end{aligned}$$

Let  $I_{n,\varepsilon}$  be the mean corrected periodogram of  $\varepsilon_t$ . Then  $I_\varepsilon(\lambda_j) = I_{n,\varepsilon}(\lambda_j) = \frac{1}{2\pi} \sum \hat{\gamma}_{\varepsilon,h} e^{-i\lambda_j h}$  and  $I_{n,\varepsilon}(0) = 0$ . We have the first term of the last line,

$$\frac{2\pi}{n} \frac{2\pi}{\sigma^2} \sum_{j=0}^{n-1} I_{n,\varepsilon}(\lambda_j) = \frac{2\pi}{n} \frac{2\pi}{\sigma^2} \frac{1}{2\pi} \sum_{j=0}^{n-1} \sum_{h=-(n-1)}^{n-1} \hat{\gamma}_{\varepsilon,h} e^{-i\lambda_j h}$$

$$\begin{aligned}
&= \frac{1}{n} \frac{2\pi}{\sigma^2} \sum_{h=-(n-1)}^{n-1} \widehat{\gamma}_{\varepsilon,h} \sum_{j=0}^{n-1} e^{-i\lambda_j h} \\
&= \frac{2\pi}{\sigma^2} \widehat{\gamma}_{\varepsilon,0}.
\end{aligned}$$

Thus, the LHS of (36) is

$$2 \cdot \frac{2\pi}{\sigma^2} \widehat{\gamma}_{\varepsilon,0} \left( \frac{2\pi}{n} \sum_{j=1}^{n-1} R(\lambda_j) \right) + \left( \frac{2\pi}{n} \sum_{j=1}^{n-1} R(\lambda_j) \right)^2. \quad (37)$$

We will show that the second term is  $O_p(n^{-2} \log^4 n)$ . It follows by Chebyshev's inequality and the fact that  $\widehat{\gamma}_{\varepsilon,0} = O_p(1)$  that the first term is  $O_p(n^{-1} \log^2 n)$ . Now

$$E \left( \sum_{j=1}^{n-1} R(\lambda_j) \right)^2 = E \left( \sum_{j=1}^{\log^2 n} R(\lambda_j) \right)^2 + 2E \left( \sum_{j=1}^{\log^2 n} \sum_{h=\log^2 n}^{n-1} R(\lambda_j) R(\lambda_h) \right) + E \left( \sum_{j=\log^2 n}^{n-1} R(\lambda_j) \right)^2.$$

By lemma 5, the first term is  $O(\log^2 n)$ , the second term is  $O(\log^4 n)$  while the third term is  $O(\log^4 n)$  and hence (37) is  $O_p(n^{-1} \log^2 n)$ .

**Lemma 4** *Under the assumptions in Theorem 2,*

$$E [R(\lambda_j) I_\varepsilon(\lambda_h) R(\lambda_k) I_\varepsilon(\lambda_\ell)] = O(j^{-d} h^{d-1} k^{-d} \ell^{d-1} \log h \log \ell) \quad (38)$$

and

$$E [R(\lambda_j) R(\lambda_h) R(\lambda_k) R(\lambda_\ell)] = O(j^{-d} h^{d-1} k^{-d} \ell^{d-1} \log h \log \ell) \quad (39)$$

uniformly for  $\log^2 n \leq j < h \leq n$ ,  $\log^2 n \leq k < \ell \leq n$ .

**Proof.** The development of this proof closely matches that of Lemma 2 of Hurvich, Deo and Brodsky (1998). We'll use the following notations,

$$I_j = I(\lambda_j), \quad f_j = f(\lambda_j) \quad \text{and} \quad I_{\varepsilon j} = I_\varepsilon(\lambda_j).$$

The LHS of (38) is

$$E \left[ \left( \frac{I_j}{f_j} - 2\pi\sigma^{-2} I_{\varepsilon j} \right) I_{\varepsilon h} \left( \frac{I_k}{f_k} - 2\pi\sigma^{-2} I_{\varepsilon k} \right) I_{\varepsilon \ell} \right]$$

$$\begin{aligned}
&= E \left[ \left( \frac{I_j}{f_j} - 1 - 2\pi\sigma^{-2}I_{\varepsilon j} + 1 \right) I_{\varepsilon h} \left( \frac{I_k}{f_k} - 1 - 2\pi\sigma^{-2}I_{\varepsilon k} + 1 \right) I_{\varepsilon \ell} \right] \\
&= E \left[ \left( \frac{I_j}{f_j} - 1 \right) I_{\varepsilon h} \left( \frac{I_k}{f_k} - 1 \right) I_{\varepsilon \ell} \right] - E \left[ \left( 2\pi\sigma^{-2}I_{\varepsilon j} - 1 \right) I_{\varepsilon h} \left( \frac{I_k}{f_k} - 1 \right) I_{\varepsilon \ell} \right] \\
&\quad - E \left[ \left( \frac{I_j}{f_j} - 1 \right) I_{\varepsilon h} \left( 2\pi\sigma^{-2}I_{\varepsilon k} - 1 \right) I_{\varepsilon \ell} \right] + E \left[ \left( 2\pi\sigma^{-2}I_{\varepsilon j} - 1 \right) I_{\varepsilon h} \left( 2\pi\sigma^{-2}I_{\varepsilon k} - 1 \right) I_{\varepsilon \ell} \right].
\end{aligned} \tag{40}$$

Note that the last expectation of (40) is zero. Let

$$E \left[ \left( \frac{I_j}{f_j} - 1 \right) I_{\varepsilon h} \left( \frac{I_k}{f_k} - 1 \right) I_{\varepsilon \ell} \right] = E(\zeta_j \zeta_h \zeta_k \zeta_\ell),$$

and

$$\begin{aligned}
\mathbf{v} &= \left( \frac{A_{xj}}{f_j^{1/2}}, \frac{B_{xj}}{f_j^{1/2}}, \frac{\sqrt{2\pi}}{\sigma} A_{\varepsilon h}, \frac{\sqrt{2\pi}}{\sigma} B_{\varepsilon h}, \frac{A_{xk}}{f_k^{1/2}}, \frac{B_{xk}}{f_k^{1/2}}, \frac{\sqrt{2\pi}}{\sigma} A_{\varepsilon \ell}, \frac{\sqrt{2\pi}}{\sigma} B_{\varepsilon \ell} \right)' \\
&= (v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8)'
\end{aligned}$$

where

$$A_{aj} = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n a_t \cos(\lambda_j t), \quad B_{aj} = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n a_t \sin(\lambda_j t).$$

The vector  $\mathbf{v}$  has a eight-dimensional multivariate Gaussian distribution with mean zero and covariance matrix  $\Sigma$ . Define  $\Psi = \Sigma^{-1}$ . Partition  $\Sigma$  and  $\Psi$  as

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \quad \text{and} \quad \Psi = \begin{bmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{bmatrix},$$

where  $\Sigma_{ij}$  and  $\Psi_{ij}$  are  $4 \times 4$  matrices. By the formulas for the inverse of a partitioned matrix,

$$\Psi_{11} = \Sigma_{11}^{-1} + \Sigma_{11}^{-1} \Sigma_{12} \left( \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \right)^{-1} \Sigma_{21} \Sigma_{11}^{-1},$$

$$\Psi_{12} = -\Sigma_{11}^{-1} \Sigma_{12} \left( \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \right)^{-1}$$

and

$$\Psi_{22} = \left( \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \right)^{-1}.$$

Letting  $\mathcal{V}_{aj} = A_{aj}$  or  $B_{aj}$ , we have from Lemma 4 of Moulines and Soulier (1999),

$$E \left( \frac{\mathcal{V}_{xj}}{f_j^{1/2}} \frac{\mathcal{V}_{xk}}{f_k^{1/2}} \right) = O(j^{-d} k^{d-1} \log k). \tag{41}$$

for  $1 \leq j < k \leq n/2$ . Following similar arguments of this lemma, it can be shown that for  $1 \leq j < k \leq n/2$

$$E\left(\frac{\mathcal{V}_{xj}^2}{f_j}\right) = \frac{1}{2} + O(j^{-1} \log j), \quad (42)$$

$$E\left(\frac{\mathcal{V}_{\varepsilon j} \mathcal{V}_{xk}}{f_k^{1/2}}\right) = O(j^{-d} k^{d-1} \log k) \quad \text{and} \quad E\left(\frac{\mathcal{V}_{xj} \mathcal{V}_{\varepsilon k}}{f_j^{1/2}}\right) = O(j^{-d} k^{d-1} \log k). \quad (43)$$

Letting

$$\mathbf{R} = \Sigma - \frac{1}{2} \mathbf{I}_8 = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix}$$

where  $\mathbf{I}_8$  is a  $8 \times 8$  identity matrix, we see from (41), (42) and (43) that  $\mathbf{R} = o(1)$  for  $\log^2 n < j < h \leq n/2$ ,  $\log^2 n < k < \ell \leq n/2$ . By the fact that  $(\mathbf{I} + \mathbf{A})^{-1} = \mathbf{I} - (\mathbf{I} + \mathbf{A})^{-1} \mathbf{A}$ , we get  $\bar{\Psi} = 2\mathbf{I}_8 - 2\mathbf{R}(\mathbf{I}_8 + 2\mathbf{R})^{-1} = O(1)$ . Let

$$\tilde{\Psi} = \begin{bmatrix} \Psi_{11} & 0 \\ 0 & \Psi_{22} \end{bmatrix}$$

and define  $\bar{\Psi} = \Psi - \tilde{\Psi}$ . We have

$$\begin{aligned} E(\zeta_j \zeta_h \zeta_k \zeta_\ell) &= (2\pi)^4 |\Psi|^{1/2} \int \dots \int \zeta_j \zeta_h \zeta_k \zeta_\ell e^{-\frac{1}{2} \mathbf{v}' \Psi \mathbf{v}} d\mathbf{v} \\ &= (2\pi)^4 |\Psi|^{1/2} \int \dots \int \zeta_j \zeta_h \zeta_k \zeta_\ell e^{-\frac{1}{2} \mathbf{v}' \tilde{\Psi} \mathbf{v}} d\mathbf{v} \end{aligned} \quad (44)$$

$$+ (2\pi)^4 |\Psi|^{1/2} \int \dots \int \zeta_j \zeta_h \zeta_k \zeta_\ell e^{-\frac{1}{2} \mathbf{v}' \tilde{\Psi} \mathbf{v}} \left\{ e^{-\frac{1}{2} \mathbf{v}' \bar{\Psi} \mathbf{v}} - 1 \right\} d\mathbf{v} \quad (45)$$

Let  $\mathbf{v}_{(jh)} = (v_1, v_2, v_3, v_4)'$ ,  $\mathbf{v}_{(k\ell)} = (v_5, v_6, v_7, v_8)'$ , the first term of the above equation is

$$(2\pi)^4 |\Psi|^{1/2} \iiint \zeta_j \zeta_h e^{-\frac{1}{2} \mathbf{v}'_{(jh)} \Psi_{11} \mathbf{v}_{(jh)}} d\mathbf{v}_{(jh)} \iiint \zeta_k \zeta_\ell e^{-\frac{1}{2} \mathbf{v}'_{(k\ell)} \Psi_{22} \mathbf{v}_{(k\ell)}} d\mathbf{v}_{(k\ell)}. \quad (46)$$

The first quadruple integral of (46) is

$$\iiint \zeta_j \zeta_h e^{-\frac{1}{2} \mathbf{v}'_{(jh)} \Sigma_{11}^{-1} \mathbf{v}_{(jh)}} e^{-\frac{1}{2} \mathbf{v}'_{(jh)} \mathbf{M}_{11} \mathbf{v}_{(jh)}} d\mathbf{v}_{(jh)} \quad (47)$$

where

$$\mathbf{M}_{11} = \Sigma_{11}^{-1} \Sigma_{12} \left( \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \right)^{-1} \Sigma_{21} \Sigma_{11}^{-1}.$$

Let  $\tau_{11}$  be the largest absolute entry of  $\mathbf{M}_{11}$ . Since  $|e^u - 1| \leq |u| e^{|u|}$  for all  $u$ ,

$$e^{-\frac{1}{2} \mathbf{v}'_{(jh)} \mathbf{M}_{11} \mathbf{v}_{(jh)}} = 1 + O\left\{ \tau_{11} \left\| \mathbf{v}_{(jh)} \right\|^2 e^{\frac{3}{2} \tau_{11} \left\| \mathbf{v}_{(jh)} \right\|^2} \right\}.$$

Thus (47) is equal to

$$\begin{aligned} & \iiint \zeta_j \zeta_h e^{-\frac{1}{2} \mathbf{v}'_{(jh)} \boldsymbol{\Sigma}_{11}^{-1} \mathbf{v}_{(jh)}} d\mathbf{v}_{(jh)} \\ & + O \left\{ \iiint \left| \zeta_j \zeta_h \right| \tau_{11} \left\| \mathbf{v}_{(jh)} \right\|^2 e^{-\frac{1}{2} \mathbf{v}'_{(jh)} (\boldsymbol{\Sigma}_{11}^{-1} - 3\tau_{11} \mathbf{I}_4) \mathbf{v}_{(jh)}} d\mathbf{v}_{(jh)} \right\}. \end{aligned} \quad (48)$$

The second term is  $O(\tau_{11}) = O(j^{-2d} k^{2d-2} \log^2 k \mathbf{1}_{(j < k)} + k^{-2d} j^{2d-2} \log^2 k \mathbf{1}_{(j > k)})$  by (41), (42) and (43). Note that

$$\boldsymbol{\Sigma}_{11}^{-1} = 2\mathbf{I}_4 - 2\mathbf{R}_{11} (\mathbf{I}_4 + 2\mathbf{R}_{11})^{-1} = 2\mathbf{I}_4 + o(1).$$

Let  $\eta_{11}$  be the largest absolute entry of  $2\mathbf{R}_{11} (\mathbf{I}_4 + 2\mathbf{R}_{11})^{-1}$ ,

$$e^{\frac{1}{2} \mathbf{v}'_{(jh)} (2\mathbf{R}_{11} (\mathbf{I}_4 + 2\mathbf{R}_{11})^{-1}) \mathbf{v}_{(jh)}} = 1 + O \left\{ \eta_{11} \left\| \mathbf{v}_{(jh)} \right\|^2 e^{\frac{3}{2} \eta_{11} \left\| \mathbf{v}_{(jh)} \right\|^2} \right\}.$$

Thus the first term of (48) is

$$\begin{aligned} & \iiint \zeta_j \zeta_h e^{-\frac{1}{2} \mathbf{v}'_{(jh)} 2\mathbf{I}_4 \mathbf{v}_{(jh)}} d\mathbf{v}_{(jh)} \\ & + O \left\{ \iiint \left| \zeta_j \zeta_h \right| \eta_{11} \left\| \mathbf{v}_{(jh)} \right\|^2 e^{\frac{1}{2} \mathbf{v}'_{(jh)} ((2-3\eta_{11}) \mathbf{I}_4) \mathbf{v}_{(jh)}} d\mathbf{v}_{(jh)} \right\} \\ = & \iint \zeta_j e^{-\frac{1}{2} \mathbf{v}'_{(j)} 2\mathbf{I}_2 \mathbf{v}_{(j)}} d\mathbf{v}_{(j)} \iint \zeta_j e^{-\frac{1}{2} \mathbf{v}'_{(h)} 2\mathbf{I}_2 \mathbf{v}_{(h)}} d\mathbf{v}_{(h)} \\ & + O \left\{ \iiint \left| \zeta_j \zeta_h \right| \eta_{11} \left\| \mathbf{v}_{(jh)} \right\|^2 e^{\frac{1}{2} \mathbf{v}'_{(jh)} ((2-3\eta_{11}) \mathbf{I}_4) \mathbf{v}_{(jh)}} d\mathbf{v}_{(jh)} \right\}. \end{aligned}$$

The first term of the RHS of the above equation is zero since the first double integral is the expectation of  $\zeta_j$  assuming the covariance matrix is  $0.5\mathbf{I}_4$ . The second term is  $O(\eta_{11}) = O(j^{-d} h^{d-1} \log h)$ . We have shown that the first quadruple integral of (46) is  $O(j^{-d} h^{d-1} \log h + j^{-2d} k^{2d-2} \log^2 k \mathbf{1}_{(j \leq k)} + j^{2d-2} k^{-2d} \log^2 j \mathbf{1}_{(k \leq j)})$ . It can be shown in the same fashion that the second quadruple integral of (46) is  $O(k^{-d} \ell^{d-1} \log \ell + j^{-2d} k^{2d-2} \log^2 k \mathbf{1}_{(j \leq k)} + j^{2d-2} k^{-2d} \log^2 j \mathbf{1}_{(j > k)})$ . Hence (44) is  $O(j^{-d} h^{d-1} k^{-d} \ell^{d-1} \log h \log \ell)$ .

Now we consider (45). By the mean value theorem,  $|e^u - 1 - u| \leq (1/2)u^2 e^{|u|}$  for all  $u$ . Thus

$$e^{-\frac{1}{2} \mathbf{v}' \bar{\boldsymbol{\Psi}} \mathbf{v}} - 1 = -\frac{1}{2} \mathbf{v}' \bar{\boldsymbol{\Psi}} \mathbf{v} + O \left( \tau^2 \left\| \mathbf{v} \right\|^4 e^{2\tau \left\| \mathbf{v} \right\|^2} \right),$$

where  $\tau$  is the largest absolute entry of  $\bar{\boldsymbol{\Psi}}$ . Note that  $\tau^2 = O(j^{-2d} k^{2d-2} \log^2 k \mathbf{1}_{(j \leq k)} + j^{2d-2} k^{-2d} \log^2 j \mathbf{1}_{(k \leq j)})$ .

Hence (45) is

$$\begin{aligned} & (2\pi)^4 |\boldsymbol{\Psi}|^{1/2} \int \cdots \int \zeta_j \zeta_h \zeta_k \zeta_\ell - \frac{1}{2} \mathbf{v}' \bar{\boldsymbol{\Psi}} \mathbf{v} e^{-\frac{1}{2} \mathbf{v}' \bar{\boldsymbol{\Psi}} \mathbf{v}} d\mathbf{v} \\ & + O \left\{ \tau^2 \int \cdots \int \left| \zeta_j \zeta_h \zeta_k \zeta_\ell \right| \left\| \mathbf{v} \right\|^4 e^{-\frac{1}{2} \mathbf{v}' (\bar{\boldsymbol{\Psi}} - 4\tau \mathbf{I}_8) \mathbf{v}} d\mathbf{v} \right\}. \end{aligned}$$



The second term is  $O(\tau^2)$ . The first term is the linear combination of  $E_{\tilde{\Psi}}[\zeta_j \zeta_h \zeta_k \zeta_\ell A_j A_k]$ ,  $E_{\tilde{\Psi}}[\zeta_j \zeta_h \zeta_k \zeta_\ell A_j B_k]$ ,  $E_{\tilde{\Psi}}[\zeta_j \zeta_h \zeta_k \zeta_\ell A_j A_\ell]$ ,  $E_{\tilde{\Psi}}[\zeta_j \zeta_h \zeta_k \zeta_\ell A_j B_\ell]$ , ..., etc., where  $E_{\tilde{\Psi}}$  denotes the expectation assuming that  $\mathbf{v}$  is multivariate normal with mean zero and covariance matrix  $\tilde{\Psi}$ . Since  $\text{cov}(\mathbf{v}) = \tilde{\Psi}$  implies that the vectors  $(A_j, B_j, A_h, B_h)$ ,  $(A_k, B_k, A_\ell, B_\ell)$  are independent. Thus, for example,  $E_{\tilde{\Psi}}[\zeta_j \zeta_h \zeta_k \zeta_\ell A_j A_k] = E_{\tilde{\Psi}}[\zeta_j \zeta_h A_j] E_{\tilde{\Psi}}[\zeta_k \zeta_\ell A_k]$ , and both of these expectations are zero because the  $\zeta_j \zeta_h$  and  $\zeta_k \zeta_\ell$  are even functions of  $(A_j, B_j, A_h, B_h)$  respectively, and because the densities for  $(A_j, B_j, A_h, B_h)$  and  $(A_k, B_k, A_\ell, B_\ell)$  are also even functions. We have shown that (45) is  $O(\tau^2) = O(j^{-2d} k^{2d-2} \log^2 k \mathbf{1}_{(j \leq k)} + j^{2d-2} k^{-2d} \log^2 j \mathbf{1}_{(k \leq j)})$ . Hence

$$E \left[ \left( \frac{I_j}{f_j} - 1 \right) I_{\varepsilon h} \left( \frac{I_k}{f_k} - 1 \right) I_{\varepsilon \ell} \right] = O(j^{-d} h^{d-1} k^{-d} \ell^{d-1} \log h \log \ell).$$

It can be shown in a similar way that the rest of the second and the third expectations of (40) are both  $O(j^{-d} h^{d-1} k^{-d} \ell^{d-1} \log h \log \ell)$  uniformly in  $\log^2 n \leq j < h \leq n/2$ ,  $\log^2 n \leq k < \ell \leq n/2$ . The order in (39) can be derived following the same lines as above.

**Lemma 5** *Under the assumptions of Theorem 2,*

$$E[R(\lambda_j)R(\lambda_h)] = O(j^{-1}h^{-1} \log h \log j + j^{-2d}h^{2d-2} \log^2 h)$$

and

$$E[R^2(\lambda_j)] = O(j^{-1} \log j).$$

uniformly for  $\log^2 n \leq j < h \leq n$ . Also  $\max_{1 \leq j \leq n} E[R^2(\lambda_j)] < \infty$ .

The proof of the first two bounds stated in this lemma is similar to that of lemma 4. The last bound is obtained by using the bounds (41), (42) and (43) and the Gaussianity of the observations.

**Lemma 6** *Let  $g(\lambda)$  be defined as (17). Then, under assumption 7,*

$$\begin{aligned} \sum_{h=1}^{n-1} g(\lambda_h) e^{-i\lambda_h m} &= O(n^\delta) \text{ if } m = 0, \\ &= O(n^{1+\delta} m^{-1}) \text{ if } m \neq 0. \end{aligned}$$

**Proof.** We'll prove the lemma by showing that

$$\begin{aligned} \frac{2\pi}{n} \sum_{h=1}^{n-1} g(\lambda_h) e^{-i\lambda_h m} &= O(n^{-1+\delta}) \text{ if } m = 0, \\ &= O(n^\delta m^{-1}) \text{ if } m \neq 0 \text{ and } |m| \leq n \end{aligned} \quad (49)$$

We first derive the result for  $m = 0$ . Note that

$$\int_0^{2\pi} g(\lambda) d\lambda = 0.$$

Hence, the LHS of (49) is

$$\begin{aligned} \frac{2\pi}{n} \sum_{h=1}^{n-1} g(\lambda_h) &= \left( \frac{2\pi}{n} \sum_{h=1}^{n-1} g(\lambda_h) - \int_0^{2\pi} g(\lambda) d\lambda \right) \\ &= \sum_{h=1}^{n-1} \int_{\lambda_{h-1}}^{\lambda_h} (g(\lambda_h) - g(\lambda)) d\lambda - \int_{\lambda_{n-1}}^{\lambda_n} g(\lambda) d\lambda \\ &= \sum_{h=1}^{n-1} g'(\lambda_{\tilde{h}}) \int_{\lambda_{h-1}}^{\lambda_h} (\lambda_h - \lambda) d\lambda - \int_{\lambda_0}^{\lambda_1} g(\lambda) d\lambda, \end{aligned}$$

where  $\lambda_{h-1} < \lambda_{\tilde{h}} < \lambda_h$  and we use the fact that  $g(\lambda)$  is symmetric around  $\pi/2$ . By assumption 7, the last equation above is

$$O\left(\sum_{h=1}^n \lambda_h^{-1-\delta} \cdot \frac{1}{2} \left(\frac{2\pi}{n}\right)^2 + \lambda_1^{1-\delta}\right) = O(n^{-1+\delta}).$$

For  $m \neq 0$ , we have by summation by parts

$$\begin{aligned} &\frac{2\pi}{n} \sum_{h=1}^{n-1} g(\lambda_h) e^{i\lambda_h m} \\ &= \frac{2\pi}{n} \sum_{h=1}^{n-2} (g(\lambda_h) - g(\lambda_{h+1})) \sum_{\ell=1}^h e^{i\lambda_\ell m} + \frac{2\pi}{n} g(\lambda_{n-1}) \sum_{\ell=1}^{n-1} e^{i\lambda_\ell m} \\ &= \frac{2\pi}{n} \sum_{h=1}^{n-2} g'(\lambda_{\tilde{h}}) (\lambda_h - \lambda_{h+1}) \sum_{\ell=1}^h e^{i\lambda_\ell m} + \frac{2\pi}{n} g(\lambda_{n-1}) (-1). \end{aligned}$$

Because  $\sum_{\ell=a}^b e^{i\lambda_\ell} = O(\lambda^{-1})$  uniformly in  $a$  and  $b$  for  $0 < \lambda < \pi$  (see Zygmund, 1977, p. 2), this is

$$O\left(\frac{1}{n} \sum_{h=1}^{n-2} \lambda_h^{-1-\delta} \lambda_1 \lambda_m^{-1} + \frac{1}{n^{1-\delta}}\right) = O\left(\frac{n^\delta}{m}\right).$$

## REFERENCES

- [1] Beran, J. (1992) “A Goodness-of-fit test for time series with long range dependence”, *J. R. Statist. Soc. B*, **54**, 749-60.
- [2] Billingsley, P. (1986), “Probability and Measure”, 2nd ed., John Wiley.
- [3] Box, G. E. P., and Pierce, D. A. (1970), “Distribution of residual autocorrelations in autoregressive-integrated moving average time series models”, *Journal of American Statistical Association*, **65**, 1509-26.
- [4] Dahlhaus, R. (1989), “Efficient parameter estimation for self-similar processes”, *Ann. Stat.*, **17**, 1749-1766.
- [5] Davies, R. B. and Harte, D. S. (1987), “Tests for Hurst effect”, *Biometrika*, **74**, 95-102.
- [6] Deo, R. S. and Chen, W. W. (2000), “On the Integral of the Squared Periodogram”, *Stochastic Processes and Their Applications*, **85**, 159-76.
- [7] Granger, C. W. J. and Joyeux, R. (1980), “An introduction to long memory time series models and fractional differencing.”, *Journal of Time Series Analysis*, **1**, 15-29.
- [8] Grenander, U. and Rosenblatt, M. (1957), *Statistical Analysis of Stationary Time Series*, John Wiley.
- [9] Hong, Y. (1996), “Consistent testing for serial correlation of unknown form”, *Econometrica*, **64**, 837-64.
- [10] Hosking, J. R. M. (1981), “Fractional differencing ”, *Biometrika*, **68**, 165-76.
- [11] Hurvich, C. M., Deo, R. S. and Brodsky, J. (1998), “The mean squared error of Geweke and Porter-Hudak’s estimator of the memory parameter of a long-memory time series”, *Journal of Time Series Analysis*, **19**, 19-46.
- [12] Ljung, G. M. and Box, G. E. P. (1978), “On a measure of lack of fit in time series models”, *Biometrika*, **65**, 297-303.
- [13] Moulines, E. and Soulier, P. (1999), “Broadband log-periodogram regression of time series with long-range dependence”, *Ann. Stat.*, **27**, 1415-39.
- [14] Mandelbrot, B. B. and Van Ness, J. W. (1968) “Fractional Brownian motions, Fractional noises and applications”, *SIAM Review*, **10**, 422-437.
- [15] Milhoj, A. (1981), “A test of fit in time series models”, *Biometrika*, **68**, 177-87.

[16] Zygmund, A. (1977), "Trigonometric Series". Cambridge University press.

**TABLE I**  
Rejection Rates in Percentage Under an AR(1) Model

$n$	$p_n$	128												512			
		8		13		21		11		20		37					
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%		
$T_n$	BAR	3.08	5.02	4.04	6.12	4.90	7.80	3.82	5.82	4.32	6.98	5.06	8.02				
	TUK	3.04	4.96	4.04	6.12	6.30	9.68	3.98	5.82	4.56	7.10	5.16	8.40				
	QS	3.64	5.64	4.52	6.90	5.04	7.74	4.06	6.52	4.74	7.64	5.58	9.26				
$H_n$	BAR	3.30	5.08	3.82	5.82	4.26	6.76	3.62	5.72	4.20	6.54	4.76	7.34				
	TUK	3.16	4.90	3.78	5.92	4.46	6.96	3.76	5.78	4.26	6.84	4.88	7.48				
	QS	3.52	5.52	4.22	6.44	4.82	7.40	4.02	6.20	4.36	7.12	5.08	8.36				
$M_n$		4.34 at 5%		7.12 at 10%		5.14 at 5%		8.88 at 10%									

Note: Model  $x_t - 0.8x_{t-1} = e_t$ .

**TABLE II**  
Rejection Rates in Percentage Under an ARFIMA (0,d,0) Model

$n$	$p_n$	128												512			
		8		13		21		11		20		37					
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%		
$T_n$	BAR	2.62	4.08	3.52	5.28	4.90	7.42	2.64	4.60	3.74	6.00	4.80	8.20				
	TUK	2.52	4.00	3.46	5.58	4.96	7.50	2.92	4.78	3.86	6.14	5.10	8.42				
	QS	3.22	4.98	4.34	6.78	6.62	9.60	3.30	5.74	4.40	7.06	5.58	9.08				
$H_n$	BAR	2.28	3.76	3.02	4.86	3.54	5.88	2.56	4.42	3.42	5.86	4.22	7.00				
	TUK	2.20	3.52	3.20	5.10	3.90	5.88	3.12	5.32	4.14	6.52	4.70	7.86				
	QS	2.82	4.46	3.66	5.36	4.10	7.04	2.72	4.54	3.70	5.98	4.44	7.44				
$M_n$		4.70 at 5%		7.58 at 10%		4.50 at 5%		8.18 at 10%									

Note: Model  $x_t = \text{ARFIMA}(0,d,0)$  with  $d = 0.3$ .

**TABLE III**  
Rejections Rates in Percentage under AR(2) Alternative fitting Model AR(1)

$n$	$p_n$	128												512			
		8		13		21		11		20		37					
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%		
$T_n$	BAR	22.48	28.60	22.92	29.04	22.88	29.62	80.18	84.96	76.02	81.64	68.54	75.74				
	TUK	21.94	28.16	22.80	28.58	22.40	28.76	79.96	82.18	74.76	81.06	65.66	73.26				
	QS	22.44	28.64	22.80	28.96	22.74	29.74	78.20	83.42	70.56	77.68	61.04	69.96				
$H_n$	BAR	23.58	30.22	23.42	29.66	22.42	28.70	80.62	85.46	75.84	81.86	68.24	75.26				
	TUK	23.18	29.36	23.22	29.42	21.98	28.22	80.32	85.24	74.90	80.90	65.02	72.66				
	QS	23.28	29.90	22.76	28.54	21.12	27.40	78.34	83.96	70.46	77.14	59.60	68.88				
$M_n$		8.84 at 5%		13.78 at 10%		17.78 at 5%		25.96 at 10%									

Note: Model  $x_t - 0.8x_{t-1} + 0.15x_{t-2} = u_t$ .

**TABLE IV**  
Rejections Rates in Percentage Under ARMA(1,1) Alternative  
fitting Model ARIMA(1,d,0)

$n$	$p_n$	128								512			
		8		13		21		11		20		37	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$T_n$	BAR	9.50	13.38	8.44	12.58	8.96	13.04	31.84	40.74	28.36	36.56	25.10	34.24
	TUK	7.24	11.28	8.04	12.06	8.80	12.80	31.34	40.54	26.94	35.12	23.36	32.48
	QS	8.74	12.26	8.04	12.02	10.04	14.74	29.20	37.80	25.00	33.62	22.78	31.50
$H_n$	BAR	12.68	17.04	11.28	15.52	8.92	13.20	33.02	42.28	28.94	37.36	24.82	33.70
	TUK	7.98	12.12	8.20	12.32	8.66	12.96	32.66	41.88	27.70	35.72	23.12	31.88
	QS	11.48	15.70	10.18	14.40	9.06	13.68	30.32	38.70	25.30	33.72	21.78	30.10
$M_n$		5.42 at 5%		8.76 at 10%		6.44 at 5%		10.38 at 10%					

Note: Model  $x_t = 0.8x_{t-1} + u_t + 0.2u_{t-1}$ .

**TABLE V**  
Rejections Rates in Percentage Under ARFIMA(0,d,0) Alternative  
fitting Model ARMA(1,1)

$n$	$p_n$	128								512			
		8		13		21		11		20		37	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$T_n$	BAR	6.46	9.38	7.50	11.10	8.90	13.30	37.28	44.78	38.54	46.16	37.02	45.06
	TUK	6.54	9.10	7.50	11.32	8.68	13.14	38.06	45.20	39.14	46.74	36.20	44.42
	QS	7.20	10.22	8.24	12.56	10.44	15.34	39.70	46.54	37.90	45.72	34.36	43.00
$H_n$	BAR	5.26	7.54	6.22	8.90	6.84	10.34	36.14	43.38	37.28	44.80	35.00	42.88
	TUK	5.32	7.48	6.46	9.12	6.92	10.54	37.16	44.08	37.90	45.46	34.22	42.36
	QS	6.04	8.64	6.68	10.18	7.32	10.78	38.42	45.56	36.72	44.56	32.26	40.10
$M_n$		5.34 at 5%		8.92 at 10%		11.56 at 5%		17.96 at 10%					

Note: Model  $x_t = \text{ARFIMA}(0,d,0)$  with  $d = 0.4$ .

**TABLE VI**  
Rejections Rates in Percentage Under ARFIMA(1,d,0) Alternative  
fitting Model ARFIMA(0,d,0)

$n$	$p_n$	128								512			
		8		13		21		11		20		37	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$T_n$	BAR	8.52	12.48	8.76	12.68	9.68	14.16	16.92	22.42	14.94	21.14	13.32	19.42
	TUK	8.16	12.10	8.10	12.14	9.10	13.60	16.26	21.78	14.50	20.50	12.80	18.32
	QS	8.24	11.74	8.82	12.86	10.88	15.54	15.76	21.34	13.16	19.22	12.62	18.24
$H_n$	BAR	7.54	10.84	7.54	11.42	7.98	11.56	16.22	21.78	14.22	20.14	12.32	17.74
	TUK	7.36	10.68	7.26	11.06	7.60	11.36	15.28	20.38	12.52	18.14	10.88	16.32
	QS	7.32	10.70	7.32	11.34	8.12	11.53	15.84	20.98	13.88	19.68	11.60	16.82
$M_n$		6.14 at 5%		9.92 at 10%		6.82 at 5%		11.40 at 10%					

Note: Model  $x_t - 0.1x_{t-1} = u_t$ ,  $u_t = \text{ARFIMA}(0,d,0)$  with  $d = 0.4$ .