A test for second order stationarity of a time series based on the Discrete Fourier Transform

Yogesh Dwivedi and Suhasini Subba Rao
Texas A&M University,
College Station, Texas 77845, USA

February 22, 2010

Abstract

We consider a zero mean discrete time series, and define its discrete Fourier transform at the canonical frequencies. It can be shown that the discrete Fourier transform is asymptotically uncorrelated at the canonical frequencies if and if only the time series is second order stationary. Exploiting this important property, we construct a Portmanteau type test statistic for testing stationarity of the time series. It is shown that under the null of stationarity, the test statistic is approximately a chi square distribution. To examine the power of the test statistic, the asymptotic distribution under the locally stationary alternative is established. It is shown to be a generalised noncentral chi-square, where the noncentrality parameter measures the deviation from stationarity. The test is illustrated with simulations, where is it shown to have good power.

Keywords and phrases α-mixing, Discrete Fourier Transform, linear time series, local stationarity, Portmanteau test, test for second order stationarity.

Primary 62M10, **Secondary** 62F10

1 Introduction

An important assumption that is often made when analysing time series is that it is at least second order stationary. A large proportion of the time series literature is based on this assumption. If the assumption is not properly tested and the analysis is performed, the resulting model may be misspecified and the forecasts obtained may be inappropriate. Therefore, it is important to check whether the time series is second order stationary.

Over the years, various statistical tests have been proposed. One of the first tests for stationarity is considered in Priestley and Subba Rao (1969), more recently tests have been proposed in Sachs and

Neumann (1999) and Paparoditis (2009). We briefly mention, that there exists several CUMSUM based tests (see Andreou and Ghysels (2008) for a review), which are specifically designed for the detection of change points in a time series, which is a particular example of nonstationarity. In contrast, the aforementioned stationarity tests are designed for the detection of smoothly varying alternatives (see Sachs and Neumann (1999) for an interesting discussion on the differences between smooth and non-smooth tests). Most of the tests for stationarity are based on comparing spectral densities over various segments. More precisely, the test statistic proposed in Paparoditis (2009) is an L_2-based statistic, which compares the local spectral density estimator with the global spectral density estimator. Whereas, Sachs and Neumann (1999) propose a test which detects for a changes in the autocovariance structure from segment to segment, and is based on the observation that under the null of stationarity the bivariate wavelets coefficient of the local spectral density is zero. Motivated by this observation, Sachs and Neumann (1999) propose estimators of the Haar wavelet coefficients and use a multiple testing procedure to test the significance of the wavelet coefficients.

The underlying important assumption, on which these tests for stationarity are based, is on a delicate, subjective, choice of segments of the data. This can make the test extremely sensitive to the segment length. Moreover, the rate of convergence to the limiting distribution (usually a Gaussian) depends on the number of segments. Therefore, for relatively small sample sizes the normal approximation may not be reliable. For example, the L_2-statistic in Paparoditis (2009), like most L_2-tests, can be quite skewed, which can lead to large type I errors. On the other hand, the rate of convergence of the Sachs and Neumann (1999) test statistic depends, implicitly, on the segment length. Furthermore, since the joint distribution of the wavelet coefficients does not exist, obtaining the power of the Sachs and Neumann (1999) test for the alternative of local stationarity can be extremely difficult. In this paper, we propose a test based on the discrete Fourier transforms, which is based on the entire length of data, thus avoiding a subjective choice of segment length and its associated problems. Unlike most tests for stationarity, which are comparison based, the proposed test is motivated by a property unique to second order stationary time series.

In Section 2 we define the Discrete Fourier transform (DFT) and show that the DFT are asymptotically uncorrelated at the canonical frequencies if and only if the time series is second order stationary. This motivates the test statistics, which is based on the DFT. The Portmanteau type test statistics we propose is based on the covariance function calculated using the DFT at the canonical frequencies, which we call the DFT covariance. The asymptotic sampling distribution of the test statistic, under the assumption that the time series is strictly stationary and either linear or α-mixing is obtained in Section 3. Under the null hypothesis that the time series is strictly stationary, we show that the asymptotic sampling distribution of the test statistic is a central chi-square. To examine the power of the test, we consider the case of locally stationary time series (see Dahlhaus (1997) and Dahlhaus and Polonik (2006)), and derive the distribution of the test statistic under this class of alternatives. In Section 4 we show the distribution under this class of alternatives, is a type of non-central chi-square, where the noncentrality parameter is of order $O(T)$. Furthermore, we show that the case of local stationarity, the DFT covariance is an estimator of the Fourier coefficients of the locally stationary spectral density. One practical advantage of our approach is that DFT covariance can be plotted, which can be used to indicate where the departures from stationarity may lie.

In Section 5, we compare our test to the test for stationarity recently proposed in Paparoditis (2009). In Section 6 we compare the two tests through simulations and we examine the type I error and power of both tests. We show that the type I error for the our test is good and the power for various alternatives is high.

The proofs can be found in the appendix, and some results concerning quadratic forms may be of independent interest.

2 The Test Statistic, motivation and sampling distribution

2.1 Motivation

Let $\{X_t\}$ be a zero mean time series. Suppose we observe $\{X_t; t = 1, \ldots, T\}$ and let $J_T(\omega_k)$ be its DFT defined as

$$J_T(\omega_k) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} X_t \exp(it\omega_k), \text{ for } 1 \leq k \leq T,$$

where $\omega_k = \frac{2\pi k}{T}$ are the canonical frequencies. It is well known that if $\{X_t\}$ is a strictly stationary time series, whose covariances are absolutely summable then for $k_1 \neq k_2$ and $k_1 \neq T - k_2$ we have $\mathrm{cov}\big(J_T(\omega_{k_1}), J_T(\omega_{k_2})\big) = O(\frac{1}{T})$. Therefore in the case of stationary processes, the discrete Fourier transform $\{J_T(\omega_k)\}_{k=1}^{T}$ is asymptotically uncorrelated. Let

$$\kappa(t,\tau) = \mathbb{E}(X_t, X_\tau) = \frac{1}{T} \sum_{k_1,k_2=1}^{T} \mathbb{E}(J_T(\omega_{k_1})\overline{J_T(\omega_{k_2})}) \exp(-it\omega_{k_1} + i\tau\omega_{k_2}), \tag{1}$$

where $\overline{z}$ is the complex conjugate of the complex variable z. From the above we observe if $\mathbb{E}(J_T(\omega_{k_1})\overline{J_T(\omega_{k_2})}) = 0$ for $k_1 \neq k_2$ or $k_1 \neq T-k_2$, then we have $\kappa(t,\tau) = \kappa(t-\tau)$ for $0 \leq t,\tau \leq T-1$. In other words, an uncorrelated discrete Fourier transform sequence implies that the original time series is second order stationary, up to lag T. This argument can be generalised. Suppose that $\mathbb{E}(J_T(\omega_{k_1})\overline{J_T(\omega_{k_2})}) = O(T^{-1}|k_1 - k_2|^{-1})$ for $k_1 \neq k_2$, then using (1) the process is second stationary, since for all T, $\kappa(t,\tau) = \kappa_T(t-\tau) + O(\log T/T)$ (where $\kappa_T(t-\tau) = \frac{1}{T}\sum_{k=1}^{T} \mathbb{E}|J_T(\omega_k)|^2 \exp(-i(t-\tau)\omega_k))$.

Let us consider a simple example, where we show that if $\{X_t\}$ are independent, heteroscedastic random variables, then the sequence $\{J_T(\omega_k)\}$ is not uncorrelated. Let us suppose $X_t = \sigma_t\varepsilon_t$, where σ_t is a deterministic, time dependent function and $\{\varepsilon_t\}$ are independent identically distributed (iid) random variables with $\mathbb{E}(\varepsilon_t) = 0$ and $\mathrm{var}(\varepsilon_t) = 1$. In this case, the covariance of the DFT at the canonical frequencies is

$$\mathbb{E}(J_T(\omega_{k_1})\overline{J_T(\omega_{k_2})}) = \frac{1}{2\pi T}\sum_{t=1}^{T}\sigma_t^2\exp(it(\omega_{k_1}-\omega_{k_2})).$$

From the above, it is clear that $\mathrm{cov}(J_T(\omega_{k_1}), J_T(\omega_{k_2})) \neq 0$ for some $k_1 \neq k_2$ (if σ_t is not constant over t).

2.2 The test statistic

The above observations lead us to the following test statistic. We note that, if the the time series is strictly stationary, then $\mathbb{E}(J_T(\omega_k)) = 0$ and $\mathrm{var}(J_T(\omega_k)) \to f(\omega_k)$ as $T \to \infty$, where $f : [0, 2\pi] \to \mathbb{R}$ is the spectral density of the original time series $\{X_t\}$ (see Priestley (1981) and Brockwell and Davis (1987)). Therefore by standardising with $\sqrt{f(\omega_k)}$, under the null of stationarity, $\{J_T(\omega_k)/\sqrt{f(\omega_k)}\}$ is close to an uncorrelated, second order stationary sequence. Therefore to test for stationarity of $\{X_t\}$ we will test for randomness of the sequence $\{J_T(\omega_k)/\sqrt{f(\omega_k)}\}$. The proposed test will be a type of Portmanteau test (see Chen and Deo (2004) for applications of the Portmanteau test in time series analysis). Of course, in reality the spectral density $f(\omega)$ is unknown, therefore we will replace $f(\cdot)$ with the estimated spectral density function $\widehat{f}_T(\cdot)$, where

$$\widehat{f}_T(\omega_k) = \sum_j \frac{1}{bT}K(\frac{\omega_k-\omega_j}{b})|J_T(\omega_j)|^2, \tag{2}$$

$K : [-1, 1] \to \mathbb{R}$ is a positive, continuous, symmetric kernel function which satisfies $\int_{-1}^{1} K(x)dx = 1$ and $\int_{-1}^{1} K(x)^2dx < \infty$ and b is a bandwidth. We mention that so long as the bandwidth is chosen such that $T^{-1/2} << b << T^{-1/4}$ then it does not play any role in the asymptotic rates.

We define the empirical covariance at lag r, which is complex valued, of the discrete Fourier transform as

$$\widehat{c}_T(r) = \frac{1}{T}\sum_{k=1}^{T}\frac{J_T(\omega_k)\overline{J_T(\omega_{k+r})}}{\sqrt{\widehat{f}_T(\omega_k)\widehat{f}_T(\omega_{k+r})}}, \quad \text{for } 1 \le r \le T-1. \tag{3}$$

In Figures 1 and 2 plots of the DFT covariance $\widehat{c}_T(r)$ over various lags of stationary and nonstationary time series are given. We observe that there is a clear difference between the DFT covariance of stationary and nonstationary time series. In the stationary case the DFT covariances tend to be smaller (at least for small lags r) than the DFT covariance of a nonstationary time series.

The proposed test statistic is based on $\widehat{c}_T(r)$. We show in Lemma A.10 that if $\{X_t\}$ is a strictly stationary time series, $\mathbb{E}(X_t^4) < \infty$ and the fourth order cumulants are absolutely summable, then both the variance of the real and imaginary parts of $\sqrt{T}\widehat{c}_T(r)$ converge to $1 + \kappa(\omega_r)$ as $T \to \infty$ where

$$\kappa(\omega_r) = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{2\pi} \frac{f_4(\lambda_1, -\lambda_1 - \omega_r, -\lambda_2)}{\sqrt{f(\lambda_1)f(\lambda_1+\omega_r)f(\lambda_2)f(\lambda_2+\omega_r)}} d\lambda_1 d\lambda_2, \tag{4}$$

$\omega_r = 2\pi r/T$ and $f_4(\lambda_1, \lambda_2, \lambda_3) = (2\pi)^{-3}\sum_{j_1,j_2,j_3=-\infty}^{\infty} \mathrm{cum}(X_0, X_{j_1}, X_{j_2}, X_{j_3}) \exp(i(\lambda_1 j_1 + \lambda_2 j_2 + \lambda_3 j_3))$ is the tri-spectra. Under the null hypothesis of strict stationarity, we show in Theorem 3.2 that

$$\sqrt{T}\left(\frac{1}{1+\kappa(\omega_1)}\Re\widehat{c}_T(1), \ldots, \frac{1}{1+\kappa(\omega_m)}\Im\widehat{c}_T(m)\right) \xrightarrow{D} \mathcal{N}(0, I_{2m}), \tag{5}$$

hence the empirical covariances $\widehat{c}_T(r)$ at different lags are asymptotically uncorrelated and $\widehat{c}_T(r) = o_p(1)$. Motivated by the above result we define the test statistic

$$\mathcal{T}_m = T\sum_{r=1}^{m} \frac{|\widehat{c}_T(r)|^2}{1+\kappa(\omega_r)},$$

where $|z|^2 = z\overline{z}$ and $r \neq 0$ or $T/2$. We note, that unlike the classical Portmanteau tests, using covariances with a large lag is not problematic as the DFT is periodic.

We derive the asymptotic distribution of the test statistic in Section 3, under the null hypothesis that either $\{X_t\}$ statisfies the MA(∞) representation

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \tag{6}$$

where $\{\varepsilon_t\}$ are iid random variables with $\mathbb{E}(\varepsilon_t) = 0$, $\mathbb{E}(\varepsilon_t^2) = 1$ and $\kappa_4 = \mathrm{cum}_4(\varepsilon_t)$ or that $\{X_t\}$ is a strictly stationary time series which is α-mixing (defined in Assumption 3.1, below), which includes a large class of time series (see, for example, Doukhan (1994) and Cline and Pu (1999)). We mention that not all linear processes are mixing hence we have separated the assumptions and the proof of asymptotic normality for both cases is different. Under these assumptions we will show in Corollary 3.1, below, that $\widehat{c}_T(r) = o_p(1)$ and $\mathcal{T}_m$ converges in distribution to χ^2_{2m}. Therefore we reject the null of strict stationarity at the $\alpha\%$ significance level if $\mathcal{T}_m > \chi^2_{2m}(1-\alpha)$.

In the case of linearity, (4) has an interesting form. It can be shown that

$$\kappa(\omega_r) = \frac{\kappa_4}{2}\left|\frac{1}{2\pi}\int_0^{2\pi} \exp(i\phi(\omega) - \phi(\omega+\omega_r))d\omega\right|^2,$$

where

$$\phi(\omega) = \arctan\frac{\Im A(\omega)}{\Re A(\omega)} \text{ and } A(\omega) = \frac{1}{\sqrt{2\pi}}\sum_{j=0}^{\infty}\psi_j \exp(i\omega j).$$

Hence in the case of linearity, if m is small then by the continuity of ϕ the test statistic can be approximated by

$$\mathcal{T}_m = T\sum_{r=1}^{m}\frac{|\widehat{c}_T(r)|^2}{(1+\kappa_4/2)} \qquad \text{where} \quad \kappa_4 = \mathrm{cum}_4(\varepsilon).$$

Remark 2.1 (Estimation of the tri-spectra) *We observe that the test statistic $\mathcal{T}_m$ requires estimates of the parameter $\kappa(\omega_r)$. Therefore to estimate this parameter we require estimators of the tri-spectra and spectral density. Brillinger and Rosenblatt (1967) propose an estimator of the tri-spectra $f_4(\cdot)$ (denote this as $\widehat{f}_{4,T}(\cdot)$), which is consistent under the assumption of strict stationarity (our null hypothesis). Therefore, an estimator of $\kappa(\omega_r)$ is*

$$\hat{\kappa}_{r,T} = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{2\pi}\frac{\widehat{f}_{4,T}(\lambda_1,-\lambda_1-\omega_r,-\lambda_2)}{\sqrt{\widehat{f}_T(\lambda_1)\widehat{f}_T(\lambda_1+\omega_r)\widehat{f}_T(\lambda_2)\widehat{f}_T(\lambda_2+\omega_r)}}d\lambda_1 d\lambda_2,$$

where $\widehat{f}_T(\lambda_2)$ is defined in (2). Since $\hat{\kappa}_{r,T}$ is a consistent estimator of $\kappa(\omega_r)$, replacing $\kappa(\omega_r)$ in the test statistic with $\hat{\kappa}_{r,T}$, does not alter the asymptotic sampling distribution of $\mathcal{T}_m$.

Remark 2.2 (Practical issues) *(i) Simulations demonstrate that the test is not sensitive to the bandwidth b, and we suggest selecting the bandwidth using the method described in Beltrao and Bloomfield (1987).*

(ii) The asymptotic distribution under the null is derived under the assumptions that the spectral density of the time series $\{X_t\}$ is bounded away from zero. In practice, even if this assumption holds, the estimated spectral density $\widehat{f}_T(\cdot)$ may be quite close to zero. Therefore, in this case, to prevent falsely rejecting the null, we suggest adding a small ridge to the spectral density estimator $\widehat{f}_T(\cdot)$ to bound it away from zero.

2.3 The power of the test and selection of m

In Section 4 we obtain the asymptotic sampling properties of the test statistic $\mathcal{T}_m$, under the alternative of local stationarity. In order to understand what nonstationary behaviour the test statistic can detect and how to select the lag m in the test statistic, we will now outline some of the results in Section 4. Suppose that $\{X_t\}$ is a nonstationary time series, where in a small neighbourhood of t the observations are close to stationary and has the local spectral density $f(\frac{t}{T},\omega)$.

In Lemma 4.1 we show that $\widehat{c}_T(r) = B(r) + o_p(1)$, and $\sqrt{T}\big((\Re\widehat{c}_T(r) - \Re B(r)), (\Im\widehat{c}_T(r) - \Im B(r))\big) \xrightarrow{D} \mathcal{N}(0, \Sigma_r)$ where Σ_r is defined in Theorem 4.2 and

$$B(r) = \frac{1}{2\pi}\int_0^{2\pi}\int_0^1 \frac{f(u,\lambda)}{[f(\lambda)f(\lambda+\omega_r)]^{1/2}} \exp(-i2\pi ur)dud\lambda. \tag{7}$$

Furthermore, we show that under the alternative of local stationarity, $\mathcal{T}_m$ has asymptotically a non-central generalised chi-squared distribution where the noncentrality parameter is, roughly speaking, $\delta_m = T\sum_{r=1}^m |B(r)|^2$. Hence for large T, the test statistic $\mathcal{T}_m = O(T)$ and will have close to 100% power.

The main parameter to influence the DFT test is the choice of m. To select m, let us consider how m may influence the power. We observe that the power of the test depends on $B(r)$ and rewriting $B(r)$ we see that

$$B(r) = \frac{1}{2\pi}\int_0^{2\pi} F_2(r,\lambda)d\lambda,$$

where $F_2(r,\lambda) = \int_0^1 \frac{f(u,\lambda)}{[f(\lambda)f(\lambda+\omega_r)]^{1/2}} \exp(-i2\pi ur)du$. We show in Lemma A.2, that under Assumption 4.1, $\sup_\lambda |F_2(r,\lambda)| \leq Kr^{-2}$, therefore $|B(r)| = O(r^{-2})$ and $\sum_r |B(r)| < \infty$. Since the non-centrality parameter $\delta_m = T\sum_{r=1}^m |B(r)|^2$, it makes sense to use only small lags r, since for large lags $|B(r)|$ will be small and the contribution to the noncentrality parameter δ^2 with be neglible. Therefore a large m will result in a loss of power, since under the null $\mathcal{T}_m \xrightarrow{D} \chi^2_{2m}$ (this is analogous to the Ljung-Box test for independence, using the estimated covariance at large lags results in a loss in power). In general, relatively smooth functions contain mainly low frequency information and a Fourier expansion of the function to 10 terms can represent a function well. Therefore, in general a 'rule of thumb' is to use a maximum of $m = 10$ in the test statistic. One method for selecting m is to plot the estimated covariances $T|\widehat{c}_T(r)|^2$ against the lags, and use only the first m large lags to construct the test (see Figures 1 and 2). The plot of $T|\widehat{c}_T(r)|^2$ can also serve as a useful visual aid for 'viewing' the nonstationarity.

Remark 2.3 *An alternative test, motivated by the asymptotic result (5) is*

$$\mathcal{T}_m^* = T\max_{1\leq r\leq m}\frac{\widehat{c}_T(r)|^2}{1+\kappa(\omega_r)}.$$

Under the null of strict stationarity the asymptotic distribution of $\mathcal{T}_m^*$ has the density $\frac{m}{2}\exp(-x/2)(1-\exp(-x/2))^{m-1}$. The power of this test under the locally stationary alternative can be determined by using Theorem 4.2. We will compare $\mathcal{T}_m$ and $\mathcal{T}_m^*$ in the simulations.

To consider the type of nonstationary behaviour the test statistic can detect let us consider the

Fourier expansion of the the function $f(u,\omega)f(\omega)^{-1}$

$$\frac{f(u,\omega)}{f(\omega)} = \frac{1}{2\pi}\sum_{r}\sum_{j}\alpha(r,j)\exp(i2\pi ru)\exp(ij\omega), \tag{8}$$

where $\alpha(r,j) = \int\int \frac{f(u,\omega)}{f(\omega)}\exp(-i2\pi ru - ij\omega)dud\omega$. We observe that when $f(u,\omega) \equiv f(\omega)$ (the stationary situation), all the Fourier coefficients will be zero except for $\alpha(0,0)$. Hence the coefficients $\alpha(r,j)$ describe the nonstationary behaviour of $\frac{f(u,\omega)}{f(\omega)}$ at various Fourier frequencies.

Now for $r << T$ we have that $\omega_r = \frac{2\pi r}{T} \to 0$ as $T \to \infty$, hence by comparing $B(r)$ with (8) we observe $B(r) \approx \alpha(r,0)$. Hence for small r $B(r)$, detects the average 'long-term' nonstationary behaviour. This implies that the test can detect very slow and gradual changes in the time series, which segment based testing methods may have problems detecting.

Remark 2.4 *The other coefficients in (8), $\alpha(r,j)$, can be estimated by using $\widehat{c}_T(r,j)$ where*

$$\widehat{c}_T(r,j) = \frac{1}{T}\sum_{k=1}^{T}\frac{J_T(\omega_k)\overline{J_T(\omega_{k+r})}}{\sqrt{\widehat{f}_T(\omega_k)\widehat{f}_T(\omega_{k+r})}}\exp(-ij\omega_k).$$

The sampling properties of $\widehat{c}_T(r,j)$ are similar to the sampling properties of $\widehat{c}_T(r)$ and the test statistic can be modified to detect for nonzero $\alpha(r,j)$.

3 Sampling properties of the test statistic under the null

We now derive the asymptotic distribution of $\widehat{c}_T(r)$ and $\mathcal{T}_m$ under the null of strict stationarity. To simplify the analysis of the test statistic $\mathcal{T}_m$ we replace the denominator in the covariance $\widehat{c}_T(r)$ with its deterministic limit. To do this, we define the unobserved DFT covariance

$$\widetilde{c}_T(r) = \frac{1}{T}\sum_{k=1}^{T}\frac{J_T(\omega_k)\overline{J_T(\omega_{k+r})}}{\sqrt{f(\omega_k)f(\omega_{k+r})}}. \tag{9}$$

We will show that the difference between $\widetilde{c}_T(r)$ and $\widetilde{c}_T(r)$ is negligible (hence the limiting distributions of $\sqrt{T}\widetilde{c}_T(r)$ and $\sqrt{T}\widetilde{c}_T(r)$ are asymptotically equivalent). This is technically quite challanging and requires some of the cumulant results developed in Brillinger (1981). In order to use these results we require the relatively strong moment assumptions given below.

Assumption 3.1 *Let us suppose that $\{X_t\}$ is a strictly stationary time series and let $f(\omega) = \frac{1}{2\pi}\sum_{r=-\infty}^{\infty} cov(X_0, X_r)\exp(ir\omega)$.*

(A) **Strictly stationary linear time series**

Let us suppose that $\{X_t\}$ satisfies (6).

(i) $\sum_{j=0}^{\infty}|j\psi_j|<\infty$ *(noting that this implies* $\sup_{\omega}|f'(\omega)|<\infty$*).*

(ii) $\mathbb{E}(\varepsilon_t^{16})<\infty$ *(noting this implies* $\mathbb{E}(X_t^{16})<\infty$*).*

(B) **Strictly stationary α-mixing time series**

Let us suppose that $\{X_t\}$ is a strictly stationary time series which is α-mixing,that is satisfies

$$\sup_{\substack{A\in\sigma(X_r,X_{r+1},\ldots)\\ B\in\sigma(X_0,X_{-1},\ldots)}}|P(A\cap B)-P(A)P(B)|\leq\alpha(r).$$

(i) $\sum_r|r\,cov(X_0,X_r)|<\infty$ *(which implies* $\sup_{\omega}|f'(\omega)|<\infty$*).*

(ii) For some $\delta>0$ we have $\sum_r|r|\cdot|\alpha(r)|^{\delta/(15(1+\delta))}<\infty$ *and* $\mathbb{E}|X_t|^{16(1+\delta)}<\infty$*.*

(iii) $\inf_{\omega}f(\omega)>0$ *and if $\{X_t\}$ is a linear time series* $\inf_{\omega}|\Re A(\omega)|^2>0$ *(where* $A(\omega)=(2\pi)^{1/2}\sum_{j=0}^{\infty}\psi_j\exp(ij\omega)$*).*

(iv) Either (a) $\sum_r|r^2||cov(X_0,X_r)|<\infty$ *(implied by* $\sum_j|j^2\psi_j|<\infty$ *in the linear case) or (b) the derivative of the spectrum $f'(\omega)$ is piecewise montone on the interval $[0,2\pi]$ (in other words $f'(\cdot)$ can be partitioned into a finite number of intervals which is either nonincreasing or nondecreasing).*

We use Assumption 3.1(iv) to obtain the rate of decay of the Fourier coefficients of the function $\frac{1}{\sqrt{f(\omega)f(\omega+\omega_r)}}$. We observe, that in the case the time series is linear, the assumptions are in some sense weaker. In the results below if we state that 'Assumption 3.1 holds', this means that either the linearity or the α-mixing assumption can hold.

Theorem 3.1 *Suppose Assumption 3.1 is satisfied and let $\widehat{c}_T(r)$ and $\widetilde{c}_T(r)$ be defined as in (3) and (9) respectively. Then we have*

$$\sqrt{T}|\widehat{c}_T(r)-\widetilde{c}_T(r)|=O_p\Big((b+\frac{1}{\sqrt{bT}})+\big(\frac{1}{bT^{1/2}}+b^2T^{1/2}\big)\big(\frac{1}{|r|}+\frac{1}{T^{1/2}}\big)\Big). \tag{10}$$

PROOF. See Appendix A.2, Lemma A.8, equation (33). □

In the following lemma we derive the asymptotic variance of $\tilde{c}_T(r)$, and show that $\tilde{c}_T(r)$ is asymptotically uncorrelated at different lags r, and at the real and imaginary parts.

Lemma 3.1 *Suppose Assumption 3.1 holds. Then we have*

$$
\begin{aligned}
& cov\big(\sqrt{T}\Re\widetilde{c}_T(r_1), \sqrt{T}\Re\widetilde{c}_T(r_2)\big) = cov\big(\sqrt{T}\Im\widetilde{c}_T(r_1), \sqrt{T}\Im\widetilde{c}_T(r_2)\big) \\
= & \left\{ \begin{array}{cc} O(T^{-1}) & r_1 \neq r_2 \\ 1 + \frac{1}{4\pi}\int_0^{2\pi}\int_0^{2\pi} \frac{f_4(\omega_1, -\omega_1-\omega_r, -\omega_2)}{\sqrt{f(\omega_1)f(\omega_1+\omega_r)f(\omega_2)f(\omega_2+\omega_r)}} d\omega_1 d\omega_2 + O(\frac{1}{T}) & r_1 = r_2 = r \end{array} \right.
\end{aligned} \tag{11}
$$

and for all $r_1, r_2 \in \mathbb{Z}$, $cov(\sqrt{T}\Im\tilde{c}_T(r_1), \sqrt{T}\Re\tilde{c}_T(r_2)) = O(\frac{1}{T})$.

PROOF. See Appendix A.3. $\square$

Remark 3.1 *Suppose Assumption 3.1(A,iii,iv) holds. Then we have*

$$
\begin{aligned}
& cov(\sqrt{T}\Re\tilde{c}_T(r), \sqrt{T}\Re\tilde{c}_T(r)) = cov(\sqrt{T}\Im\tilde{c}_T(r), \sqrt{T}\Im\tilde{c}_T(r)) \\
= \;& 1 + \frac{\kappa_4}{2}\left|\frac{1}{2\pi}\int_0^{2\pi} \exp(i\phi(\omega) - \phi(\omega + \frac{2\pi r}{T}))d\omega\right|^2 + O(\frac{1}{T}).
\end{aligned}
$$

The proof of this result can be found in Appendix A.3.

We now show normality of $\widehat{c}_T(r)$, which we use to obtain the distribution of $\mathcal{T}_m$.

Theorem 3.2 *Suppose Assumption 3.1 holds. Then for fixed m we have*

$$
\sqrt{T}\left(\frac{1}{1+\kappa(\omega_1)}\Re\widehat{c}_T(1), \ldots, \frac{1}{1+\kappa(\omega_m)}\Im\widehat{c}_T(m)\right) \xrightarrow{D} \mathcal{N}(0, I_{2m}), \tag{12}
$$

as $m(b + \frac{1}{\sqrt{bT}}) + \left(\frac{1}{bT^{1/2}} + b^2T^{1/2}\right)\sum_{r=1}^{m}\left(\frac{1}{|r_m|} + \frac{1}{T^{1/2}}\right) \to 0$ and $T \to \infty$, where I_{2m} is the identity matrix and $\kappa(\cdot)$ is defined in (4)

PROOF. See Appendix A.4. $\square$

By using the above we are able to obtain the asymptotic distribution of $\mathcal{T}_m$.

Corollary 3.1 *Suppose Assumption 3.1 holds. Then for fixed m we have $\mathcal{T}_m \xrightarrow{D} \chi^2_{2m}$ with $m(b + \frac{1}{\sqrt{bT}}) + \left(\frac{1}{bT^{1/2}} + b^2T^{1/2}\right)\sum_{r=1}^{m}\left(\frac{1}{|r|} + \frac{1}{T^{1/2}}\right) \to 0,$ as $T \to \infty$.*

PROOF. The result immediately follows from Theorem 3.2. $\square$

4 Sampling properties of the test statistic under the alternative of local stationarity

It is useful to investigate the behaviour of the test statistic in the case that the null hypothesis does not hold. If the covariance structure varies over time, then the limit of $\widehat{c}_T(r)$ will not be zero. This suggests that the test statistic will have a type of non-central χ^2 distribution. However, in the case that time-varying covariance has no structure it is not clear what the limit of $\widehat{c}_T(r)$ will be. Therefore we consider the behaviour of the test statistic for the class of locally stationary processes, which as the name suggests is a class of processes which can locally be approximated by a stationary time series. The asymptotic considerations for this class of processes are based on infill asymptotics (see Robinson (1989), Dahlhaus (1997)), where we observe $\{X_{t,T}\}$ on a increasingly finer grid. Infill asymptotics will lead to a well defined limit of $\widehat{c}_T(r)$. Following Dahlhaus and Polonik (2006), $\{X_{t,T}\}$ is said to have a time-varying MA(∞) representation if it satisfies

$$X_{t,T} = \sum_{j=0}^{\infty} \psi_{t,T}(j)\varepsilon_{t-j}, \tag{13}$$

where $\{\varepsilon_t\}$ are iid random variables, $\mathbb{E}(\varepsilon_t) = 0$, $\mathrm{var}(\varepsilon_t) = 1$.

In order for $\{X_{t,T}\}$ to be a locally stationary time series, we will assume that $\psi_{t,T}(j)$ closely approximates the smooth function $\psi_j(\cdot)$. Hence the time-varying MA parameters $\{\psi_{t,T}(j)\}$ vary slowly over time. It can be shown that in this case, $\{X_{t,T}\}$ is a locally stationary time series because it can locally be approximated by a stationary time series. We will use the following assumptions.

Assumption 4.1 *Let us suppose that $\{X_{t,T}\}$ satisfies (13). Suppose, there exists a sequence of functions $\psi_j(u)$, such that $\psi_j(u)$ is Lipschitz continuous and $|\psi_j(\frac{t}{T}) - \psi_{t,T}(j)| \leq T^{-1}\ell(j)^{-1}$, where $\{\ell(j)^{-1}\}$ is a positive monotonically decreasing function which satisfies $\sum_j |j|^2\ell(j)^{-1} < \infty$. Let $f(u,\omega) = (2\pi)^{-1}|\sum_{j=0}^{\infty}\psi_j(u)\exp(ij\omega)|^2$ (hence $\sup_{u,\omega}|\frac{\partial^2 f(u,\omega)}{\partial\omega^2}| < \infty$).*

(i) $\sup_u |\psi_j(u)| < \ell(j)^{-1}$ *and* $\sup_u |\frac{d\psi_j(u)}{du}| < K\ell(j)^{-1}$ *(hence* $\sup_{u,\omega}|\frac{\partial f(u,\omega)}{\partial u}| < \infty$*).*

(ii) $\mathbb{E}(\varepsilon_t^{16}) < \infty$.

(iii) Define the integrated spectral density $f(\omega) = \int_0^1 f(u,\omega)du$, *and assume that* $\inf_\omega f(\omega) > 0$.

(iv) Either (a) $\sup_u \sum_j |\psi_j''(u)| < \infty$ *(hence* $\sup_{u,\omega}|\frac{\partial^2 f(u,\omega)}{\partial u^2}| < \infty$*) or (b)* $A(u,\omega)$ *and* $\frac{\partial A(u,\omega)}{\partial u}$ *are piecewise monotone functions with respect to* u.

We will show in Lemma 4.1, below, that in the locally stationary case the spectral density estimator $\widehat{f}_T(\cdot)$ defined in (2) estimates the integrated spectrum $f(\omega)$, where $f(\omega)$ is defined in Assumption

4.1(iii). Roughly speaking, one can consider the integrated spectrum as the average of the locally stationary spectrums.

As in Section 3, it is difficult to directly obtain the distribution of $\widehat{c}_T(r)$. Instead we replace the random denominator with its deterministic limit (that is $J_T(\omega_k)/\sqrt{\widehat{f}_T(\omega_k)}$ with $J_T(\omega_k)/\sqrt{f(\omega_k)}$), and define

$$\widetilde{c}_T(r) = \frac{1}{T}\sum_{k=1}^{T}\frac{J_T(\omega_k)\overline{J_T(\omega_{k+r})}}{\sqrt{f(\omega_k)f(\omega_{k+r})}}, \tag{14}$$

where $f(\cdot)$ is the integrated spectrum. The following result is the locally stationary analogue of Theorem 3.1.

Theorem 4.1 *Suppose Assumption 3.1 is statisfied, and let $\widehat{c}_T(r)$ and $\widetilde{c}_T(r)$ be defined as in (3) and (14) respectively. Then we have*

$$\sqrt{T}|\widehat{c}_T(r) - \widetilde{c}_T(r)| = O_p\bigg(\frac{1}{\sqrt{bT}} + \big(\frac{1}{bT^{1/2}} + b^2T^{1/2}\big)\big(\frac{1}{|r|} + \frac{1}{T^{1/2}}\big)\bigg).$$

PROOF. In Appendix A.2, Lemma A.8, equation (34). □

From the lemma above we see that in order for the sampling properties of $\sqrt{T}\widehat{c}_T(r)$ and $\sqrt{T}\widetilde{c}_T(r)$ to coincide, we require that $T^{-1/2} << b << T^{-1/4}$.

We now obtain the mean and variance of $\tilde{c}_T(r)$ under the alternative hypothesis of local stationarity.

Lemma 4.1 *Suppose Assumption 4.1 are satisfied and let $f(\omega)$ and $f(u,\omega)$ be the integrated and local spectrum (defined in Assumption 4.1) respectively. Then we have*

$$\mathbb{E}\big(\hat{f}_T(\omega) - f(\omega)\big)^2 = O\big(b^2 + \frac{1}{bT}\big), \tag{15}$$

$\mathbb{E}(\tilde{c}_T(r)) \to B(r)$ as $T \to \infty$, and

$$\begin{aligned} cov(\Re\sqrt{T}\tilde{c}_T(r_1), \Re\sqrt{T}\tilde{c}_T(r_2)) &\to \Sigma^{(1,1)}_{T,r_1,r_2} \quad cov(\Re\sqrt{T}\tilde{c}_T(r_1), \Im\sqrt{T}\tilde{c}_T(r_2)) \to \Sigma^{(1,2)}_{T,r_1,r_2} \\ cov(\Im\sqrt{T}\tilde{c}_T(r_1), \Im\sqrt{T}\tilde{c}_T(r_2)) &\to \Sigma^{(2,2)}_{T,r_1,r_2}, \end{aligned} \tag{16}$$

$b \to 0$, $bT \to \infty$ as $T \to \infty$, where

$$\Sigma^{(1,1)}_{T,r_1,r_2} = \frac{1}{4}\big(\Gamma^{(1)}_{T,r_1,r_2} + \Gamma^{(2)}_{T,r_1,r_2} + \Gamma^{(2)}_{T,r_2,r_1} + \Gamma^{(3)}_{T,r_1,r_2}\big) + O(\frac{\log T}{T}),$$

$$\Sigma^{(1,2)}_{T,r_1,r_2} = \frac{-i}{4}\big(\Gamma^{(1)}_{T,r_1,r_2} + \Gamma^{(2)}_{T,r_1,r_2} - \Gamma^{(2)}_{T,r_2,r_1} - \Gamma^{(3)}_{T,r_1,r_2}\big) + O(\frac{\log T}{T}),$$

$$\Sigma^{(2,2)}_{T,r_1,r_2} = \frac{1}{4}\big(\Gamma^{(1)}_{T,r_1,r_2} - \Gamma^{(2)}_{T,r_1,r_2} - \Gamma^{(2)}_{T,r_2,r_1} + \Gamma^{(3)}_{T,r_1,r_2}\big) + O(\frac{\log T}{T}),$$

and $\Gamma^{(i)}_{T,r_2,r_1}$ $(i = 1,2,3)$ *are defined in Lemma A.12 (in Appendix A.3).*

PROOF. See Appendix A.3. $\square$

We use the above to obtain the asymptotic distribution of $\mathcal{T}_m$ under the alternative. First we recall that we estimated the standardisation of $\widehat{c}_T(r)$, κ_r, in Remark 2.1. In the case of local stationarity, $\hat{\kappa}_{r,T}$ is an estimator of

$$\kappa(\omega_r) = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{2\pi} \frac{f_4(\lambda_1, -\lambda_1-\omega_r, -\lambda_2)}{\sqrt{f(\lambda_1)f(\lambda_1+\lambda_2)f(\lambda_2)f(\lambda_2+\omega_r)}} d\lambda_1 d\lambda_2,$$

where $f(\cdot)$ is the integrated spectral density and and $f_4(\lambda_1,\lambda_2,\lambda_3) = \int_0^1 f_4(u,\lambda_1,\lambda_2,\lambda_3)du$, with $f_4(u,\lambda_1,\lambda_2,\lambda_3) = \frac{1}{2\pi}A(u,-\lambda_1-\lambda_2-\lambda_3)\prod_{i=1}^3 A(u,\omega_i)$.

Theorem 4.2 *Suppose Assumption 4.1 holds. Let*

$$\Sigma = \begin{pmatrix} \Sigma^{(1,1)}_T & \Sigma^{(1,2)}_T \\ \Sigma^{(2,1)}_T & \Sigma^{(2,2)}_T \end{pmatrix},$$

where $\Sigma^{(1,1)}_{T,r_1,r_2}$, $\Sigma^{(1,2)}_{T,r_1,r_2}$, $\Sigma^{(2,2)}_{T,r_1,r_2}$ *are defined in Lemma 4.1 and* $\Sigma^{(1,2)}_{T,r_1,r_2} = \overline{\Sigma^{(2,1)}_{T,r_1,r_2}}$. *Furthermore define* $\boldsymbol{\mu}' = (\Re B(1),\ldots,\Re B(m),\Im B(1),\ldots,\Im B(m))$, *where* $B(r)$ *is defined in (7). Then we have*

$$\sqrt{T}\Big(\big(\Re\widehat{c}_T(1),\Re\widehat{c}_T(2),\ldots,\Im\widehat{c}_T(m)\big) - \boldsymbol{\mu}'\Big) \xrightarrow{D} \mathcal{N}(0,\Sigma), \tag{17}$$

and

$$\mathcal{T}_m \xrightarrow{D} \sum_{n=1}^{m} \frac{\big(X_n^2 + Y_n^2\big)}{(1+\kappa(\omega_r))},$$

with $\frac{m}{\sqrt{bT}} + \big(\frac{1}{bT^{1/2}} + b^2T^{1/2}\big)\sum_{r=1}^m\big(\frac{1}{|r|} + \frac{1}{T^{1/2}}\big) \to 0$ *as* $T\to\infty$, *where* $\boldsymbol{X}_{2m}$ *is a normally distributed random vector with* $\boldsymbol{X}_{2m} = (X_1,\ldots,X_m,Y_1,\ldots,Y_m)'$ *and* $\boldsymbol{X}_{2m} \sim \mathcal{N}(\sqrt{T}\boldsymbol{\mu},\Sigma)$. *Note the small abuse of notation, when we say* $A \xrightarrow{D} B$, *we mean that the distribution of random variable* A *converges to the distribution of random variable* B. *Hence* $\mathcal{T}_m$ *is a non-central (determined by* $\{T|B(r)|^2\}_{r=1}^m$*) generalised (determined by* Σ *and* $(1+\kappa(\omega_r))$*) chi-squared distribution with* $2m$ *degrees of freedom.*

Remark 4.1 *It is natural to ask whether there exists locally stationary tvMA(*∞*) processes, which the DFT test,* $\mathcal{T}_m$*, cannot detect. In other words,* $B(r) = 0$ *for all* $r \neq 0$. *Since one would only* *use relatively small r (since $B(r) = O(r^{-2})$) in the test, $B(r) \approx \alpha(r,0)$ (see equation (8)), this is equivalent to the existence of locally stationary processes were $\alpha(r,0) = 0$ for all $r \neq 0$.*

We first consider the general case of local stationarity. We note that since $f(u,\omega)/f(\omega)$ are positive and symmetric functions (symmetric over ω), there exists a time-varying positive definite sequence $\{a(u,j)\}_j$ which has the spectrum $f(u,\omega)/f(\omega)$. Moreover, since $\int_0^1 f(u,\omega)du = f(\omega)$, the functions $\{a(u,s)\}$ are such that $\int_0^1 a(u,s)du = 0$ for all $s \neq 0$ and $\int_0^1 a(u,0)du = 1$. Based on this construction, if the test cannot detect the alternative, this implies $\alpha(r,0) = 0$ for all r, since it is straightforward to show that $a(u,0) = (2\pi)^{-1}\sum_j \alpha(r,0)\exp(ir\omega)$. Therefore, $a(u,0) = 1$ for all u, and for the test not to detect the alternative, $f(u,\omega)/f(\omega)$ must satisfy

$$\frac{f(u,\omega)}{f(\omega)} = 1 + \sum_{s\neq 0} a(u,s)\exp(is\omega), \tag{18}$$

where $\{1, \{a(u,s); s \neq 0\}\}$ is a positive definite sequence which depends on u and $\int_0^1 a(u,s)du = 0$. Such a situation could possibly arise, though the conditions are so stringent it would be extremely rare. However, in the case that we are unable to reject the null we could plot $\{T|\widehat{c}_T(r,1)|^2\}_r$ and look for significant coefficients. One could also re-do the test, but use $T|\widehat{c}_T(r,1)|^2$ instead off $T|\widehat{c}_T(r)|^2$ (in the case of nonGaussianity, a small change of the variance of $\sqrt{T}\widehat{c}_T(r,1)$ would have to be made).

5 Comparisons with the Paparoditis test

We compare the test statistic with the test statistic recently proposed in Paparoditis (2009). Paparoditis (2009) proposes the test statistic

$$\mathcal{S}_T = \frac{1}{N}\sum_{s=1}^{N}\int_0^{2\pi} V_T^2\Big(\frac{m/2 + (s-1)m}{T}, \lambda\Big)d\lambda \tag{19}$$

where m is the segment length, $N = T/m$ is number of segments,

$$V_T(u,\lambda) = \frac{1}{m}\sum_{j=-(m-1)}^{m-1} K\Big(\frac{\lambda - \lambda_j}{\tilde{b}}\Big)\Big(\frac{I_m(u,\lambda)}{\widehat{f}_T(\lambda)} - 1\Big),$$

$\tilde{b}$ is a local bandwidth ($\tilde{b} << m$), $\widehat{f}_T$ the spectral density estimator defined in (2),

$$I_m(u,\lambda) = \frac{1}{2\pi H_{2,m}(0)}\Big|\sum_{t=1}^{m} h\Big(\frac{t}{m}\Big)X_{t+[uT]-m/2-1}\exp(i\lambda t)\Big|^2,$$

$h(\cdot)$ is a taper and $H_{k,m}(\lambda) = \sum_{s=1}^{m} h(t/m)^k \exp(-i\lambda s)$. Under the assumption that $\{X_t\}$ is a stationary linear time series Paparoditis (2009) shows that

$$\sqrt{N\tilde{b}}m\mathcal{S}_T - \mu_T(K, \kappa_4, H) \xrightarrow{D} \mathcal{N}(0, \tau^2(K, \kappa_4, H)), \tag{20}$$

where $\mu_T(K, \kappa_4, H) = O(\sqrt{\frac{N}{\tilde{b}}} + \sqrt{N\tilde{b}})$ and $\tau^2(K, \kappa_4, H) = O(1)$ are functions of the kernel K, fourth order cumulant κ_4 and the taper H. All terms are defined in Paparoditis (2009).

We observe that $\mathcal{S}_T$ is the sum of squared random variables, hence $\mathcal{S}_T$ will be quite skewed for small sample sizes. Therefore, for small T, the normal approximation will not be particularly reliable, this is a well known problem with many L_2-based tests.

It is not possible to make a direct theoretical comparision of the power of both the Paparoditis test and the DFT test, since both distributions under the alternative hypothesis are different and quite complicated. However, it is interesting to note that they both share similar noncentrality parameters. Paparoditis (2009) shows that under the null of local stationarity $\mathcal{S}_T$ is normal, $\mathcal{S}_T = O(T^{1/2})$, and the mean of $\mathcal{S}_T$ is dominated by the term

$$T \int_0^1 \frac{1}{2\pi} \int_0^{2\pi} \big(\frac{f(u,\omega)}{f(\omega)} - 1\big)^2 d\omega du$$

By using (8) we can show that

$$T \int_0^1 \frac{1}{2\pi} \int_0^{2\pi} \big(\frac{f(u,\omega)}{f(\omega)} - 1\big)^2 d\omega du = T\big(\sum_{r,j} |\alpha(r,j)|^2 - 1\big),$$

where the coefficients $\{\alpha(r,j)\}$ are defined in (8). In comparison, in Section 2.3 we showed that that $\widehat{c}_T(r) \xrightarrow{P} B(r) \approx \alpha(r,0)$, and the power of the DFT test, $\mathcal{T}_m$, is determined by $\{T|\alpha(r,0)|^2\}$.

6 Simulations

We do the comparison for both stationary and nonstationary time series. In each case, we replicate the time series 1000 times, and for each replication we do the test. We do the test for sample sizes $T = 64, 128, 256$ and 512. For every simulation we select the bandwidth b using the cross-valiadation method suggested in Beltrao and Bloomfield (1987) and use the Tukey kernel. To select the local bandwidth $\tilde{b}$ in the Paparoditis test we use his suggested rule $\tilde{b} = b(T/m)^{0.25}$. We use the split cosine bell taper and $N = 4$ in the Paparoditis test. For the DFT tests we use the test statistic $\mathcal{T}_m$ ($m = 1, 5, 10$) and the maximum test statistic $\mathcal{T}_m^*$ ($m = 5, 10$). We do the tests at both the 5% and 1% level. The rejection rates can be found in the tables below. We also give the average $|\widehat{c}_T(r)|^2$ for each lag $r = 1, \ldots, m$ over the 1000 realisations, we denote this as $|\bar{c}(r)|^2$ (we note that the larger $|\bar{c}(r)|^2$ the more likely we are to reject the null).

6.1 Stationary time series

We consider the the following models.

(i) Model 1. $AR(2)$ model. $X_t = 0.75X_{t-1} - 0.4X_{t-2} + \varepsilon_t$. The results can be found in Tables 1 and 2. The plot of the estimated densities can be found in Figure 3.

We observe from Table 3 that for both models the rejection rates using the test $\mathcal{T}_m$ for $m = 1, 5, 10$ tends to be about the same. There seems to be very little difference between the rejection rates for different sample sizes. The maximum test $\mathcal{T}_m^*$ tends to be a little more conservative. The Paparoditis test performs well, and for large sample sizes tends to reject the null less than the DFT tests.

(ii) Model 2: $ARMA(1, 2)$ model. $X_t = 0.8X_{t-1} + \varepsilon_t + 0.3\varepsilon_{t-1} + 2\varepsilon_{t-2}$. The results can be found in Tables 3 and 4.

The results for Model 2 tend to be slightly worse than the results for Model 1. This is probably due to the more complicated dependence structure of the ARMA$(1, 2)$ model than the AR(2) model. For the relatively large sample sizes $T = 256, 512$ the number of rejects of the DFT tests are mainly within in the 5% and 1% rejection level. In comparison, the Paparoditis test seems to falsely reject a large proportion of the realisations, and this only improves for $T = 512$.

Under the null both the DFT tests $\mathcal{T}_m$ and $\mathcal{T}_m^*$ tend to reject within the stated rejection values - suggesting that the chi-squared approximation of the test statistic is relatively good. There seems to be little difference in rejection rates for different m, and different stationary models. Furthermore, the average $|\bar{c}(r)|^2$ for both models is small (see 2 and 4). It is interesting to note that the rejection rate at the 5% and 1% for the Paparoditis test tends to be about the same (this is probably because the Paparoditis test statistic tends to be quite skewed). Moreover, the small sample performance of the Paparoditis test for stationary models seems to be depend on the model.

6.2 Nonstationary time series

We now consider three nonstationary time series model. The first model combines the two stationary models considered above. The second model is an autoregressive model where the variance of the innovations is time-varying, this model is different to the other models in the sense that no rescaling is used, ie. the first half of $T = 512$ is the time series for $T = 256$. The last nonstationary model are independent random variables with time-varying variance.

(i) Model 3: First 1/4 of time series is Model 1 and last 3/4 of time series is Model 2.

$$X_t = \begin{cases} 2.425(0.75X_{t-1} - 0.4X_{t-2} + \varepsilon_t) & t = 1, \ldots, 0.25T \\ 0.198(0.8X_{t-1} + \varepsilon_t + 0.3\varepsilon_{t-2} + 2\varepsilon_{t-3}) & t = (0.25T + 1), \ldots, T \end{cases}$$

We have standardised both time series such that both time series have variance one. The results can be found in Tables 5 and 6. A plot of one realisation of $T|\widehat{c}_T(r)|^2$ is given in Figure 1.

As one would expect the rejection rate, for all the tests, increases with the sample size. The Paparoditis test correctly rejects the null slighly more often than the DFT test when the sample size is small. The rejection rate for the Paparoditis test at the 1% level is consistently greater than the DFT tests at the 1% level.

The rejection rate for $\mathcal{T}_1, \mathcal{T}_5$ and $\mathcal{T}_{10}$ seem to be similar for this model. Studying Table 6 we observe that average covariance $|\bar{c}(1)|^2$ seems to hold the most information about the nonstationary behaviour, hence no information is lost by using only $\mathcal{T}_1$ in this test.

(ii) Model 4: The variance of the innovation of the AR process changes smoothly over time. $X_t = 0.8X_{t-1} + \sigma_t \varepsilon_t$, where $\sigma_t = (\frac{1}{2} + \sin(\frac{2t\pi}{512}) + 0.3\cos(\frac{2t\pi}{512}))$. The results can be found in Tables 7 and 8.

The results in this test are quite different to the results for Model 3. Neither the DFT test or the Paparoditis test does well for small sample sizes, though the Paparoditis is slightly better than the DFT for $T = 64$ (9% compared to less than 5.5%).

However, there is a substantial increase in the rejection rate once the sample size increases to $T = 128$ the rejection rate for all tests. Interestingly the rejection rate increases as the sample sizes grows, except for $\mathcal{T}_1$, where the rejection rate drops from 97.9% ($T = 128$) to 58% ($T = 256$) and then increases to 100% ($T = 512$). An explanation for this can be found in the average coefficients $\bar{c}(r)$ for different sample sizes. This model is different to the other models in the sense that the time series with $T = 256$ is not a rescaled version of $T = 128$, this means the average coefficients $\bar{c}(r)$ can change substantially with the sample size. This can be seen from the average coefficients of $\bar{c}(1)$ in Tables 8. $\bar{c}(1)$ is large for $T = 128$ and $T = 512$ but smaller for $T = 256$. This explains why $\mathcal{T}_1$ gives a relatively low rejection rate when $T = 256$.

(iv) Model 5: Independent random variables with time-varying variance. Define the piecewise varying function $\sigma : [0, 1] \to \mathbb{R}$

$$\sigma(u) = \begin{cases} 1 & \text{for } u \in \{[\frac{5}{20}, \frac{6}{20}), [\frac{14}{20}, \frac{15}{20}), [\frac{16}{20}, \frac{17}{20}), [\frac{18}{20}, \frac{19}{20})\} \\ 2 & \text{for } u \in \{[\frac{8}{20}, \frac{12}{20}), [\frac{13}{20}, \frac{14}{20}), [\frac{19}{20}, 1]\} \\ 3 & \text{for } u \in \{[0, \frac{5}{20}), [\frac{6}{20}, \frac{7}{20}), [\frac{12}{20}, \frac{13}{20}), [\frac{15}{20}, \frac{16}{20}), [\frac{17}{20}, \frac{18}{20})\} \end{cases}$$

and the time series $X_{t,T} = \sigma(\frac{t}{T})\varepsilon_t$. The results can be found in Tables 9 and 10.

The rejection rates for this model steadily increase as the sample size grows for all the DFT tests. Though the rejection rates for $\mathcal{T}_m$ tends to be slightly higher for larger m. This is because the size of the average covariances $|\bar{c}(r)|^2$ vary quite a lot in magnitude over r (see Table 10), and the larger values of $|\bar{c}(r)|^2$ are not concentrated around $r = 1$. One can also see why this is true by looking at a realisation of $T|\widehat{c}_T(r)|^2$ given in Figure 2, where we see that the magnitudes are highly variable over different lags. On the other hand the the proportion of rejects for the Paparoditis test is a lot less than the DFT tests.

Overall the DFT tests tend to perform consistently well for these model. But the max test $\mathcal{T}_m^*$ is more conservative than the sum test $\mathcal{T}_m$ and tends to reject the null less often. We observe that there is a some difference between how $\mathcal{T}_m$ can perform for different values of m. Therefore to select m it is worthwhile plotting $T|\hat{c}_T(r)|^2$ against r, as in Figure 1 and 2, which indicates how to choose m.

The DFT and Paparoditis test tend to comparable for Model 3 and 4. Though it seems that a large sample size is required for the Paparoditis test to detect the nonstationarity in Model 5 (this could be because the observations are independent).

7 Conclusions

In this article we have considered a test for second order stationarity. We proved the above result under the assumption that the time series is strictly stationary with absolutely summable covariances. But we believe this result is true even if the process is strictly stationary with long memory. However, the proof is beyond the scope of this paper and is future research. We believe a more indepth investigation of the real and imaginary parts of the DFT covariance $\widehat{c}_T(r)$ and $\widehat{c}_T(r, j)$ (defined in Remark 2.4) may give an insight into the nonstationary behaviour of the time series. Indeed $\widehat{c}_T(r, j)$ can be used to estimate the locally stationary spectral density $f(u, \omega)$. Simple two-dimensional plots of $\{\widehat{c}_T(r, j)\}$ may be a simple method for 'visualising' the nonstationarity. We also believe it is straightforward to extend our results to tests for second order stationarity of spatial random fields which are defined on a lattice.

Acknowledgements

The authors are grateful to Professor Manfred Deistler for several interesting discussions. This work has been partially supported by the NSF grant DMS-0806096.

	DFT sum			DFT max		Paparoditis test
	$\mathcal{T}_1$	$\mathcal{T}_5$	$\mathcal{T}_{10}$	$\mathcal{T}_5^*$	$\mathcal{T}_{10}^*$	$\mathcal{S}_T$
$T = 64$ 5% level	2.3	2.8	3.5	1.8	1.1	6.1
$T = 64$ 1% level	0.1	0.5	1	0.3	0.1	5.7
$T = 128$ 5% level	2.7	2.6	2.4	2	1.8	12.1
$T = 128$ 1% level	0.3	0.4	0.7	0	0.1	10.9
$T = 256$ 5% level	2.1	1.9	2.9	2.5	2.1	1.3
$T = 256$ 1% level	0.5	0.1	0.3	0.4	0.4	1
$T = 512$ 5% level	4.1	2.7	4	3	4.2	0.2
$T = 512$ 1% level	1	0.3	0.9	0.6	1.7	0.1

Table 1: The rejection rates for stationary Model 1 at the 5% and 1% level

$\|\bar{c}(1)\|^2$	$\|\bar{c}(2)\|^2$	$\|\bar{c}(3)\|^2$	$\|\bar{c}(4)\|^2$	$\|\bar{c}(5)\|^2$	$\|\bar{c}(6)\|^2$	$\|\bar{c}(7)\|^2$	$\|\bar{c}(8)\|^2$	$\|\bar{c}(9)\|^2$	$\|\bar{c}(10)\|^2$
0.004	0.004	0.003	0.003	0.004	0.004	0.004	0.004	0.004	0.004

Table 2: The average $|\hat{c}_T(r)|^2$ for stationary Model 1. $T = 512$

	DFT sum			DFT max		Paparoditis test
	$\mathcal{T}_1$	$\mathcal{T}_5$	$\mathcal{T}_{10}$	$\mathcal{T}_5^*$	$\mathcal{T}_{10}^*$	$\mathcal{S}_T$
$T = 64$ 5% level	4.6	7.1	8.5	3.1	2.3	41.3
$T = 64$ 1% level	1.8	4.2	6.5	1.4	0.7	39.2
$T = 128$ 5% level	4.4	6.3	8.3	4.2	4	76.2
$T = 128$ 1% level	1	3.2	5.4	1	1.3	74
$T = 256$ 5% level	5.3	6.1	6.9	5.1	5	83
$T = 256$ 1% level	1.3	4.1	5	1.8	1.7	81.8
$T = 512$ 5% level	3.7	4.3	4.6	4.4	3.6	23.6
$T = 512$ 1% level	0.7	1.8	1.7	0.9	0.9	20.2

Table 3: The rejection rates for stationary Model 2 at the 5% and 1% level

$\|\bar{c}(1)\|^2$	$\|\bar{c}(2)\|^2$	$\|\bar{c}(3)\|^2$	$\|\bar{c}(4)\|^2$	$\|\bar{c}(5)\|^2$	$\|\bar{c}(6)\|^2$	$\|\bar{c}(7)\|^2$	$\|\bar{c}(8)\|^2$	$\|\bar{c}(9)\|^2$	$\|\bar{c}(10)\|^2$
0.004	0.004	0.004	0.004	0.004	0.003	0.004	0.004	0.004	0.003

Table 4: The average $|\hat{c}_T(r)|^2$ for stationary Model 2. $T = 512$

	DFT sum			DFT max		Paparoditis test
	$\mathcal{T}_1$	$\mathcal{T}_5$	$\mathcal{T}_{10}$	$\mathcal{T}_5^*$	$\mathcal{T}_{10}^*$	$\mathcal{S}_T$
$T = 64$ 5% level	48.9	45.3	40.8	31.3	21.5	41.9
$T = 64$ 1% level	23.1	27.9	26.9	11.5	7.9	38.1
$T = 128$ 5% level	88.1	86.6	80	80.2	71.6	93.3
$T = 128$ 1% level	71.4	71.4	64.8	59.2	51.1	91.5
$T = 256$ 5% level	99.6	99.7	99.3	99.5	99.2	99.8
$T = 256$ 1% level	98.8	99.1	97.9	97.7	96.6	99.7
$T = 512$ 5% level	100	100	100	100	100	100
$T = 512$ 1% level	100	100	100	100	100	100

Table 5: The rejection rates for nonstationary Model 3 at the 5% and 1% level

$\lvert\bar{c}(1)\rvert^2$	$\lvert\bar{c}(2)\rvert^2$	$\lvert\bar{c}(3)\rvert^2$	$\lvert\bar{c}(4)\rvert^2$	$\lvert\bar{c}(5)\rvert^2$	$\lvert\bar{c}(6)\rvert^2$	$\lvert\bar{c}(7)\rvert^2$	$\lvert\bar{c}(8)\rvert^2$	$\lvert\bar{c}(9)\rvert^2$	$\lvert\bar{c}(10)\rvert^2$
0.111	0.057	0.016|	0.006	0.012|	0.012	0.008	0.006	0.008	0.009

Table 6: The average $|\hat{c}_T(r)|^2$ for nonstationary Model 3. $T = 512$

	DFT sum			DFT max		Paparoditis test
	$\mathcal{T}_1$	$\mathcal{T}_5$	$\mathcal{T}_{10}$	$\mathcal{T}_5^*$	$\mathcal{T}_{10}^*$	$\mathcal{S}_T$
$T = 64$ 5% level	3.1	3.8	5.5	1.9	2.2	9
$T = 64$ 1% level	0.1	1.1	2.7	0.1	0.3	8.3
$T = 128$ 5% level	97.9	95	84.7	93.9	89.4	89.9
$T = 128$ 1% level	90.7	79.6	67.3	77	68.8	87.2
$T = 256$ 5% level	58.1	100	100	100	100	93
$T = 256$ 1% level	35.8	100	100	100	100	89.1
$T = 512$ 5% level	100	100	100	100	100	100
$T = 512$ 1% level	100	100	100	100	100	100

Table 7: The rejection rates for nonstationary Model 4 at the 5% and 1% level

$T = 512$	$\lvert\bar{c}(1)\rvert^2$	$\lvert\bar{c}(2)\rvert^2$	$\lvert\bar{c}(3)\rvert^2$	$\lvert\bar{c}(4)\rvert^2$	$\lvert\bar{c}(5)\rvert^2$	$\lvert\bar{c}(6)\rvert^2$	$\lvert\bar{c}(7)\rvert^2$	$\lvert\bar{c}(8)\rvert^2$	$\lvert\bar{c}(9)\rvert^2$	$\lvert\bar{c}(10)\rvert^2$
	0.347	0.088	0.008|	0.006	0.006|	0.007	0.007	0.007	0.007	0.007
$T = 256$	$\lvert\bar{c}(1)\rvert^2$	$\lvert\bar{c}(2)\rvert^2$	$\lvert\bar{c}(3)\rvert^2$	$\lvert\bar{c}(4)\rvert^2$	$\lvert\bar{c}(5)\rvert^2$	$\lvert\bar{c}(6)\rvert^2$	$\lvert\bar{c}(7)\rvert^2$	$\lvert\bar{c}(8)\rvert^2$	$\lvert\bar{c}(9)\rvert^2$	$\lvert\bar{c}(10)\rvert^2$
	0.035	0.159	0.062	0.035	0.025	0.021	0.019	0.017	0.017	0.016
$T = 128$	$\lvert\bar{c}(1)\rvert^2$	$\lvert\bar{c}(2)\rvert^2$	$\lvert\bar{c}(3)\rvert^2$	$\lvert\bar{c}(4)\rvert^2$	$\lvert\bar{c}(5)\rvert^2$	$\lvert\bar{c}(6)\rvert^2$	$\lvert\bar{c}(7)\rvert^2$	$\lvert\bar{c}(8)\rvert^2$	$\lvert\bar{c}(9)\rvert^2$	$\lvert\bar{c}(10)\rvert^2$
	0.13	0.042	0.029|	0.024	0.022	0.022	0.021	0.022	0.021	0.02
$T = 64$	$\lvert\bar{c}(1)\rvert^2$	$\lvert\bar{c}(2)\rvert^2$	$\lvert\bar{c}(3)\rvert^2$	$\lvert\bar{c}(4)\rvert^2$	$\lvert\bar{c}(5)\rvert^2$	$\lvert\bar{c}(6)\rvert^2$	$\lvert\bar{c}(7)\rvert^2$	$\lvert\bar{c}(8)\rvert^2$	$\lvert\bar{c}(9)\rvert^2$	$\lvert\bar{c}(10)\rvert^2$
	0.027	0.026	0.026	0.026	0.027	0.027	0.027	0.028	0.026	0.027

Table 8: The average $|\hat{c}_T(r)|^2$ for nonstationary Model 4, calculated using $T = 512, 256, 128$ and 64.

A Appendix

In this appendix we prove the results from the main section.

	DFT sum			DFT max		Paparoditis test
	$\mathcal{T}_1$	$\mathcal{T}_5$	$\mathcal{T}_{10}$	$\mathcal{T}_5^*$	$\mathcal{T}_{10}^*$	$\mathcal{S}_T$
$T = 64$ 5% level	18.3	26.8	36	17.7	18.4	3.4
$T = 64$ 1% level	5.5	11.9	19.3	4.3	4.4	2.6
$T = 128$ 5% level	33.3	47.8	66.5	35.2	42.5	11.9
$T = 128$ 1% level	14.8	27.2	44.1	15.4	18.7	9.5
$T = 256$ 5% level	57.7	78.5	97.1	62.6	78.2	37.9
$T = 256$ 1% level	36	58.4	87.2	38.8	52.1	32.9
$T = 512$ 5% level	82.4	97.8	100	91.9	99.6	76.4
$T = 512$ 1% level	65.5	92.8	100	76.7	94.9	73

Table 9: The rejection rates for nonstationary Model 5 at the 5% and 1% level

$\lvert\bar{c}(1)\rvert^2$	$\lvert\bar{c}(2)\rvert^2$	$\lvert\bar{c}(3)\rvert^2$	$\lvert\bar{c}(4)\rvert^2$	$\lvert\bar{c}(5)\rvert^2$	$\lvert\bar{c}(6)\rvert^2$	$\lvert\bar{c}(7)\rvert^2$	$\lvert\bar{c}(8)\rvert^2$	$\lvert\bar{c}(9)\rvert^2$	$\lvert\bar{c}(10)\rvert^2$
0.027	0.023	0.007	0.016	0.009	0.007	0.029	0.01	0.028	0.005

Table 10: The average $|\hat{c}_T(r)|^2$ for nonstationary Model 5. $T = 512$

For short hand, when it is clear that T plays a role, we use the notation $J_k := J_T(\omega_k)$, $\bar{J}_k = J_T(-\omega_k)$, $\hat{f}_k := \hat{f}_T(\omega_k)$ and $f_k = f(\omega_k)$.

A.1 Some results on DFTs and Fourier coefficients

In the sections below, under various assumptions of the dependence of $\{X_t\}$, we will show asymptotic normality and obtain the mean and variance of $\widetilde{c}_T(r)$. In the case that $\{X_t\}$ is a short memory, stationary time series, then it is relatively straightforward to evaluate the variance of $\widetilde{c}_T(r)$, since $\{J_T(\omega_k)\overline{J_T(\omega_{k+r})}\}$ are close to uncorrelated random variables. However, in the nonstationary case this no longer holds and we have to use some results in Fourier analysis to derive $\mathrm{var}(\widetilde{c}_T(r))$. To do this we start by studying the general random variable

$$H_T = \frac{1}{T}\sum_{k=1}^{T} H(\omega_k) J_T(\omega_k)\overline{J_T(\omega_{k+r})},$$

where $J_T(\omega_k) = \frac{1}{\sqrt{2\pi T}}\sum_{t=1}^{T} X_t \exp(\frac{2\pi i t k}{T})$. We will show that under certain conditions on $H(\cdot)$, H_T can be written as the weighted average of X_t. Expanding $J_T(\omega_k)\overline{J_T(\omega_k)}$ we see that H_T can be written as

$$H_T = \frac{1}{T}\sum_{t,\tau} X_t X_\tau \exp(-i\omega_r \tau)\Big(\frac{1}{T}\sum_{k=1}^{T} H(\omega_k)\exp(i\omega_k(t-\tau))\Big).$$

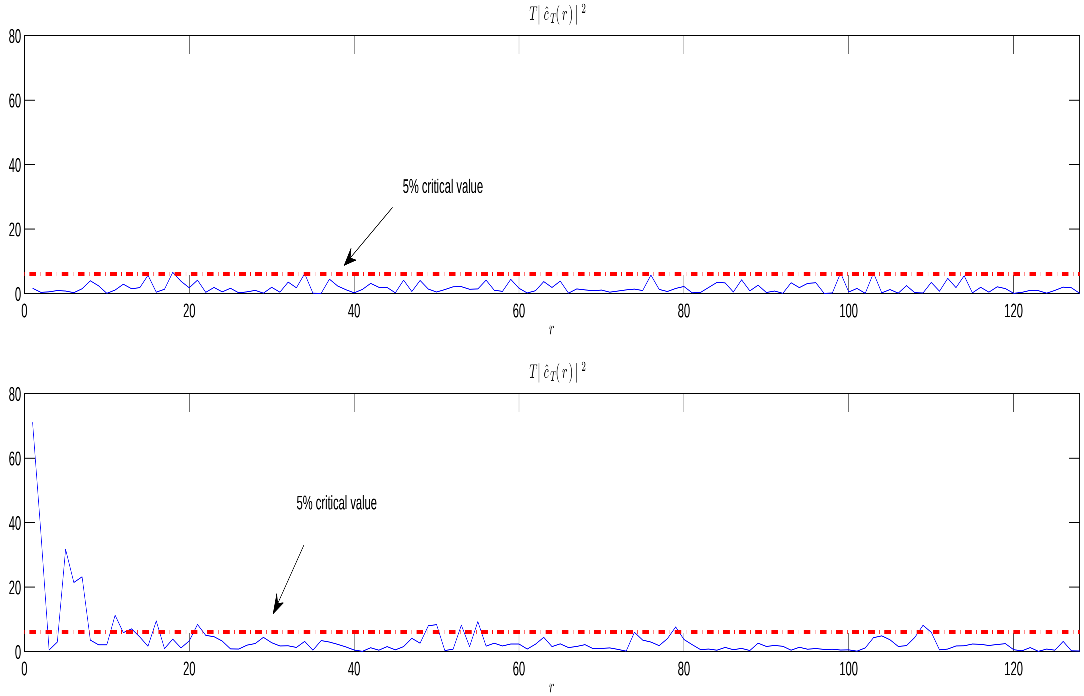

Figure 1: The top plot is the plot of the squared DFT covariance $T|\hat{c}_T(r)|^2$ from one realisation of the stationary time series model 1 ($T = 512$ and $r = 1, \ldots, 128$). The lower plot is the plot of the squared DFT covariance $T|\hat{c}_T(r)|^2$ from one realisation of the nonstationary time series model 3 ($T = 512$ and $r = 1, \ldots, 128$). The line is for $\chi_2^2(0.95)$. If several DFT squared covariances are above this line, this may indicate nonstationary behaviour

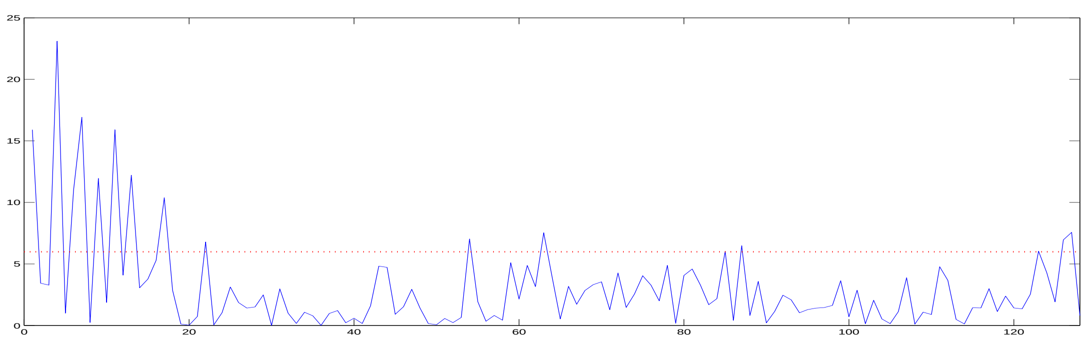

Figure 2: The plot is the plot of the squared DFT covariance $T|\hat{c}_T(r)|^2$ from one realisation of the nonstationary time series model 5 ($T = 512$ and $r = 1, \ldots, 128$). The line is for $\chi_2^2(0.95)$.
.

Without any smoothness assumptions on $H(\omega_k)$, it is not clear whether the inner sum of the above converges to zero as $|t - \tau| \to \infty$, and if the variance of H_T converges to zero. However, let us

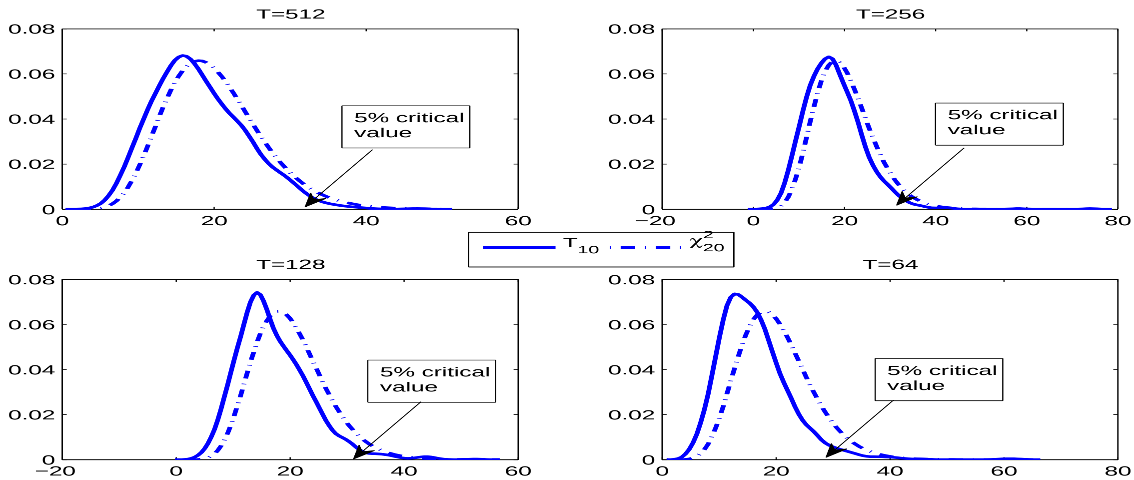

Figure 3: Thick line: Estimated density of test statistic of $\mathcal{T}_5$ for 1000 realisations of the stationary series model 1 ($T = 64, 128, 256, 512$). The dotted lines are the chi-squared with 10 degrees of freedom

suppose that $\sup_\omega |H'(\omega)| < \infty$. In this case, the DFT of $H(\omega)$, $\frac{1}{T}\sum_{k=1}^{T} H(\omega_k)\exp(i\omega_k(t-\tau))$, is an approximation of the Fourier coefficient $h(t-\tau) = \int H(\omega)\exp(i(t-\tau)\omega)d\omega$ (the error in this approximation is discussed below). Noting that the Fourier coefficients $h(k) \to 0$ as $k \to \infty$, we have

$$H_T \approx \frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T} X_t X_\tau h(t-\tau)\exp(-i\omega_r\tau).$$

Hence H_T can be approximated by a quadratic form, where the weights decay as $|t-\tau| \to \infty$. We now justify some of the approximations discussed above and state some well know results in Fourier analysis. An interesting overview of results in Fourier analysis is given in Briggs and Henson (1997).

The following theorem is well known, see for example, Briggs and Henson (1997), Theorem 6.2, for the proof.

Theorem A.1 *Suppose that $g[0,\Omega] \to \mathbb{R}$ is a periodic function with period Ω. We shall assume that either (a) $\sup_x |g''(x)| < \infty$ or (b) $\sup_x |g'(x)| < \infty$ and $g'(\cdot)$ is piecewise montone function. Let*

$$a(s) = \frac{1}{\Omega}\int_0^\Omega g(x)\exp(isx)dx \quad \text{and} \quad a_T(s) = \frac{1}{T}\sum_{k=1}^{T} g(\frac{\Omega k}{T})\exp(i\frac{\Omega k}{T}).$$

Therefore for all s, we have in case (a) $|a(s)| \le C\sup_x |g''(x)| s^{-2}$ and in case (b) $|a(s)| \le C\sup_x |g'(x)| s^{-2}$, where C is constant independent of s and $g(\cdot)$.

Moreover $\sup_{1\leq s\leq T}|a(s)-a_T(s)|\leq CT^{-2}$.

We now apply the above result to our setup. We will use Lemma A.1, below, to prove the asymptotic normality result in Section A.4.

Lemma A.1 *Suppose Assumption 3.1 or 4.1 is satisfied. And let $f(\omega)$ be the spectral density of the stationary linear time series or the integrated spectral density of the locally stationary time series. Let*

$$\begin{aligned} G_{T,\omega_r}(s) &= \frac{1}{T}\sum_{k=1}^{T}\frac{1}{(f(\omega_k)f(\omega_k+\omega_r))^{1/2}}\exp(is\omega_k) && (21)\\ \text{and } G_{\omega_r}(s) &= \frac{1}{2\pi}\int_0^{2\pi}\frac{1}{(f(\omega)f(\omega+\omega_r))^{1/2}}\exp(is\omega)d\omega. && (22)\end{aligned}$$

Under the stated assumptions we have either $(f(\omega)f(\omega+\omega_r))^{-1/2}$ has a bounded second derivative (under Assumption 3.1(iv-a) or Assumption 4.1) or that $(f(\omega)f(\omega+\omega_r))^{-1/2}$ has a bounded first derivative and is piecewise monotone (under Assumption 3.1(iv-b)). Therefore, $|G_{\omega_r}(s)|\leq Ks^{-2}$ (K is a finite function independent of ω_r), $\sup_{\omega_r}\sum_s|G_{\omega_r}(s)|<\infty$ and $\sup_{\omega_r}\sum_{s=1}^{T}|G_{T,\omega_r}(s)|<\infty$.

PROOF. The above is a straightforward application of Theorem A.1. □

We will use the result below in Sections A.2 and A.3.

Lemma A.2 *Suppose Assumption 4.1 is satisfied. For $2\leq n\leq 16$, define the nth order cumulant spectra as*

$$f_n(u,\omega_1,\ldots,\omega_{n-1})=\frac{1}{(2\pi)^{(n/2)-1}}\left\{\prod_{j=1}^{n-1}A(u,\omega_j)\right\}A(u,-\sum_{j=1}^{n-1}\omega_j),$$

and the Fourier transform

$$F_n(k;\omega_1,\ldots,\omega_{n-1})=\int_0^1 f_n(u,\omega_1,\ldots,\omega_{n-1})\exp(i2\pi ku)du. \tag{23}$$

(i) *If Assumption 4.1(iv)(a) holds, then $\sup_{u,\omega_1,\ldots,\omega_n}|\frac{\partial^2 f_n(u,\omega_1,\ldots,\omega_{n-1})}{\partial u^2}|<\infty$ and*

$$\sup_{\omega_1,\ldots,\omega_{n-1}}|F_n(k;\omega_1,\ldots,\omega_{n-1})|\leq C\sup_{u,\omega_1,\ldots,\omega_n}|\frac{\partial^2 f_n(u,\omega_1,\ldots,\omega_{n-1})}{\partial u^2}|\frac{1}{|k|^2}. \tag{24}$$

(ii) *If Assumption 4.1(iv)(b) holds, then $\sup_{u,\omega_1,\ldots,\omega_n}|\frac{\partial f_n(u,\omega_1,\ldots,\omega_{n-1})}{\partial u}|<\infty$, $\frac{\partial f_n(u,\omega_1,\ldots,\omega_{n-1})}{\partial u}$ is a* *piecewise monotone function in u and*

$$\sup_{\omega_1,\ldots,\omega_n}|F_n(k;\omega_1,\ldots,\omega_{n-1})| \leq C\sup_{u,\omega_1,\ldots,\omega_n}|\frac{\partial f_n(u,\omega_1,\ldots,\omega_{n-1})}{\partial u}|\frac{1}{|k|^2}. \tag{25}$$

We note that the constant C is independent of the function $f_n(\cdot)$ and k.

PROOF. We only prove (ii), the proof of (i) is similar. We note that

$$\begin{aligned}(2\pi)^{(n/2)-1}\frac{df_n(u,\omega_1,\ldots,\omega_{n-1})}{du} &= A(u,-\sum_{j=1}^{n-1}\omega_j)\sum_{r=1}^{n-1}\frac{dA(u,\omega_r)}{du}\prod_{j\neq r}A(u,\omega_j)+\\ &\quad \frac{\partial A(u,-\sum_{j=1}^{n-1}\omega_j)}{\partial u}\prod_{j=1}^{n-1}A(u,\omega_j).\end{aligned}$$

Now under Assumption 4.1, $\sup_{u,\omega}|A(u,\omega)|$ and $\sup_{u,\omega}|\frac{\partial A(u,\omega)}{\partial u}|$ are bounded function, hence we see from the above that $\sup_{u,\omega}|\frac{\partial f_n(u,\omega_1,\ldots,\omega_{n-1})}{\partial u}|$ is bounded. Moreover, by using Theorem A.1 we have (24). The proof of (ii) is similar, and we omit the details. □

We observe that in the stationary case $A(u,\omega)\equiv A(\omega)$, then $F_n(k;\omega_1,\ldots,\omega_{n-1})=0$ for $k\neq 0$.

In the following lemma we consider the error in approximation of the DFT with the Fourier coefficient.

Lemma A.3 *Suppose Assumption 4.1 is satisfied. Let $F_n(k;\omega_1,\ldots,\omega_{n-1})$ be defined as in (23) and let*

$$F_{n,T}(s;\omega_1,\ldots,\omega_{n-1}) = \frac{1}{T}\sum_{t=1}^{T}f_n(\frac{t}{T},\omega_1,\ldots,\omega_{n-1})\exp(i2s\pi t/T). \tag{26}$$

Then under Assumption 4.1(v)(a) we have

$$\sup_{\omega_1,\ldots,\omega_{n-1}}\big|F_{n,T}(s;\omega_1,\ldots,\omega_{n-1})-F_n(s;\omega_1,\ldots,\omega_{n-1})\big| \leq C\sup_{u,\omega_1,\ldots,\omega_{n-1}}|\frac{\partial^2 f_n(u,\omega_1,\ldots,\omega_{n-1})}{\partial u^2}|\frac{1}{T^2},$$

and under Assumption 4.1(v)(b) we have

$$\sup_{\omega_1,\ldots,\omega_{n-1}}\big|F_{n,T}(s;\omega_1,\ldots,\omega_{n-1})-F_n(s;\omega_1,\ldots,\omega_{n-1})\big| \leq C\sup_{u,\omega_1,\ldots,\omega_{n-1}}|\frac{\partial f_n(u,\omega_1,\ldots,\omega_{n-1})}{\partial u}|\frac{1}{T^2},$$

where C is a constant independent of $f_n(\cdot)$.

PROOF. The proof follows immediately from Theorem A.1 and Lemma A.2. □

A.2 Proof of Theorems 3.1 and 4.1

We first gives some results which connect the sum of cumulants and mixing. The result below is analogous to Ibragimov's inequality for higher order cumulants. It is motivated by Neumann (1996), Remark 3.1.

Lemma A.4 *Let us suppose that $\{X_t\}$ is a stationary time series and for some $\delta > 0$ we have $\sum_r |r|\alpha(r)^{\frac{\delta}{(k-1)(1+\delta)}} < \infty$ and $\mathbb{E}|X_t^{k(1+\delta)}| < \infty$. Then we have*

$$\sum_{t_2,\ldots,t_k} \prod_{i=2}^{k}(1+|t_i|)cum(X_{t_1},\ldots,X_{t_k})| \leq C_k\big(\mathbb{E}|X_t^{k(1+\delta)}|\big)^{1/(1+\delta)}\big(\sum_r |r|\alpha(r)^{\frac{1}{(k-1)(1+\delta)}}\big)^{k-1} < \infty, \quad (27)$$

where C_k is a finite constant which depends only on k.

PROOF. To prove the result we use a result from (Statulevicius & Jakimavicius, 1988), Theorem 3, part (2), which states that if $t_1 \leq t_2 \leq \ldots \leq t_k$, then for all $2 \leq i \leq k$ we have

$$\big|\mathrm{cum}(X_{t_1}, X_{t_2}, \ldots, X_{t_k})\big| \leq 3(k-1)!2^{k-1}\alpha(t_i - t_{i-1})^{\delta/(1+\delta)}\big(\mathbb{E}|X_t^{k(1+\delta)}|\big)^{1/(1+\delta)}.$$

Now by taking the $1/(k-1)$th root of the above for all $2 \leq i \leq k$ we have

$$\big|\mathrm{cum}(X_{t_1}, X_{t_2}, \ldots, X_{t_k})\big|^{1/(k-1)} \leq C_k^{1/(k-1)}\alpha(t_i - t_{i-1})^{\delta/(k-1)(1+\delta))}\big(\mathbb{E}|X_t^{k(1+\delta)}|\big)^{1/(k-1)(1+\delta)},$$

where $C_k = 3(k-1)!2^{k-1}$. Since the above bound holds for all i, multiplying the above over i gives

$$\big|\mathrm{cum}(X_{t_1}, X_{t_2}, \ldots, X_{t_k})\big| \leq \big(3(k-1)!2^{k-1}\big)\big(\mathbb{E}|X_t^{k(1+\delta)}|\big)^{1/(1+\delta)}\prod_{i=2}^{k}\alpha(t_i - t_{i-1})^{\delta/(k-1)(1+\delta)}. \quad (28)$$

Finally we rewrite $\sum_{t_2,\ldots,t_k=1}^{\infty}$ as the sum of orderings, that is $\sum_{t_2,\ldots,t_k=1}^{\infty} = \sum_{t_2\leq\ldots\leq t_k=1}^{\infty} + \ldots$. Now since the number of orderings is finite, we can use (28) to obtain

$$\sum_{t_2,\ldots,t_k=1}^{\infty}\big|\mathrm{cum}(X_{t_1}, X_{t_2}, \ldots, X_{t_k})\big| \leq C_k k!\big(\mathbb{E}|X_t^{k(1+\delta)}|\big)^{1/(1+\delta)}\big\{\sum_r |r|\alpha(r)^{\delta/((k-1)(1+\delta))}\big\}^{k-1} < \infty,$$

which gives result. □

Lemma A.5 *Suppose either Assumption 3.1(A or B) is satisfied. Then for $2 \leq n \leq 16$ we have*

$$\sum_{k_2,\ldots,k_n}(1+|\kappa_n|)|\kappa_n(k_1,\ldots,k_n)| < \infty, \quad (29)$$

and

$$cum(J_T(\omega_1),\ldots,J_T(\omega_n)) = \frac{(2\pi)^{(n/2)-1}}{T^{n/2}} f_n(\omega_1,\ldots,\omega_{n-1}) \sum_{t=1}^{T} \exp(it\sum_{j=1}^{n}\omega_j) + O(\frac{1}{T^{n/2}}), \qquad (30)$$

where f_n is the nth order cumulant spectral density and $\kappa_n(j_1,\ldots,j_{n-1}) = cum(X_t, X_{t+j_1},\ldots,X_{t+j_{n-1}})$.

PROOF. We first prove (29). For stationary linear time series, the proof is of the above is well known, for stationary time series which is mixing the result immediately follows from Lemma A.4. To proof of (30) follows immediately from Brillinger (1981), Theorem 4.3.2. □

The follow result is due to Paparoditis (2009), Lemma 6.2, and is a generalisation of Brillinger (1981), Theorem 4.3.2, to locally stationary time series.

Lemma A.6 (Paparoditis (2009), Lemma 6.2) *Suppose that Assumption 4.1 holds and let $f_n(\cdot)$ be defined as in (23). Then for $\omega_1,\ldots,\omega_n \in [0,2\pi]$ and $2 \le n \le 16$ we have*

$$cum(J_T(\omega_1),\ldots,J_T(\omega_n)) = \frac{(2\pi)^{(n/2)-1}}{T^{n/2}} \sum_{t=1}^{T} f_n(\frac{t}{T},\omega_1,\ldots,\omega_{n-1}) \exp(it\sum_{j=1}^{n}\omega_j) + O(\frac{(\log T)^{n-1}}{T^{n/2}}).$$

Now the following corollary immediately follows from Lemmas A.3 and A.6.

Corollary A.1 *Suppose that Assumption 4.1 holds, and let $F_n(\cdot)$ be defined as in (23). Then we have*

$$cum(J_T(\omega_{j_1}),\ldots,J_T(\omega_{j_n})) = \frac{(2\pi)^{(n/2)-1}}{T^{(n/2)-1}} F_n(j_1+\ldots+j_n;\omega_{j_1},\ldots,\omega_{j_{n-1}}) + O(\frac{(\log T)^{n-1}}{T^{n/2}} + \frac{1}{T^2})$$

and

$$\big|cum(J_T(\omega_{j_1}),\ldots,J_T(\omega_{j_n}))\big| \le C\frac{1}{T^{(n/2)-1}|j_1+\ldots+j_n|} + \frac{C(\log T)^{n-1}}{T^{n/2}} + \frac{C}{T^2},$$

where $\omega_{j_r} = 2\pi j_r/T$, $-T \le j_r \le T$ and C is constant independent of ω.

In the lemma below we use the well known result which represents moments in terms of cumulants Let us suppose that $\sup_t \mathbb{E}(|X_t|^n) < \infty$. Then we have

$$\mathbb{E}(X_{t_1},\ldots,X_{t_n}) = \sum_{\pi}\prod_{B\in\pi} \mathrm{cum}(X_i; i\in B), \qquad (31)$$

where π is a partition of $\{t_1,t_2,\ldots,t_n\}$ and the sum $\sum_\pi$ is done over all partitions of $\{t_1,t_2,\ldots,t_n\}$.

Lemma A.7 *Let $f(\cdot)$ denote the spectral density or integrated spectral density. Then we have* *(i) Suppose either Assumption 3.1 or Assumption 4.1 holds, then*

$$\mathbb{E}\bigg(J_{k_1}\bar{J}_{k_1+r}\bar{J}_{k_2}J_{k_2+r}\big\{\hat{f}_{k_1}\hat{f}_{k_1+r}-\mathbb{E}(\hat{f}_{k_1}\hat{f}_{k_1+r})\big\}\big\{\hat{f}_{k_2}\hat{f}_{k_2+r}-\mathbb{E}(\hat{f}_{k_2}\hat{f}_{k_2+r})\big\}\bigg) = \frac{C}{bT}\bigg(\frac{1}{|k_1-k_2|^2}+\frac{\log T}{T}\bigg),$$

(ii) Suppose either Assumption 3.1 or Assumption 4.1 holds, then

$$\mathbb{E}\bigg\{\hat{f}_k\hat{f}_{k+r}-\mathbb{E}(\hat{f}_k\hat{f}_{k+r})\bigg\}^4 \leq \frac{C}{(bT)^2}$$

(iii) Suppose Assumption 3.1(i) holds, then we have $\mathbb{E}\big\{\hat{f}_k\hat{f}_{k+r}\big\}-f_kf_{k+r}=O(b+\frac{1}{bT})$. *Suppose Assumption 3.1(iv-a) or Assumption 4.1 holds, then we have* $\mathbb{E}\big\{\hat{f}_k\hat{f}_{k+r}\big\}-f_kf_{k+r}=O(b^2+\frac{1}{bT})$.

(iv) Suppose Assumption 4.1 holds, then we have

$$\begin{aligned}\mathbb{E}(J_{s_1}J_{s_2}J_{s_3}J_{s_4}) \;&\leq\; C\big(\frac{1}{|s_1+s_2|^2}+\delta_T\big)\big(\frac{1}{|s_3+s_4|^2}+\delta_T\big)+C\big(\frac{1}{|s_1+s_3|^2}+\delta_T\big)\big(\frac{1}{|s_2+s_4|^2}+\delta_T\big)\\ &+C\big(\frac{1}{|s_1+s_4|^2}+\delta_T\big)\big(\frac{1}{|s_2+s_3|^2}\big)+\delta_T\big)\\ &+\frac{C}{T}\frac{1}{|s_1+s_2+s_3+s_4|^2}+O\big(\frac{(\log T)^3}{T^2}+\frac{1}{T^2}\big),\end{aligned}$$

where $\delta_T=O(\frac{\log T}{T}+\frac{1}{T^2})$.

PROOF. To simplify notation in the proof we let $w_j=\frac{1}{bT}K(\frac{\omega_j}{b})$. By expanding the expectation in (i) we have

$$\begin{aligned}&\mathbb{E}\bigg(J_{k_1}\bar{J}_{k_1+r}\bar{J}_{k_2}J_{k_2+r}\big\{\hat{f}_{k_1}\hat{f}_{k_1+r}-\mathbb{E}(\hat{f}_{k_1}\hat{f}_{k_1+r})\big\}\big\{\hat{f}_{k_2}\hat{f}_{k_2+r}-\mathbb{E}(\hat{f}_{k_2}\hat{f}_{k_2+r})\big\}\bigg)\\ =\;&\sum_{j_1,j_2,j_3,j_4} w_{j_1}w_{j_2}w_{j_3}w_{j_4}\\ &\bigg\{\mathbb{E}(J_{k_1-j_1}\bar{J}_{k_1-j_1}J_{k_1+r-j_2}\bar{J}_{k_1+r-j_2}J_{k_1}\bar{J}_{k_1+r}\bar{J}_{k_2}J_{k_2+r}J_{k_2-j_3}\bar{J}_{k_2-j_3}J_{k_2+r-j_4}\bar{J}_{k_2+r-j_4})\\ &-\mathbb{E}(J_{k_1-j_1}\bar{J}_{k_1-j_1}J_{k_1+r-j_2}\bar{J}_{k_1+r-j_2})\mathbb{E}(J_{k_1}\bar{J}_{k_1+r}\bar{J}_{k_2}J_{k_2+r}J_{k_2-j_3}\bar{J}_{k_2-j_3}J_{k_2+r-j_4}\bar{J}_{k_2+r-j_4})\\ &-\mathbb{E}(J_{k_2-j_3}\bar{J}_{k_2-j_3}J_{k_2+r-j_4}\bar{J}_{k_2+r-j_4})\mathbb{E}(J_{k_1}\bar{J}_{k_1+r}\bar{J}_{k_2}J_{k_2+r}J_{k_1-j_1}\bar{J}_{k_1-j_1}J_{k_1+r-j_2}\bar{J}_{k_1+r-j_2})\\ &+\mathbb{E}(J_{k_1-j_3}\bar{J}_{k_1-j_3}J_{k_1+r-j_4}\bar{J}_{k_1+r-j_4})\mathbb{E}(J_{k_1}\bar{J}_{k_1+r}\bar{J}_{k_2}J_{k_2+r})\mathbb{E}(J_{k_2-j_3}\bar{J}_{k_2-j_3}J_{k_2+r-j_4}\bar{J}_{k_2+r-j_4})\bigg\}. \quad (32)\end{aligned}$$

To prove the result we first represent the above moments in terms of cumulants using (31). We observe that many of the terms will cancel, however those that do remain will involve at least one cumulant which has elements belong to the set $\{J_{k_1-j_1},\bar{J}_{k_1-j_1},J_{k_1+r-j_2},\bar{J}_{k_1+r-j_2}\}$ and the set $\{J_{k_2-j_3},\bar{J}_{k_2-j_3},J_{k_2+r-j_4},\bar{J}_{k_2+r-j_4}\}$, since these terms can only arise in the cumulant expansion of $\mathbb{E}(J_{k_1-j_1}\bar{J}_{k_1-j_1}J_{k_1+r-j_2}\bar{J}_{k_1+r-j_2}J_{k_1}\bar{J}_{k_1+r}\bar{J}_{k_2}J_{k_2+r}J_{k_2-j_3}\bar{J}_{k_2-j_3}J_{k_2+r-j_4}\bar{J}_{k_2+r-j_4})$. Now by a careful analysis of all cumulants involving elements from both these two sets we observe that the largest cumulant terms are $\mathrm{cum}(J_{k_1-j_1}, J_{k_2-j_3})$ and $\mathrm{cum}(\bar{J}_{k_1-j_1}, \bar{J}_{k_2-j_3})$ (the rest are of a lower order). Therefore recalling that $\sum_{j_1,j_2,j_3,j_4} w_{j_1}w_{j_2}w_{j_3}w_{j_4} = \sum_{j_1,j_2,j_3,j_4=1}^{bT} \frac{1}{(bT)^4}\prod_j K(\frac{\omega_j}{b})$ and using Corollary A.1 gives

$$\begin{aligned}
&\mathbb{E}\Big(J_{k_1}\bar{J}_{k_1+r}\bar{J}_{k_2}J_{k_2+r}\big\{\hat{f}_{k_1}\hat{f}_{k_1+r} - \mathbb{E}(\hat{f}_{k_1}\hat{f}_{k_1+r})\big\}\big\{\hat{f}_{k_2}\hat{f}_{k_2+r} - \mathbb{E}(\hat{f}_{k_2}\hat{f}_{k_2+r})\big\}\Big) \\
\leq\ & \sum_{j_1,j_2,j_3,j_4=1}^{bT} \frac{1}{(bT)^4}\prod_{j_i} K(\frac{\omega_{j_i}}{b})\big(\frac{1}{|k_1+k_2-j_1-j_3|^2} + \frac{\log T}{T}\big)^2 \\
\leq\ & \frac{C}{bT}\Big(\frac{1}{|k_1-k_2|^2} + \frac{\log T}{T}\Big),
\end{aligned}$$

where C is a finite constant independent of k_1 and k_2.

The proof of (ii) is similar to the proof of (i), hence we omit the details. We note that if we were to show asymptotic normality of $\widehat{f}_k\widehat{f}_{k+r}$, then (ii) would immediately follow from this.

We now prove (iii). By definition of $\hat{f}_k\hat{f}_{k+r}$, using Lemmas A.6 and A.1, under Assumption 3.1(i) we have

$$\begin{aligned}
\mathbb{E}\big\{\hat{f}_k\hat{f}_{k+r}\big\} - f_kf_{k+r} &= \sum_{j_1,j_2}\frac{1}{(bT)^2}K(\frac{w_{j_1}}{b})K(\frac{w_{j_2}}{b})\mathbb{E}\big\{J_{k-j_1}\bar{J}_{k-j_1}J_{k+r-j_2}\bar{J}_{k+r-j_2}\big\} - f(\omega_k)f(\omega_{k+r}) \\
&= \sum_{j_1}\frac{1}{bT}K(\frac{w_{j_1}}{b})f_2(\omega_{k-j_1})\sum_{j_2}\frac{1}{bT}K(\frac{w_{j_2}}{b})f_2(\omega_{k+r-j_1}) - f(\omega_k)f(\omega_{k+r}) + O(\frac{1}{bT}) \\
&= O(b+\frac{1}{bT}).
\end{aligned}$$

Using a similar proof we can show that under Assumption 3.1(iv-a) or Assumption 4.1 we have $\mathbb{E}\big\{\hat{f}_k\hat{f}_{k+r}\big\} - f_kf_{k+r} = O(b^2+\frac{1}{bT})$. Thus proving (iii).

The proof of (iv) uses Lemma A.6 and Corollary A.1 and is straightforward, hence we omit the details. □

In the lemma below we prove Theorems 3.1 and 4.1.

Lemma A.8 *Suppose Assumption 3.1 holds. Then we have*

$$\sqrt{T}|\widehat{c}_T(r) - \widetilde{c}_T(r)| = O\Big(\frac{1}{\sqrt{bT}} + (b+\frac{1}{bT}) + \big(\frac{1}{bT^{1/2}} + b^2T^{1/2}\big)\big(\frac{1}{|r|} + \frac{1}{T^{1/2}}\big)\Big). \tag{33}$$

Suppose Assumption 4.1 holds. Then we have

$$\sqrt{T}|\widehat{c}_T(r) - \widetilde{c}_T(r)| = O\Big(\frac{1}{\sqrt{bT}} + \big(\frac{1}{bT^{1/2}} + b^2T^{1/2}\big)\big(\frac{1}{|r|} + \frac{1}{T^{1/2}}\big)\Big). \tag{34}$$

PROOF. The proofs of (33) and (34) are very similar. Most of the time we will be obtaining the bounds under Assumption 4.1, however in a few places the bounds under Assumption 3.1 can be better than those under Assumption 4.1. In this case we will obtain the bounds under each of the Assumptions (separately).

To prove both (33) and (34) we first note that by the mean value theorem evaluated to the second order we have $x^{-1/2} - y^{-1/2} = (-1/2)x^{-3/2}(x-y) + (1/2)(-1/2)(-3/2)x_y^{-5/2}(x-y)^2$, where x_y lies between x and y. Applying this to the difference $\widehat{c}_T(r) - \widetilde{c}_T(r)$ we have the expansion

$$\sqrt{T}|\widehat{c}_T(r) - \widetilde{c}_T(r)| \leq \frac{1}{2}I + \frac{3}{8}II,$$

where

$$I = \frac{1}{\sqrt{T}}\sum_k \frac{J_k\bar{J}_{k+r}}{(f_k f_{k+r})^{3/2}}\big\{\hat{f}_k\hat{f}_{k+r} - f_k f_{k+r}\big\} \text{ and } II = \frac{1}{\sqrt{T}}\sum_k \frac{J_k\bar{J}_{k+r}}{(\bar{f}_k\bar{f}_{k+r})^{5/2}}\big\{\hat{f}_k\hat{f}_{k+r} - f_k f_{k+r}\big\}^2.$$

We consider the terms I and II separately. We first obtain a bound for $\mathbb{E}|I^2|$. Observe that $\mathbb{E}|I^2| = 3\big(\mathbb{E}|I_1^2| + \mathbb{E}|I_2^2|\big)$, where

$$\begin{aligned} I_1 &= \frac{1}{\sqrt{T}}\sum_k \frac{J_k\bar{J}_{k+r}}{(f_k f_{k+r})^{3/2}}\big\{\hat{f}_k\hat{f}_{k+r} - \mathbb{E}(\hat{f}_k\hat{f}_{k+r})\big\} \\ I_2 &= \frac{1}{\sqrt{T}}\sum_k \frac{J_k\bar{J}_{k+r}}{(f_k f_{k+r})^{3/2}}\big\{\mathbb{E}(\hat{f}_k\hat{f}_{k+r}) - f_k f_{k+r}\big\}, \end{aligned}$$

hence we will obtain the bounds $\mathbb{E}(I_1^2)$ and $\mathbb{E}(I_2^2)$. Expanding $\mathbb{E}(I_1^2)$ and using Lemma A.7(i) and that $f_{k_1}f_{k_1+r}$ is bounded away from zero gives

$$\mathbb{E}(I_1^2) \leq \frac{1}{T}\sum_{k_1,k_2} \frac{1}{(f_{k_1}f_{k_1+r}f_{k_2}f_{k_2+r})^{3/2}}\frac{C}{bT}\Big\{\frac{1}{|k_1-k_2|^2} + \frac{\log T}{T}\Big\} = O(\frac{1}{bT}). \tag{35}$$

Expanding $\mathbb{E}|I_2^2|$ gives

$$\begin{aligned} \mathbb{E}(I_2^2) \;&\leq\; \frac{C}{T}\sum_{k_1,k_2} \frac{1}{(f_{k_1}f_{k_1+r}f_{k_2}f_{k_2+r})^{3/2}}\mathbb{E}(J_{k_1}\bar{J}_{k_1+r}\bar{J}_{k_2}J_{k_2+r}) \times \\ &\quad \Big(\mathbb{E}\big\{\hat{f}_{k_1}\hat{f}_{k_1+r}\big\} - f_{k_1}f_{k_1+r}\Big)\Big(\mathbb{E}\big\{\hat{f}_{k_2}\hat{f}_{k_2+r}\big\} - f_{k_2}f_{k_2+r}\Big). \end{aligned} \tag{36}$$

The bounds for $\mathbb{E}|I_2^2|$ differ slightly, depending on the assumption. Under Assumption 3.1, by using Brillinger (1981), Theorem 4.3.2, it can be shown that $\mathbb{E}(J_{k_1}\bar{J}_{k_1+r}\bar{J}_{k_2}J_{k_2+r}) = O(T^{-1})$ (since $r \neq 0$). Moreover, by using Lemma A.7(iii) we have $\mathbb{E}\big\{\hat{f}_{k_1}\hat{f}_{k_1+r}\big\} - f_{k_1}f_{k_1+r} = O(b + (bT)^{-1})$. Therefore under Assumption 3.1 we have

$$\mathbb{E}(I_2^2) = O\big((b+\frac{1}{bT})^2\big). \tag{37}$$

Therefore, under Assumption 3.1, using (35) and (37) gives $\mathbb{E}|I|^2 = O(\frac{1}{bT}+(b+\frac{1}{bT})^2)$ and

$$I = O_p(b+\frac{1}{bT}+\frac{1}{\sqrt{bT}}). \tag{38}$$

On the other hand, under Assumption 4.1 we do not have that $\mathbb{E}(J_{k_1}\bar{J}_{k_1+r}\bar{J}_{k_2}J_{k_2+r}) = O(T^{-1})$, instead we substitute Lemma A.7(iv) into (36) and obtain

$$\mathbb{E}(I_2^2) \ \ = \ \ O\big(\{\frac{T}{r^2}+1\}(b^2+\frac{1}{bT})^2\big). \tag{39}$$

Therefore, under Assumption 4.1, using (35) and (39) gives $\mathbb{E}|I|^2 = O(\frac{1}{bT}+(\frac{T}{r^2}+1)(b^2+\frac{1}{bT})^2)$ and

$$I = O_p\bigg(\frac{1}{\sqrt{bT}}+\big(\frac{\sqrt{T}}{|r|}+1\big)(b^2+\frac{1}{bT})\bigg). \tag{40}$$

We now obtain a bound for II. Since the spectral density $f(\omega)$ is bounded away from zero and $\sup_\omega|\widehat{f}_T(\omega)-f(\omega)| \stackrel{P}{\to} 0$ (see (Paparoditis, 2009), Lemma 6.1(ii)), we have $II = O_p(\tilde{I}I)$, where

$$\tilde{I}I = \frac{1}{\sqrt{T}}\sum_k \big(J_k\bar{J}_{k+r}\big\{\hat{f}_k\hat{f}_{k+r}-f_kf_{k+r}\big\}^2\big).$$

To obtain a bound for $\tilde{I}I$ we use that $|\tilde{I}I| \leq 3\tilde{I}I_1+3\tilde{I}I_2$, where

$$\tilde{I}I_1 \ \ = \ \ \frac{1}{\sqrt{T}}\sum_k\big(|J_k\bar{J}_{k+r}|\big\{\hat{f}_k\hat{f}_{k+r}-\mathbb{E}(\hat{f}_k\hat{f}_{k+r})\big\}^2\big) \text{ and } \tilde{I}I_2 = \frac{1}{\sqrt{T}}\sum_k\big(|J_k\bar{J}_{k+r}|\big\{\mathbb{E}(\hat{f}_k\hat{f}_{k+r})-f_kf_{k+r}\big\}^2\big).$$

Using Cauchy-Schwarz inequality, Lemma A.7(ii,iv) we have

$$\mathbb{E}|\tilde{I}I_1| \ \ \leq \ \ \frac{1}{\sqrt{T}}\sum_k\mathbb{E}\big(|J_k\bar{J}_{k+r}|^2\big)^{1/2}\big(\mathbb{E}\big\{\hat{f}_k\hat{f}_{k+r}-\mathbb{E}(\hat{f}_k\hat{f}_{k+r})\big\}^4\big)^{1/2} = O\big(\frac{T^{1/2}}{bT}\big(\frac{1}{|r|}+\frac{1}{T^{1/2}}\big)\big).$$

We now obtain a bound for $\mathbb{E}|\tilde{I}I_2|$. Using Lemma A.7(iii,iv) we have

$$\mathbb{E}(\tilde{I}I_2) \ \ \leq \ \ \frac{1}{\sqrt{T}}\sum_k\mathbb{E}\big(|J_k\bar{J}_{k+r}|^2\big)^{1/2}\big\{\mathbb{E}(\hat{f}_k\hat{f}_{k+r})-f_kf_{k+r}\big\}^2 = O(\sqrt{T}(b+\frac{1}{bT})^2\big(\frac{1}{|r|}+\frac{1}{T^{1/2}}\big)).$$

Therefore

$$\tilde{I}I = O\Big((\frac{1}{|r|} + \frac{1}{T^{1/2}})\frac{1}{bT^{1/2}} + b^2T^{1/2} + \frac{1}{b^2T^{5/2}} + T^{-1/2}\Big). \qquad (41)$$

Hence (41) and (38) give (33) and (41) and (40) give (34). $\square$

A.3 The expectation and variance of the covariance $\tilde{c}_T(r)$

A.3.1 The moments under the null of strict stationarity

It is straightforward to show, under Assumption 3.1, that $\mathbb{E}(\sqrt{T}\tilde{c}_T(r)) = O(T^{-1/2})$. We now obtain the asymptotic variance of the $\tilde{c}_T(r)$ under the null of stationarity.

The following lemma immediately follows from (Brillinger, 1981), Theorem 4.3.2. We use this result to obtain the asymptotic variance of $\widetilde{c}_T(r)$, below.

Lemma A.9 *Let $\{X_t\}$ be a stationary time series where we denote the second and fourth order cumulants as κ_2 and κ_4. Suppose $\sum_k (1+|k|)|\kappa_2(k)| < \infty$ and $\sum_{k_1,k_2,k_3}(1+|k_i|)|\kappa_4(k_1,k_2,k_3)| < \infty$. Then we have*

$$\begin{aligned} cov(J_{k_1}J_{k_2}, J_{k_3}J_{k_4}) &= \Big(\frac{f(\omega_{k_1})}{T}\sum_{t=1}^{T} e^{it\omega_{k_1-k_3}} + O(\frac{1}{T})\Big)\Big(\frac{f(\omega_{k_2})}{T}\sum_{t=1}^{T} e^{it\omega_{k_2-k_4}} + O(\frac{1}{T})\Big) \\ &\quad + \Big(\frac{f(\omega_{k_1})}{T}\sum_{t=1}^{T} e^{it\omega_{k_1-k_4}} + O(\frac{1}{T})\Big)\Big(\frac{f(\omega_{k_2})}{T}\sum_{t=1}^{T} e^{it\omega_{k_2-k_3}} + O(\frac{1}{T})\Big) \\ &\quad + (2\pi)f_4(\omega_{k_1},\omega_{k_2},-\omega_{k_3})\frac{1}{T^2}\sum_{t=1}^{T} e^{it\omega_{k_1+k_2-k_3-k_4}} + O(\frac{1}{T^2}). \end{aligned} \qquad (42)$$

Lemma A.10 *Suppose the assumptions in Lemma A.9 hold. Then we have*

$$\begin{aligned} & cov\big(\sqrt{T}\Re\widetilde{c}_T(r_1), \sqrt{T}\Re\widetilde{c}_T(r_2)\big) = cov\big(\sqrt{T}\Im\widetilde{c}_T(r_1), \sqrt{T}\Im\widetilde{c}_T(r_2)\big) \\ = & \left\{ \begin{array}{cc} O(T^{-1}) & r_1 \neq r_2 \\ 1 + \frac{1}{2}\frac{2\pi}{2T^2}\sum_{k_1,k_2=1}^{T} g_{T,k_1}^{(r)} g_{T,k_2}^{(r)} f_4(\omega_{k_1}, -\omega_{k_1+r}, -\omega_{k_2}) + O(\frac{1}{T}) & r_1 = r_2 = r \end{array} \right. \end{aligned} \qquad (43)$$

PROOF. To prove the result we use that $\Re\widetilde{c}_T(r) = \frac{1}{2}(\widetilde{c}_T(r) + \bar{\widetilde{c}}_T(r))$ and $\Im\widetilde{c}_T(r) = \frac{-i}{2}(\widetilde{c}_T(r) + \bar{\widetilde{c}}_T(r))$, and evaluate $\mathrm{cov}(\sqrt{T}\tilde{c}_T(r_1), \sqrt{T}\tilde{c}_T(r_2))$, $\mathrm{cov}(\sqrt{T}\tilde{c}_T(r_1), \sqrt{T}\bar{\tilde{c}}_T(r_2))$ and $\mathrm{cov}(\sqrt{T}\bar{\tilde{c}}_T(r_1), \sqrt{T}\bar{\tilde{c}}_T(r_2))$. Expanding $\mathrm{cov}(\sqrt{T}\tilde{c}_T(r_1), \sqrt{T}\tilde{c}_T(r_2))$ gives

$$\mathrm{cov}(\sqrt{T}\tilde{c}_T(r_1), \sqrt{T}\tilde{c}_T(r_2)) = \frac{1}{T}\sum_{k_1,k_2=1}^{T} g_{T,k_1}^{(r_1)} g_{T,k_2}^{(r_2)} \mathrm{cov}(J_{k_1}\overline{J}_{k_1+r_1}, J_{k_2}\overline{J}_{k_2+r_2}),$$

where $g_{T,k}^{(r)} = f(\omega_k)^{-1/2} f(\omega_k + \omega_r)^{-1/2}$. Substituting (42) into the above it is easy to show that for $r_1 \neq r_2$ we have $\text{cov}(\sqrt{T}\tilde{c}_T(r_1), \sqrt{T}\tilde{c}_T(r_2)) = O(T^{-1})$ and for $r := r_1 = r_2$ we have

$$\begin{aligned}
& \text{cov}(\sqrt{T}\tilde{c}_T(r), \sqrt{T}\tilde{c}_T(r)) \\
= \; & \frac{2}{T}\sum_{k=1}^{T} (g_{T,k}^{(r)})^2 f(\omega_k) f(\omega_{k+r}) + \frac{2\pi}{T^2} \sum_{k_1,k_2=1}^{T} g_{T,k_1}^{(r)} g_{T,k_2}^{(r)} f_4(\omega_{k_1}, -\omega_{k_1+r}, -\omega_{k_2}) + O(\frac{1}{T}) \\
= \; & 2 + \frac{2\pi}{T^2} \sum_{k_1,k_2=1}^{T} g_{T,k_1}^{(r)} g_{T,k_2}^{(r)} f_4(\omega_{k_1}, -\omega_{k_1+r}, -\omega_{k_2}) + O(\frac{1}{T}).
\end{aligned}$$

The same method gives us a similar bound for $\text{cov}(\overline{\sqrt{T}\tilde{c}_T(r)}, \overline{\sqrt{T}\tilde{c}_T(r)})$. Similarly it can be shown that unless $r_2 = T - r_1$ we have $\text{cov}(\sqrt{T}\tilde{c}_T(r_1), \overline{\sqrt{T}\tilde{c}_T(r_2)}) = O(T^{-1})$. Also, for $r_1 \neq r_2$ we have $\text{cov}(\overline{\sqrt{T}\tilde{c}_T(r_1)}, \overline{\sqrt{T}\tilde{c}_T(r_2)}) = O(T^{-1})$ Altogether this gives us (43). □

PROOF of Lemma 3.1 Under the stated assumptions the spectral density f and the tri-spectra $f_4(\omega_1, \omega_2, \omega_3)$ is Lipschitz continuous, therefore using Lemma A.10 we have

$$\begin{aligned}
& \text{cov}\big(\sqrt{T}\Re\widetilde{c}_T(r_1), \sqrt{T}\Re\widetilde{c}_T(r_2)\big) = \text{cov}\big(\sqrt{T}\Im\widetilde{c}_T(r_1), \sqrt{T}\Im\widetilde{c}_T(r_2)\big) \\
= \; & \begin{cases} O(T^{-1}) & r_1 \neq r_2 \\ 1 + \frac{1}{4\pi}\int_0^{2\pi}\int_0^{2\pi} \frac{f_4(\omega_1, -\omega_1 - \omega_r, -\omega_2)}{\sqrt{f(\omega_1)f(\omega_1+\omega_r)f(\omega_2)f(\omega_2+\omega_r)}} d\omega_1 d\omega_2 + O(\frac{1}{T}) & r_1 = r_2 = r \end{cases}
\end{aligned} \tag{44}$$

and for all r_1, r_2, $\text{cov}(\sqrt{T}\Re\widetilde{c}_T(r_1), \sqrt{T}\Im\widetilde{c}_T(r_2)) = O(\frac{1}{T})$. And we have the required result. □

PROOF of Remark 3.1 It is well know that for a linear time series the tri-spectra can be written in terms of the transfer function $A(\omega)$ that is

$$f_4(\omega_1, \omega_2, \omega_3) = \frac{\kappa_4}{2\pi} A(\omega_1)A(\omega_2)A(\omega_3)A(-\omega_1 - \omega_2 - \omega_3).$$

Now we recall that for a linear time series $A(\omega) = \sqrt{f(\omega)}\exp(i\phi(\omega))$ hence substituting this into the ratio in (43) gives

$$2\pi \frac{f_4(\omega_{k_1}, -\omega_{k_2} - \omega_r, -\omega_{k_2})}{\sqrt{f(\omega_{k_1})f(\omega_{k_1} + \omega_2)f(\omega_{k_2})f(\omega_{k_2} + \omega_r)}} = \kappa_4 \exp\big(i[\phi(\omega_{k_1}) - \phi(\omega_{k_1} + \omega_r) - \phi(\omega_{k_2}) + \phi(\omega_{k_2} + \omega_r)]\big).$$

Substituting the above into (43) gives

$$T\text{cov}(\widetilde{c}_T(r), \widetilde{c}_T(r)) = 1 + \frac{\kappa_4}{2}\left|\frac{1}{T}\sum_k \exp(i[\phi(\omega_k) - \phi(\omega_k + \omega_r)])\right|^2 + O(\frac{1}{T}).$$

Finally we note that under Assumption 3.1(i,iii) and using the chain rule on $\phi(\omega) = \arctan\frac{\Im A(\omega)}{\Re A(\omega)}$ we can show that $\sup_\omega |\phi'(\omega)| < \infty$. Hence the phase $\phi(\cdot)$ is Lipschitz continuous, thus $\exp(i\phi(\omega))$

is Lipschitz continuous, and we can replace the summand in the above with an integral to give the result. $\square$

A.3.2 The moments under the alternative of local stationarity

We now consider some of the moment properties of $\widetilde{c}_T(r)$ under the assumption of local stationarity.

Lemma A.11 *Suppose Assumption 4.1 holds. Then we have*

$$
\begin{aligned}
cov(J_{k_1}J_{k_2}, J_{k_3}J_{k_4}) &= \big\{F_2(k_1-k_3;\omega_{k_1})F_2(k_2-k_4;\omega_{k_2}) + F_2(k_1-k_4;\omega_{k_1})F_2(k_2-k_3;\omega_{k_2})\big\} \\
&+ \frac{(2\pi)}{T}F_4(k_1+k_2-k_3-k_3;\omega_{k_1},\omega_{k_2},-\omega_{k_3}) \\
&+ O\big(\frac{(\log T)^3}{T^2} + \frac{\log T}{T}\big(F_2(k_1-k_3;\omega_{k_1}) + F_2(k_2-k_4;\omega_{k_2}) + \\
&\quad F_2(k_1-k_4;\omega_{k_1}) + F_2(k_2-k_3;\omega_{k_2})\big)\big), \qquad (45)
\end{aligned}
$$

where $\{F_2(\cdot;\omega)\}$ and $\{F_4(\cdot;\omega_{k_1},\omega_{k_2},\omega_{k_3})\}$ are defined in (26).

PROOF. Expanding $\mathrm{cov}(J_{k_1}J_{k_2}, J_{k_3}J_{k_4})$ in terms of cumulants gives

$$
\mathrm{cov}(J_{k_1}J_{k_2}, J_{k_3}J_{k_4}) = \mathrm{cov}(J_{k_1},\bar{J}_{k_3})\mathrm{cov}(J_{k_2},\bar{J}_{k_4}) + \mathrm{cov}(J_{k_1},\bar{J}_{k_4})\mathrm{cov}(J_{k_2},\bar{J}_{k_3}) + \mathrm{cum}(J_{k_1},J_{k_2},\bar{J}_{k_3},\bar{J}_{k_4}),
$$

finally substituting Corollary A.1 into the above gives the result. $\square$

PROOF of Lemma 4.1 equations (15) and (7) We first prove (15). We note that from the definition of $\widehat{f}_T(\omega_k)$ in (2) and under Assumption 4.1 we have

$$
\big|\mathbb{E}(\widehat{f}_T(\omega_k)) - f(\omega_k)\big| = \big|\sum_j \frac{1}{bT}K(\frac{\omega_k-\omega_j}{b})\mathbb{E}(|J_T(\omega_j)|^2) - f(\omega_k)\big| = O(b^2). \qquad (46)
$$

We now obtain $\mathrm{var}(\widehat{f}_T(\omega_k))$. We observe that

$$
\mathrm{var}(\widehat{f}_T(\omega_k)) = \sum_{j_1,j_2}\frac{1}{(bT)^2}K(\frac{\omega_k-\omega_{j_1}}{b})K(\frac{\omega_k-\omega_{j_2}}{b})\mathrm{cov}(|J_T(\omega_{j_1})|^2, |J_T(\omega_{j_2})|^2).
$$

Now we substitute (45) into the above to give

$$
\begin{aligned}
\mathrm{var}(\widehat{f}_T(\omega_k)) \;\; &\leq \;\; C\sum_{j_1,j_2}\frac{1}{(bT)^2}K(\frac{\omega_k-\omega_{j_1}}{b})K(\frac{\omega_k-\omega_{j_2}}{b})\{|F_2(j_1-j_2;\omega_{j_1})F_2(j_2-j_1;\omega_{j_1})| + \\
&\quad |F_2(j_2-j_1;\omega_{j_1})F_2(j_1-j_2;-\omega_{j_1})| + \frac{2\pi}{T}F_4(0;\omega_{j_1},-\omega_{j_1},\omega_{j_2})\} + O\big(\frac{(\log T)^3}{T^2}\big).
\end{aligned}
$$

We observe from the above that the covariance terms dominate the fourth order cumulant term. Moreover, by using Lemma A.2 we have $\sup_\omega \sum_s |F_2(s;\omega)| < \infty$, which gives $\mathrm{var}(\widehat{f}_T(\omega_k)) = O(\frac{1}{bT})$. This together with (46) gives (15).

We now prove (7). Using Lemma A.6 for $n = 2$ gives

$$\mathbb{E}(\tilde{c}_T(r)) = \frac{1}{T}\sum_{k=1}^{T} \frac{1}{[f(\omega_k)f(\omega_k+\omega_r)]^{1/2}} \frac{1}{T}\sum_{t=1}^{T} f(\frac{t}{T},\omega_k)\exp(-it\omega_r) + O(\frac{\log T}{T}). \tag{47}$$

Now by using replacing sum with integral and using Lemma A.3 (noting $F_{2,T}(-r;\omega_k) = \frac{1}{T}\sum_{t=1}^{T} f(\frac{t}{T},\omega_k)\exp(-it\omega_r)$ and $F_2(-r;\omega_k) = \int_0^1 f(u,\omega_k)\exp(-i2\pi ru)du$) gives

$$\mathbb{E}(\tilde{c}_T(r)) = \frac{1}{2\pi}\int_0^1\int_0^{2\pi} \frac{f(u,\omega)}{[f(\omega)f(\omega+\omega_r)]^{1/2}} \exp(-i2\pi ur)dud\omega + O(\frac{\log T}{T} + \frac{1}{T^2}), \tag{48}$$

thus we have (7). $\square$

To prove (16) in Lemma 4.1, we evaluate the limiting variance of $\tilde{c}_T(t)$ under the alternative of local stationarity.

Lemma A.12 *Suppose Assumption 4.1 holds. Then we have*

$$cov(\Re\sqrt{T}\tilde{c}_T(r_1), \Re\sqrt{T}\tilde{c}_T(r_2)) = \frac{1}{4}\big(\Gamma^{(1)}_{T,r_1,r_2} + \Gamma^{(2)}_{T,r_1,r_2} + \Gamma^{(2)}_{T,r_2,r_1} + \Gamma^{(3)}_{T,r_1,r_2}\big) + O(\frac{\log T}{T}), \tag{49}$$

$$cov(\Re\sqrt{T}\tilde{c}_T(r_1), \Im\sqrt{T}\tilde{c}_T(r_2)) = \frac{-i}{4}\big(\Gamma^{(1)}_{T,r_1,r_2} + \Gamma^{(2)}_{T,r_1,r_2} - \Gamma^{(2)}_{T,r_2,r_1} - \Gamma^{(3)}_{T,r_1,r_2}\big) + O(\frac{\log T}{T}),$$

$$cov(\Im\sqrt{T}\tilde{c}_T(r_1), \Im\sqrt{T}\tilde{c}_T(r_2)) = \frac{1}{4}\big(\Gamma^{(1)}_{T,r_1,r_2} - \Gamma^{(2)}_{T,r_1,r_2} - \Gamma^{(2)}_{T,r_2,r_1} + \Gamma^{(3)}_{T,r_1,r_2}\big) + O(\frac{\log T}{T}),$$

where

$$\Gamma^{(1)}_{T,r_1,r_2} = \frac{1}{T}\sum_{k_1,k_2} g^{(r_1)}_{T,k_1} g^{(r_2)}_{T,k_2}\big\{F_2(k_1-k_2;\omega_{k_1})F_2(-k_1-r_1+k_2+r_2;-\omega_{k_1+r_1}) +$$
$$F_2(k_1+k_2+r_2;\omega_{k_1})F_2(-k_1-r_1-k_2;-\omega_{k_1+r_1})\big\} + \frac{1}{T^2}\sum_{k_1,k_2} g^{(r_1)}_{T,k_1} g^{(r_2)}_{T,k_2} F_4(r_2-r_1;\omega_{k_1},-\omega_{k_1+r_1},-\omega_{k_2}),$$

$$\Gamma^{(2)}_{T,r_1,r_2} = \frac{1}{T}\sum_{k_1,k_2} g^{(r_1)}_{T,k_1} g^{(r_2)}_{T,k_2}\big\{F_2(k_1+k_2;\omega_{k_1})F_2(-k_1-r_1-k_2-r_2;-\omega_{k_1+r_1}) +$$

$$F_2(k_1-k_2-r_2;\omega_{k_1})F_2(-k_1-r_1+k_2;-\omega_{k_1+r_1})\big\} + \frac{1}{T^2}\sum_{k_1,k_2} g^{(r_1)}_{T,k_1} g^{(r_2)}_{T,k_2} F_4(-r_2-r_1;\omega_{k_1},-\omega_{k_1+r_1},-\omega_{k_2}),$$

$$\Gamma^{(3)}_{T,r_1,r_2} = \frac{1}{T}\sum_{k_1,k_2} g^{(r_1)}_{T,k_1} g^{(r_2)}_{T,k_2}\big\{F_2(-k_1+k_2;-\omega_{k_1})F_2(k_1+r_1-k_2-r_2;\omega_{k_1+r_1}) +$$

$$F_2(-k_1-k_2-r_2;-\omega_{k_1})F_2(k_1+r_1+k_2;\omega_{k_1+r_1})\big\} + \frac{1}{T^2}\sum_{k_1,k_2} g^{(r_1)}_{T,k_1} g^{(r_2)}_{T,k_2} F_4(r_1-r_2;-\omega_{k_1},\omega_{k_1+r_1},-\omega_{k_2}),$$

and the coefficients $F_2(\cdot)$ *and* $F_4(\cdot)$ *are defined in (26) and* $g^{(r)}_{T,k} = \big\{f(\omega_k)f(\omega_{k+r})\big\}^{-1/2}$.

PROOF. To prove (49) we use $\Re\widetilde{c}_T(r) = \frac{1}{2}(\widetilde{c}_T(r) + \bar{\widetilde{c}}_T(r))$ and $\Im\widetilde{c}_T(r) = \frac{-i}{2}(\widetilde{c}_T(r) + \bar{\widetilde{c}}_T(r))$, and $\mathrm{cov}(\sqrt{T}\tilde{c}_T(r_1), \sqrt{T}\bar{\tilde{c}}_T(r_2))$ and $\mathrm{cov}(\sqrt{T}\bar{\tilde{c}}_T(r_1), \sqrt{T}\bar{\tilde{c}}_T(r_2))$. Expanding $\mathrm{cov}(\sqrt{T}\tilde{c}_T(r_1), \sqrt{T}\tilde{c}_T(r_2))$ we have

$$\mathrm{cov}(\sqrt{T}\tilde{c}_T(r_1), \sqrt{T}\tilde{c}_T(r_2)) = \frac{1}{T}\sum_{k_1,k_2} g^{(r_1)}_{T,k_1} g^{(r_2)}_{T,k_2}\mathrm{cov}(J_{k_1}\overline{J}_{k_1+r_1}, J_{k_2}\overline{J}_{k_2+r_2}),$$

now by substituting (45) into the above we obtain

$$\mathrm{cov}(\sqrt{T}\tilde{c}_T(r_1), \sqrt{T}\tilde{c}_T(r_2)) = \Gamma^{(1)}_{T,r_1,r_2} + O(\frac{\log T}{T}).$$

Similar results can be obtained for $\mathrm{cov}(\sqrt{T}\tilde{c}_T(r_1), \sqrt{T}\bar{\tilde{c}}_T(r_2))$ and $\mathrm{cov}(\sqrt{T}\bar{\tilde{c}}_T(r_1), \sqrt{T}\tilde{c}_T(r_2))$. Using this we obtain the required result. $\square$

PROOF of Lemma 4.1, equation (16) This immediately follows from Lemma A.12. $\square$

A.4 Asymptotic normality

In this section we prove asymptotic normality of $\sqrt{T}\tilde{c}_T(r)$. We will derive normality under two set-ups (a) $\{X_t\}$ is strictly stationary, mixing and satisfies Assumption 3.1(B) and (b) $\{X_t\}$ is a stationary linear time series or a locally stationary linear time series and satisfies Assumption 3.1(A) or 4.1. Since the locally stationary linear time series model includes the stationary time series model as a special case we show asymptotic normality of the more general locally stationary model.

A.4.1 Asymptotic normality under strict stationarity and α-mixing

We observe that $\sqrt{T}\tilde{c}_T(r)$ can be written in the following quadratic form

$$\begin{aligned}
\sqrt{T}\tilde{c}_T(r) &= \frac{1}{\sqrt{T}}\sum_{t,\tau} X_t X_\tau \exp(i\tau\omega_r)\frac{1}{T}\sum_{k}\frac{1}{\sqrt{f(\omega_k)f(\omega_{k+r})}}\exp(i\omega_k(t-\tau)) \\
&= \frac{1}{\sqrt{T}}\sum_{t,\tau} G_{T,\omega_r}(t-\tau)\exp(ir\omega\tau)X_t X_\tau \exp(i\tau\omega_r) \\
&= \frac{1}{\sqrt{T}}\sum_{t,\tau} G_{\omega_r}(t-\tau)\exp(i\tau\omega_r)X_t X_\tau + O_p\big(\frac{1}{T^{1/2}}\big), \qquad (50)
\end{aligned}$$

where $G_{\omega_r}(t-\tau)$ is defined in (21) and we use that the Fourier coefficients of $\frac{1}{\sqrt{f(\omega)f(\omega+\omega_r)}}$ decay at the rate $|t-\tau|^{-2}$ (see Lemma A.1).

Theorem A.2 *Suppose Assumption 3.1(A,iv) holds. Then we have*

$$\sqrt{T}\bigg(\frac{1}{1+\varphi(\omega_1)}\Re\tilde{c}_T(1),\ldots,\frac{1}{1+\varphi(\omega_m)}\Im\tilde{c}_T(m)\bigg) \xrightarrow{D} \mathcal{N}(0, I_{2m}). \qquad (51)$$

PROOF. We observe from (52) that

$$\sqrt{T}\tilde{c}_T(r) = \frac{1}{\sqrt{T}}\sum_{t,\tau} G_{\omega_r}(t-\tau)\exp(i\tau\omega_r)X_t X_\tau + O_p\big(\frac{1}{T^{1/2}}\big). \qquad (52)$$

The limiting variance of $\sqrt{T}\tilde{c}_T(r)$ is given in Lemma 3.1. Now we use the result in Lee and Subba Rao (2010) for quadratic forms of mixing random variables and the Cramer-Wold device to obtain the result (similar results on quadratic forms can be found in Lin (2009)). $\square$

A.4.2 Asymptotic normality under linearity (both stationary and locally stationary)

We start by approximating $\sqrt{T}\tilde{c}_T(r)$ with a random variable which only involves current innovations $\{\varepsilon_t\}_{t=1}^T$. We make this approximation in order to use the martingale central limit theorem to prove asymptotic normality of $\sqrt{T}\tilde{c}_T(r)$. In this section, we will make frequent appeals to Lemma A.1.

Using that the locally stationary time series model $X_{t,T}$ satisfies (13) we have can write $\sqrt{T}\tilde{c}_T(r)$

as

$$
\begin{aligned}
&\sqrt{T}\tilde{c}_T(r) \\
&= \frac{1}{T^{3/2}}\sum_{k=1}^{T}\frac{1}{f(\omega_k)^{1/2}f(\omega_k+\omega_r)^{1/2}}\sum_{t,\tau=1}^{T}\exp(i(t-\tau)\omega_k)\exp(-i\tau\omega_r)\sum_{j_1,j_2=0}^{\infty}\psi_{t,T}(j_1)\psi_{\tau,T}(j_2)\varepsilon_{t-j_1}\varepsilon_{\tau-j_2} \\
= &\ \frac{1}{T^{1/2}}\sum_{t,\tau=1}^{T}G_{T,\omega_r}(t-\tau)\exp(-i\tau\omega_r)\sum_{j_1,j_2=0}^{\infty}\psi_{t,T}(j_1)\psi_{\tau,T}(j_2)\big(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2}-\mathbb{E}(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2})\big),
\end{aligned}
$$

where $\{G_{T,\omega_r}(s)\}$ is the DFT defined in (21).

We now partition $\sqrt{T}\tilde{c}_T(r)$ into terms which involve 'present' and 'past' innovations. More precisely

$$
\sqrt{T}\big(\tilde{c}_T(r)-\mathbb{E}(\tilde{c}_T(r))\big)=\sqrt{T}\big(d_T(r)+V_T(r)\big), \tag{53}
$$

where

$$
\begin{aligned}
d_T(r) &= \frac{1}{T}\sum_{t,\tau=1}^{T}G_{T,\omega_r}(t-\tau)\exp(-i\tau\omega_r)\sum_{0\le j_1\le t-1}\ \sum_{0\le j_2\le \tau-1}\psi_{t,T}(j_1)\psi_{\tau,T}(j_2)\big(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2}-\mathbb{E}(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2})\big) \\
V_T(r) &= \frac{1}{T}\sum_{t,\tau=1}^{T}G_{T,\omega_r}(t-\tau)\exp(-i\tau\omega_r)\sum_{j_1\ge t-1 \text{ or } j_2\ge \tau-1}\psi_{t,T}(j_1)\psi_{\tau,T}(j_2)\big(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2}-\mathbb{E}(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2})\big).
\end{aligned}
$$

In the following lemma we obtain a bound for the remainder $V_T(r)$. Later we will show asymptotic normality of $d_T(r)$.

Lemma A.13 *Suppose Assumption 3.1(A, iii,iv) or 4.1 hold. Then we have*

$$
\big(T^{1/2}\mathbb{E}|V_T(r)|^2\big)^{1/2}\le CT^{-1/2},
$$

for some finite constant C.

PROOF. We first observe that $\sqrt{T}V_T(r)=I_1+I_2+I_3$, where

$$
I_1 = \frac{1}{T^{1/2}}\sum_{t,\tau=1}^{T}G_{T,\omega_r}(t-\tau)\exp(-i\tau\omega_r)\sum_{j_1\ge t-1}\ \sum_{0\le j_2\le \tau-1}\psi_{t,T}(j_1)\psi_{\tau,T}(j_2)\big(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2}-\mathbb{E}(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2})\big)
$$

$$
I_2 = \frac{1}{T^{1/2}}\sum_{t,\tau=1}^{T}G_{T,\omega_r}(t-\tau)\exp(-i\tau\omega_r)\sum_{j_2\ge \tau-1}\ \sum_{0\le j_1\le t-1}\psi_{t,T}(j_1)\psi_{\tau,T}(j_2)\big(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2}-\mathbb{E}(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2})\big)
$$

$$I_3 = \frac{1}{T^{1/2}} \sum_{t,\tau=1}^{T} G_{T,\omega_r}(t-\tau)\exp(-i\tau\omega_r) \sum_{j_2 \geq \tau-1} \sum_{j_1 \geq t-1} \psi_{t,T}(j_1)\psi_{\tau,T}(j_2)\big(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2} - \mathbb{E}(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2})\big).$$

We first show that $\mathbb{E}(I_1^2)^{1/2} = O(T^{-1/2})$. By the Minkowski's inequality we have

$$\mathbb{E}(I_1^2)^{1/2} \leq \frac{1}{T^{1/2}} \sum_{t,\tau=1}^{T} |G_{T,\omega_r}(t-\tau)| \bigg\{ \mathbb{E}\bigg(\sum_{j_1 \geq t-1} \sum_{0 \leq j_2 \leq \tau-1} \psi_{t,T}(j_1)\psi_{\tau,T}(j_2)\big(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2} - \mathbb{E}(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2})\big)\bigg)^2 \bigg\}^{1/2}.$$

It can be shown that

$$\begin{aligned} & \mathbb{E}\bigg(\sum_{j_1 \geq t-1} \sum_{0 \leq j_2 \leq \tau-1} \psi_{t,T}(j_1)\psi_{\tau,T}(j_2)\big(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2} - \mathbb{E}(\varepsilon_{t-j_1}\varepsilon_{\tau-j_2})\big)\bigg)^2 \\ \leq \;& \big(\mathrm{var}(\varepsilon_t)^2 + \mathrm{var}(\varepsilon_t^2)\big) \Big[\sum_{j_1 \geq t} |\psi_{t,T}(j_1)|^2 \Big] \Big[\sum_{j_2=0}^{\infty} |\psi_{\tau,T}(j_2)|^2 \Big]. \end{aligned}$$

Substituting this into the bound for $\mathbb{E}(I_1^2)$, under Assumption 4.1 and using Lemma A.1 we have

$$\begin{aligned} \mathbb{E}(I_1^2)^{1/2} \;\leq\; & \frac{1}{T^{1/2}} \sup_{\tau} \Big[\sum_{j_2=0}^{\infty} |\psi_{\tau,T}(j_2)|^2 \Big]^{1/2} \Big[\sum_{s=1}^{T} |G_{T,\omega_r}(s)| \Big] \sum_{t=1}^{T} \Big[\sum_{j_1 \geq t} |\psi_{t,T}(j_1)|^2 \Big]^{1/2} \\ & \frac{1}{T^{1/2}} \sup_{\tau} \Big[\sum_{j_2=0}^{\infty} |\psi_{\tau,T}(j_2)|^2 \Big]^{1/2} \Big[\sum_{s=1}^{T} |G_{T,\omega_r}(s)| \Big] \sum_{t=1}^{T} \sum_{j_1 \geq t} |\psi_{t,T}(j_1)| = O(T^{-1/2}). \end{aligned}$$

Using a similar method we can show that $\mathbb{E}(I_2^2)^{1/2} = O(T^{-1/2})$ and $\mathbb{E}(I_3^2)^{1/2} = O(T^{-1/2})$. Thus we obtain the result. $\square$

Therefore the above lemma shows that $\sqrt{T}\big(\tilde{c}_T(r) - \mathbb{E}(\tilde{c}_T(r)\big) = \sqrt{T}d_T(r) + O_p(T^{-1/2})$.

Remark A.1 *Now it is worth noting that in the case that $\{X_t\}$ is a stationary linear time series, then $d_T(r)$ has an interesting form. That is, it is straightforward to show (using (Priestley, 1981), Theorem 6.2.1) that*

$$\sqrt{T}\tilde{c}_T(r) = \frac{1}{\sqrt{T}} \sum_{k=1}^{T} J_\varepsilon(\omega_k)\overline{J_\varepsilon(\omega_{k+r})} \exp\big(i(\phi(\omega_k) - \phi(\omega_{k+r}))\big) + O_p(T^{-1/2}),$$

where $J_\varepsilon(\omega) = (2\pi T)^{-1/2} \sum_{t=1}^{T} \varepsilon_t \exp(it\omega_k)$.

We use the martingale central limit theorem to show asymptotic normality of $\sqrt{T}d_T(r)$, which will imply asymptotic normality of $\sqrt{T}\big(\tilde{c}_T(r) - \mathbb{E}(\tilde{c}_T(r)\big)$. To do this we rewrite $\sqrt{T}d_T(r)$ as the sum of martingale differences

$$
\begin{aligned}
&\sqrt{T}d_T(r) \\
&= \frac{1}{T^{1/2}}\sum_{s_1,s_2=1}^{T}\big(\varepsilon_{s_1}\varepsilon_{s_2}-\mathbb{E}(\varepsilon_{s_1}\varepsilon_{s_2})\big)\sum_{s_1\leq t\leq T}\sum_{s_2\leq\tau\leq T}G_{T,\omega_r}(t-\tau)\exp(-i\tau\omega_r)\psi_{t,T}(t-s_1)\psi_{\tau,T}(\tau-s_2) \\
&= \frac{1}{T^{1/2}}\sum_{s=1}^{T}M_T(s) \quad \text{ where } \quad M_T(s)=\big(\varepsilon_s^2-1\big)A_T(s,s)+\varepsilon_s\sum_{s_1=1}^{s-1}\varepsilon_{s_1}\big(A_T(s_1,s)+A_T(s_1,s)\big)
\end{aligned}
$$

and

$$
A_T(s_1,s_2)=\sum_{s_1\leq t\leq T}\sum_{s_2\leq\tau\leq T}G_{T,\omega_r}(t-\tau)\exp(-i\tau\omega_r)\psi_{t,T}(t-s_1)\psi_{\tau,T}(\tau-s_2).
$$

We now show that the coefficients in the martingale differences are absolutely summable.

Lemma A.14 *Suppose Assumption 3.1(A, iii,iv) or 4.1 holds. Then we have*

$$
\sup_T\sum_{s_1=1}^{s-1}(|A_T(s,s_1)|+|A_T(s_1,s)|)<\infty.
$$

PROOF. To prove the result we note that under Assumption 4.1 and using Lemma A.1 we have

$$
\begin{aligned}
\sum_{s\leq\tau\leq T}\sum_{s_1=1}^{s-1}|A_T(s_1,s)| &\leq \sum_{s_1=1}^{s-1}\sum_{s_1\leq t\leq T}|G_{T,\omega_r}(t-\tau)|\cdot|\psi_{t,T}(t-s_1)|\cdot|\psi_{\tau,T}(\tau-s)| \\
&\leq \Big[\sum_t|G_{T,\omega_r}(t)|\Big]\sup_{t,T}\Big[\sum_s|\psi_{t,T}(s)|\Big]^2,
\end{aligned}
$$

which gives the required result. □

In the theorem below we show asymptotic normality of $d_T(r)$. To accommodate both the stationary and nonstationary case we will let the asymptotic variance of $d_T(r)$ be V_r and specify its value later.

Theorem A.3 *Suppose Assumption 3.1(A,iii,iv) or 4.1 holds. Furthermore suppose that* $var(\sqrt{T}d_T(r),\sqrt{T}d_T(r))\to V_r<\infty$ *as* $T\to\infty$*. Then we have*

$$
\sqrt{T}\begin{pmatrix}\Re\sqrt{T}d_T(r)\\ \Im\sqrt{T}d_T(r)\end{pmatrix}\xrightarrow{D}\mathcal{N}\big(0,V_r\big) \qquad T\to\infty.
$$

PROOF. We use the martingale central limit theorem to prove the result. We will show asymptotic normality of $\Re\sqrt{T}d_T(r)$. However, using the same method it straightforward to show asymptotic normality for all linear combinations of $\Re\sqrt{T}d_T(r)$ and $\Im\sqrt{T}d_T(r)$ and thus by the Cramer-Wold device to show asymptotic normality of the random vector $\big(\Re\sqrt{T}d_T(r), \Im\sqrt{T}d_T(r)\big)$. Let $M_{1,T} = \Re M_T(s)$ and $V_{r,1} = (V_r)_{(1,1)}$. To apply the martingale central limit theorem we need to verify that the variance of $T^{-1/2}\sum_{d=1}^{T} M_{1,T}(s)$ is finite (which is assumed), Lindeberg's condition is satisfied and $\frac{1}{T}\sum_{s=1}^{T}\mathbb{E}(M_{1,T}(s)^2|\mathcal{F}_{s-1}) \xrightarrow{P} V_{r,1}$ (see (Hall & Heyde, 1980), Theorem 3.2). To verify Lindeberg's condition, we require that for all $\delta > 0$,

$$L_T = \frac{1}{T}\sum_{s=1}^{T}\mathbb{E}(M_{1,T}(s)^2 I(T^{-1/2}|M_{1,T}(s)| > \delta)|\mathcal{F}_{s-1}) \xrightarrow{P} 0,$$

$T \to \infty$, where $I(\cdot)$ is the identity function and $\mathcal{F}_s = \sigma(M_{1,T}(s), M_{1,T}(s-1), \ldots, M_{1,T}(1))$. By using Hölder and Markov inequalities, we obtain a bound for the following L_T

$$L_T \leq (T\delta)^{-1}\frac{1}{T}\sum_{s=1}^{T}\mathbb{E}(M_{1,T}(s)^4|\mathcal{F}_{s-1}). \tag{54}$$

Now by using Lemma A.14 we have $\sum_{s_1}\big(|A_T(s,s_1)| + A_T(s_1,s)|\big) < \infty$, therefore

$$\sup_T \mathbb{E}\big(\frac{1}{T}\sum_{s=1}^{T}\mathbb{E}(M_{1,T}(s)^4|\mathcal{F}_{s-1})\big) = \frac{1}{T}\sup_T\sum_{s=1}^{T}\mathbb{E}(M_{1,T}(s)^4) < \infty.$$

Since $\frac{1}{T}\sum_{s=1}^{T}\mathbb{E}(M_{1,T}(s)^4|\mathcal{F}_{s-1})$ is a positive random variable, the above result implies $\frac{1}{T}\sum_{s=1}^{T}\mathbb{E}(M_{1,T}(s)^4|\mathcal{F}_{s-1}) = O_p(1)$. Substituting this into (54) gives $L_T \xrightarrow{P} 0$ as $T \to \infty$.
Finally we need to show that

$$\frac{1}{T}\sum_{s=1}^{T}\mathbb{E}(M_{1,T}(s)^2|\mathcal{F}_{s-1}) = \frac{1}{T}\sum_{s=1}^{T}\big[\mathbb{E}(M_{1,T}(s)^2|\mathcal{F}_{s-1}) - \mathbb{E}(M_{1,T}(s)^2)\big] + \frac{1}{T}\sum_{s=1}^{T}\mathbb{E}(M_{1,T}(s)^2) \xrightarrow{P} V_{r,1}. \tag{55}$$

Under the stated assumptions, we have $\frac{1}{T}\sum_{s=1}^{T}\mathbb{E}(M_{1,T}(s)^2) \to V_{r,1}$ as $T \to \infty$. Therefore it remains to show

$$P_T := \frac{1}{T}\sum_{s=1}^{T}\big|\mathbb{E}(M_{1,T}(s)^2|\mathcal{F}_{s-1}) - \mathbb{E}(M_{1,T}(s)^2)\big| \xrightarrow{P} 0,$$

which will give us (55). We will show that $\mathbb{E}(P_T^2) \to 0$. To do this we note that $\mathbb{E}(P_T) = 0$ and

$$\mathrm{var}(P_T) = \frac{1}{T^2}\sum_{d=1}^{T}\mathrm{var}(\mathbb{E}(M_{1,T}(s)^2|\mathcal{F}_{s-1})) + \frac{2}{T^2}\sum_{s_1>s_2}^{T}\mathrm{cov}(\mathbb{E}(M_{1,T}(s_1)^2|\mathcal{F}_{s_1-1}), \mathbb{E}(M_{1,T}(s_2)^2|\mathcal{F}_{s_2-1})). \tag{56}$$

Now by using the Cauchy Schwartz inequality and conditional expectation arguments for $\mathcal{F}_{s_2} \subset \mathcal{F}_{s_1}$

we have

$$\begin{aligned} & \mathrm{cov}(\mathbb{E}(M_{1,T}(s_1)^2|\mathcal{F}_{s_1-1}), \mathbb{E}(M_{1,T}(s_2)^2|\mathcal{F}_{s_2-1})) \\ \leq & \Big[\mathbb{E}\big(\mathbb{E}(M_{1,T}(s_2)^2|\mathcal{F}_{s_2-1}) - \mathbb{E}(M_{1,T}(s_2)^2)\big)^2\Big]^{1/2}\Big[\mathbb{E}\big(\mathbb{E}(M_{1,T}(s_1)^2|\mathcal{F}_{s_2-1}) - \mathbb{E}(M_{1,T}(s_1)^2)\big)^2\Big]^{1/2}. \end{aligned}$$

We now show that $\sup_T \sum_{s_1=s_2}^{T} \Big[\mathbb{E}\big(\mathbb{E}(M_{1,T}(s_1)^2|\mathcal{F}_{s_2-1}) - \mathbb{E}(M_{1,T}(s_1)^2)\big)^2\Big]^{1/2} < \infty$. Let $\mathcal{G}_s = \sigma(\varepsilon_s, \varepsilon_{s-1}, \ldots)$, then it is clear that for all s, $\mathcal{F}_s \subset \mathcal{G}_s$. Therefore, we have $\mathbb{E}\big[\mathbb{E}(M_{1,T}(s_1)^2|\mathcal{F}_{s_2-1})^2\big] \leq \mathbb{E}\big[\mathbb{E}(M_{1,T}(s_1)^2|\mathcal{G}_{s_2-1})^2\big]$ which gives

$$\begin{aligned} \mathbb{E}\Big[\mathbb{E}(M_{1,T}(s_1)^2|\mathcal{F}_{s_2-1}) - \mathbb{E}(M_{1,T}(s_1)^2)\Big]^2 &= \mathbb{E}\big[\mathbb{E}(M_{1,T}(s_1)^2|\mathcal{F}_{s_2-1})^2\big] - \big[\mathbb{E}(M_{1,T}(s_1)^2)\big]^2 \\ &\leq \mathbb{E}\big[\mathbb{E}(M_{1,T}(s_1)^2|\mathcal{G}_{s_2-1})^2\big] - \big[\mathbb{E}(M_{1,T}(s_1)^2)\big]^2. \end{aligned}$$

Expanding $M_{1,T}(s_1)$ in terms of $\{\varepsilon_t\}$ and using $\sup_{T,t}\sum_j |\psi_{t,T}(j)| < \infty$, it can be shown that $\mathbb{E}\big[\mathbb{E}(M_{1,T}(s_1)^2|\mathcal{G}_{s_2-1})^2\big] - \big[\mathbb{E}(M_{1,T}(s_1)^2)\big]^2 \to 0$ as $s_1 \to \infty$, and

$$\mathrm{var}(P_T) \leq \frac{1}{T}\sum_{s_2=1}^{T} \sup_{s_2,T} \sum_{s_1=s_2}^{T} \Big(\mathbb{E}\big[\mathbb{E}(M_{1.T}(s_1)^2|\mathcal{G}_{s_2-1})^2\big] - \big[\mathbb{E}(M_{s_1}^2)\big]^2\Big)^{1/2} < \infty.$$

Substituting the above into (57) we have $\mathrm{var}(P_T) = O(T^{-1})$, hence we have shown (55), and the conditions of the martingale central limit theorem are satisfied, giving the required result. □

Theorem A.4 *Suppose Assumption 3.1(A,iii,iv) or 4.1 holds. Furthermore suppose that var*$(\Re\sqrt{T}\widetilde{c}_T(1), \Im\sqrt{T}\widetilde{c}_T(m)) \to V < \infty$ *as* $T \to \infty$. *Then we have*

$$\sqrt{T}\bigg(\frac{1}{1+\varphi(\omega_1)}(\Re\widetilde{c}_T(1) - \mathbb{E}(\Re\widetilde{c}_T(1))), \ldots, \frac{1}{1+\varphi(\omega_m)}(\Im\widetilde{c}_T(m) - \mathbb{E}(\Im\widetilde{c}_T(m)))\bigg) \xrightarrow{D} \mathcal{N}\big(0, V\big).$$

PROOF. Using (53) and Lemma A.13 we have

$$\begin{aligned} & \sqrt{T}\bigg(\frac{1}{1+\varphi(\omega_1)}(\Re\widetilde{c}_T(1) - \mathbb{E}(\Re\widetilde{c}_T(1))), \ldots, \frac{1}{1+\varphi(\omega_m)}(\Im\widetilde{c}_T(m) - \mathbb{E}(\Im\widetilde{c}_T(m)))\bigg) \\ = & \sqrt{T}\big(\Re d_T(1), \ldots, \Im d_T(m)\big) + O_p(\frac{1}{T^{1/2}}), \end{aligned}$$

now by using same proof as that in Theorem A.3 we obtain the result. □

A.4.3 Proof of Theorems 3.2 and 4.2

Proof of Theorem 3.2 Using Lemma 3.1 we have

$$\begin{aligned}\sqrt{T}\Big(\Re\widetilde{c}_T(1),\ldots,\Im\widetilde{c}_T(m)\Big) &= \sqrt{T}\Big(\Re\widehat{c}_T(1),\ldots,\Im\widehat{c}_T(m)\Big) + \\ & O_p\left(m\left[(b+\frac{1}{\sqrt{bT}}) + (\frac{1}{bT^{1/2}} + b^2T^{1/2})\sum_{n=1}^{m}(\frac{1}{|r_n|} + \frac{1}{T^{1/2}})\right]\right).\end{aligned}$$

Lemma 3.1, implies that $T\mathrm{var}\Big(\Re\widehat{c}_T(1),\ldots,\Im\widehat{c}_T(m)\Big) \to \mathrm{diag}(1+\varphi(\omega_1),\ldots,\mathrm{diag}(1+\varphi(\omega_m))$. Combining this with either Theorems A.2 or A.4, gives

$$\sqrt{T}\left(\frac{1}{1+\varphi(\omega_1)}\Re\widetilde{c}_T(1),\ldots,\frac{1}{1+\varphi(\omega_m)}\Im\widetilde{c}_T(m)\right) \xrightarrow{D} \mathcal{N}(0, I_{2m}). \tag{57}$$

Finally, since $m(b+\frac{1}{\sqrt{bT}}) + \left(\frac{1}{bT^{1/2}} + b^2T^{1/2}\right)\sum_{r=1}^{m}\left(\frac{1}{|r|}+\frac{1}{T^{1/2}}\right) \to 0$, using (57) we have (51). □

PROOF of Theorem 4.2. The proof is identical to the proof of Theorem 3.2. Hence we omit the details. □

References

Andreou, E., & Ghysels, E. (2008). Structural breaks in financial time series. In T. G. Anderson, R. A. Davis, J.-P. Kreiss, & T. Mikosch (Eds.), *Handbook of financial time series* (p. 839-866). Berlin: Springer.

Beltrao, K. I., & Bloomfield, P. (1987). Determining the bandwidth of a kernel spectrum estimate. *Journal of Time Series Analysis*, *8*, 23-38.

Briggs, W., & Henson, V. E. (1997). *The DFT: An owner's manual for the discrete fourier transform.* SIAM.

Brillinger, D. (1981). *Time series, data analysis and theory.* San Francisco: SIAM.

Brillinger, D., & Rosenblatt, M. (1967). Asymptotic theory of estimates of k-th order spectra. *Proceedings of the National Academy of Sciences, USA*, *57*, 206-210.

Brockwell, P. J., & Davis, R. A. (1987). *Time series: Theory and methods.* New York: Springer.

Chen, W., & Deo, R. (2004). A generalized Portmanteau goodness of fit test for time series. *Econometric Theory*, *20*, 382-416.

Cline, D., & Pu, H. (1999). Geometric ergodicity of nonlinear time series. *Statistica Sinica*, *9*, 1103-118.

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.*, *16*, 1-37.

Dahlhaus, R., & Polonik, W. (2006). Nonparametric quasi-maximum likelihood estimation for Gaussian locally stationary processes. *Ann. Stat.*, *34*, 2790-2824.

Doukhan, P. (1994). *Mixing, properties and examples.* New York: Springer.

Hall, P., & Heyde, C. (1980). *Martingale Limit Theory and its Application.* New York: Academic Press.

Lee, J., & Subba Rao, S. (2010). A short note on asymptotic normality of quadratic forms for stationary and nonstationary, mixing random variables. *Preprint*.

Lin, F. (2009). The asymptotic behaviour of quadratic forms in ϕ-mixing random variables. *Journal of Computational and Applied Mathematics*, *233*, 437-448.

Neumann, M. (1996). Spectral density estimation via nonlinear wavelet methods for stationary nongaussian time series. *J. Time Series Analysis*.

Paparoditis, E. (2009). Testing temporal constancy of the spectral structure of a time series. *Bernoulli*, *15*, 1190-1221.

Priestley, M. B. (1981). *Spectral analysis and time series.* London: Academic Press.

Priestley, M. B., & Subba Rao, T. (1969). A test for non-stationarity of a time series. *J. Roy. Stat. (B)*, *31*, 140-149.

Robinson, P. (1989). Nonparametric estimation of time-varying parameters. In P. Hackl (Ed.), *Statistical analysis and Forecasting of Economic Structural Change* (p. 253-264). Berlin: Springer.

Sachs, R. von, & Neumann, M. H. (1999). A wavelet-based test for stationarity. *Journal of Time Series Analysis*, *21*, 597-613.

Statulevicius, V., & Jakimavicius. (1988). Estimates of semiivariant and centered moments of stochastic processes with mixing: I. *Lithuanian Math. J.*, *28*, 226-238.