A test for second order stationarity of a multivariate time series

Carsten Jentsch* and Suhasini Subba Rao†

August 1, 2013

Abstract

It is well known that the discrete Fourier transforms (DFT) of a second order stationary time series between two distinct Fourier frequencies are asymptotically uncorrelated. In contrast for a large class of second order nonstationary time series, including locally stationary time series, this property does not hold. In this paper these starkly differing properties are used to define a global test for stationarity based on the DFT of a vector time series. It is shown that the test statistic under the null of stationarity asymptotically has a chi-squared distribution, whereas under the alternative of local stationarity asymptotically it has a noncentral chi-squared distribution. Further, if the time series is Gaussian and stationary, the test statistic is pivotal. However, in many econometric applications, the assumption of Gaussianity can be too strong, but under weaker conditions the test statistic involves an unknown variance that is extremely difficult to directly estimate from the data. To overcome this issue, a scheme to estimate the unknown variance, based on the stationary bootstrap, is proposed. The properties of the stationary bootstrap under both stationarity and nonstationarity are derived. These results are used to show consistency of the bootstrap estimator under stationarity and to derive the power of the test under nonstationarity. The method is illustrated with some simulations. The test is also used to test for stationarity of FTSE 100 and DAG 30 stock indexes from January 2011-December 2012.

Keywords and Phrases: Discrete Fourier transform; Local stationarity; Nonlinear time series; Stationary bootstrap; Testing for stationarity

*Department of Economics, University of Mannheim, L7, 3-5, 68131 Mannheim, Germany, cjentsch@mail.uni-mannheim.de

†Department of Statistics, Texas A&M University, College Station, Texas 77845, USA, suhasini@stat.tamu.edu (corresponding author)

1 Introduction

In several disciplines, including finance, the geo sciences and the biological sciences, there has been a dramatic increase in the availability of multivariate time series data. In order to model this type of data, several multivariate time series models have been proposed, including the Vector Autoregressive model and the vector GARCH model, to name but a few (see, for example, Lütkepohl (2005) and Laurent, Rombouts, and Violante (2011)). The majority of these models are constructed under the assumption that the underlying time series is stationary. For some time series this assumption can be too strong, especially over relatively long periods of time. However, relaxing this assumption, to allow for nonstationary time series models, can lead to complex models with a large number of parameters, which may not be straightforward to estimate. Therefore, before fitting a time series model, it is important to check whether or not the multivariate time series is second order stationary.

Over the years, various tests for second order stationarity for univariate time series have been proposed. These include, Priestley and Subba Rao (1969), Loretan and Phillips (1994), von Sachs and Neumann (1999), Paparoditis (2009, 2010), Dahlhaus and Polonik (2009), Dwivedi and Subba Rao (2011), Dette, Preuss, and Vetter (2011), Dahlhaus (2012), Example 10, Jentsch (2012), Lei, Wang, and Wang (2012) and Nason (2013). However, as far as we are aware there does not exist any tests for second order stationarity of multivariate time series (Jentsch (2012) does propose a test for multivariate stationarity, but the test was designed to the detect the alternative of a multivariate periodically stationary time series). One crude solution is to individually test for stationarity for each of the univariate processes. However, there are a few drawbacks with this approach. The first is that such a multiple testing scheme does not take into account that each of the test statistics are independent (since the difficulty with multivariate time series is the dependencies between the marginal univariate time series) leading to incorrect type I errors. The second problem is that such a strategy can lead to misleading conclusions. For example if each of the marginal time series are second order stationary, but the cross-covariances are second order nonstationary, the above testing scheme would not be able detect the alternative. Therefore there is a need to develop a test for stationarity of a multivariate time series, which is the aim in this paper.

The majority of the univariate tests, are local, in the sense that they are based on comparing the local spectral densities over various segments. This approach suffers from some possible disadvantages. In particular, the spectral density may locally vary over time, but this does not imply that the process is second order nonstationary, for example Hidden Markov models can be stationary but the spectral density can vary according to the regime. For these reason, we propose a global test for multivariate second order stationarity.

Our test is motivated by the tests for detecting periodic stationarity (see, for example, Goodman (1965), Hurd and Gerr (1991) and Bloomfield, Hurd, and Lund (1994)) and the test for second order stationarity proposed in Dwivedi and Subba Rao (2011), all these tests use fundamental properties of the discrete Fourier transform (DFT). More precisely, the above mentioned periodic stationarity tests are based on the property that the discrete Fourier transform is correlated if the difference in the frequencies is a multiple of $2\pi/P$ (where P denotes the periodicity), whereas Dwivedi and Subba Rao (2011) use the idea that the DFT asymptotically uncorrelates stationary time series, but not nonstationary time series. Motivated by Dwivedi and Subba Rao (2011), in this paper, we exploit the uncorrelating property of the DFT to construct the test. However, the test proposed here differs from Dwivedi and Subba Rao (2011) in several important ways, these include (i) our test takes into account the multivariate nature of the time series (ii) the proposed test is defined such that it can detect a wider range of alternatives and (iii) most tests for stationarity assume Gaussianity or linearity of the underlying time series, which in several econometric applications is unrealistic, whereas our test allows for testing of nonlinear stationary time series.

In Section 2, we motivate the test statistic by comparing the covariance between the DFT of stationary and nonstationary time series, where we focus on the large class of nonstationary processes called locally stationary time series (see Dahlhaus (1997), Dahlhaus and Polonik (2006) and Dahlhaus (2012) for a review). Based on these observations, we define DFT covariances which in turn are used to define a Portmanteau-type test statistic. Under the assumption of Gaussianity, the test statistic is pivotal, however for non-Gaussian time series the test statistic involves a variance which is unknown and extremely difficult to estimate. If we were to ignore this variance (and thus implicitly assume Gaussianity) then the test can be unreliable. Therefore in Section 2.4 we propose a bootstrap procedure, based on the stationary bootstrap (first proposed in Politis and Romano (1994)), to estimate the variance. In Section 3, we derive the asymptotic sampling properties of the DFT covariance. We show that under the null hypothesis, the mean of the DFT covariance is asymptotically zero. In contrast, under the alternative of local stationarity, we show that the DFT covariance estimates nonstationary characteristics in the time series. These results are used to derive the sampling distribution of the test statistic. Since the stationary bootstrap is used to estimate the unknown variance, in Section 4, we analyze the stationary bootstrap when the underlying time series is stationary and nonstationary.

Some of these results may be of independent interest. In Section 5 we show that under (fourth order) stationarity the bootstrap variance estimator is a consistent estimator of the true variance. In addition, we analyze the bootstrap variance estimator under nonstationarity and show how it influences the power of the test. The test statistic involves some tuning parameters and in Section 6.1, we give some suggestions on how to select these tuning parameters. In Section 6.2, we analyze the performance of the test statistic under both the null and the alternative and compare the test statistic when the variance is estimated using the bootstrap and when Gaussianity is assumed. In the simulations we include both stationary GARCH and Markov switching models and for nonstationary models we consider time-varying linear models and the random walk. In Section 6.3, we apply our method to analyze the FTSE 100 and DAX 30 stock indexes. Typically, stationary GARCH-type models are used to model this type of data. However, even over the relatively short period January 2011- December 2012, the results from the test suggest that the log returns are nonstationary.

The proofs can be found in the Appendix.

2 The test statistic

2.1 Motivation

Let us suppose $\{\underline{X}_t = (X_{t,1}, \ldots, X_{t,d})', t \in \mathbb{Z}\}$ is a d-dimensional constant mean, multivariate time series and we observe $\{\underline{X}_t\}_{t=1}^{T}$. We define the vector discrete Fourier transform (DFT) as

$$\underline{J}_T(\omega_k) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} \underline{X}_t e^{-it\omega_k}, \quad k = 1, \ldots, T,$$

where $\omega_k = 2\pi \frac{k}{T}$ are the Fourier frequencies. Suppose that $\{\underline{X}_t\}$ is a second order stationary multivariate time series, where the autocovariance matrices of $\{\underline{X}_t\}$ satisfy

$$\sum_{h=-\infty}^{\infty} |h| \cdot |Cov(X_{0,m} X_{h,n})| < \infty \text{ for all } m, n = 1, \ldots, d. \tag{2.1}$$

Then, it is well known for $k_1 - k_2 \neq 0$, that $Cov(J_{T,m}(\omega_{k_1}), J_{T,n}(\omega_{k_2})) = O(\frac{1}{T})$ (uniformly in T, k_1 and k_2), in other words the DFT has transformed a stationary time series into a sequence which is approximately uncorrelated. The behavior in the case that the vector time series is second order *nonstationary* is very different. To obtain an asymptotic expression for the covariance between the DFTs, we will use the rescaling device introduced by Dahlhaus (1997) to study locally stationary time series, which is a class of nonstationary processes. $\{\underline{X}_{t,T}\}$ is called a locally second order stationary time series, if its covariance structure changes slowly over time such that there exists a smooth matrix function $\{\boldsymbol{\kappa}(u;r)\}$ which can approximate the time-varying covariance. More precisely, $|\mathrm{cov}(\underline{X}_{t,T}, \underline{X}_{\tau,T}) - \boldsymbol{\kappa}(\frac{t}{T}; t-\tau)| \leq T^{-1}\boldsymbol{\kappa}(t-\tau)$, where $\{\boldsymbol{\kappa}(h)\}_h$ is a matrix sequence whose elements are absolutely summable. An example of a locally stationary model which satisfies these conditions is the time-varying moving average model defined in Dahlhaus (2012), equations (63)–(65) (with $\ell(j) = \log(|j|)^{1+\varepsilon}|j|^2$ for $|j| \neq 0$). It is worth mentioning that Dahlhaus (2012) uses the slightly weaker condition $\ell(j) = \log(|j|)^{1+\varepsilon}|j|$. In the Appendix (Lemma A.8), we show that

$$\mathrm{cov}(\underline{J}_T(\omega_{k_1}), \underline{J}_T(\omega_{k_2})) = \int_0^1 \mathbf{f}(u;\omega_{k_1}) \exp(i2\pi u(k_1-k_2))du + O(\frac{1}{T}), \tag{2.2}$$

uniformly in T, k_1 and k_2, where $\mathbf{f}(u;\omega) = \frac{1}{2\pi}\sum_{r=-\infty}^{\infty} \boldsymbol{\kappa}(u;r)\exp(-ir\omega)$ is the local spectral density matrix (see Lemma A.8 for details). We recall if $\{\underline{X}_t\}_t$ is second order stationary then the 'spectral density' function $\mathbf{f}(u;\omega)$ does not depend on u and the above expression reduces to $Cov(\underline{J}_T(\omega_{k_1}), \underline{J}_T(\omega_{k_2})) = O(\frac{1}{T})$ for $k_1 - k_2 \neq 0$. It is interesting to observe that for locally stationary time series its DFT sequence mimics the behavior of a time series, in the sense that the correlation between the DFTs decay the further apart the frequencies.

Equation (2.2) highlights the starkly differing properties of the covariance of the DFTs between stationary and nonstationary time series, and we will exploit this difference in order to construct the test statistic.

2.2 The weighted DFT covariance

The discussion in the previous section suggests that to test for stationarity, we can transform the time series into the frequency domain and test if the vector sequence $\{\underline{J}_T(\omega_k)\}$ is asymptotically uncorrelated. Testing for uncorrelatedness of a multivariate time series is a well established technique in time series analysis (see, for example, Hosking (1980, 1981) and Escanciano and Lobato (2009)). Most of these tests are based on constructing a test statistic which is a function of sample autocovariance matrices of the time series. Motivated by these methods, we will define the weighted (standardized) covariance DFT and use this to define the test statistic.

To summarize the previous section, if $\{\underline{X}_t\}$ is a second order stationary time series which satisfies (2.1), then $\mathrm{E}(\underline{J}_T(\omega_k)) = \underline{0}$ (for $k \neq 0, T/2, T$) and $\mathrm{var}(\underline{J}_T(\omega_k)) \to \mathbf{f}(\omega_k)$ as $T \to \infty$, where $\mathbf{f} : [0, 2\pi] \to \mathbb{C}^{d\times d}$ with

$$\mathbf{f}(\omega) = \{f_{m,n}(\omega); m, n = 1, \ldots, d\}$$

is the spectral density matrix of $\{\underline{X}_t\}$. If the spectral density $\mathbf{f}(\omega)$ is non-singular on $[0, 2\pi]$, then its Cholesky decomposition is unique and well defined on $[0, 2\pi]$. More precisely,

$$\mathbf{f}(\omega) = \mathbf{B}(\omega)\overline{\mathbf{B}(\omega)}', \tag{2.3}$$

where $\mathbf{B}(\omega)$ is a lower triangular matrix and $\overline{\mathbf{B}(\omega)}'$ denotes the transpose and complex conjugate of $\mathbf{B}(\omega)$. Let $\mathbf{L}(\omega_k) := \mathbf{B}^{-1}(\omega_k)$, thus $\mathbf{f}^{-1}(\omega_k) = \overline{\mathbf{L}(\omega_k)}'\mathbf{L}(\omega_k)$. Therefore, if $\{\underline{X}_t\}$ is a second order stationary time series, then the vector sequence, $\{\mathbf{L}(\omega_1)\underline{J}_T(\omega_1), \ldots, \mathbf{L}(\omega_T)\underline{J}_T(\omega_T)\}$, is asymptotically an uncorrelated sequence with a constant variance.

Of course, in reality the spectral density matrix $\mathbf{f}(\omega)$ is unknown and has to be estimated from the data. Let $\widehat{\mathbf{f}}_T(\omega)$ be a nonparametric estimate of $\mathbf{f}(\omega)$, where

$$\widehat{\mathbf{f}}_T(\omega) = \frac{1}{2\pi T}\sum_{t,\tau=1}^{T} \lambda_b(t-\tau)\exp(i(t-\tau)\omega)\big(\underline{X}_t - \underline{\bar{X}}\big)\big(\underline{X}_\tau - \underline{\bar{X}}\big)' \quad \omega \in [0, 2\pi], \tag{2.4}$$

$\{\lambda_b(r) = \lambda(br)\}$ are the lag weights and $\underline{\bar{X}} = \frac{1}{T}\sum_{t=1}^{T} \underline{X}_t$. Below we state the assumptions we require on the lag window, which we use throughout this article.

Assumption 2.1 (The lag window and bandwidth) *(K1) The lag window $\lambda : \mathbb{R} \to \mathbb{R}$, where $\lambda(\cdot)$ has a compact support $[1, 1]$, is symmetric about 0, $\lambda(0) = 1$, the derivative $\lambda'(u)$ exists in $(0, 1)$ and is bounded. Some consequences of the above conditions are $\sum_r |\lambda_b(r)| = O(b^{-1})$, $\sum_r |r| \cdot |\lambda_b(r)| = O(b^{-2})$ and $|\lambda_b(r) - 1| \leq \sup_u |\lambda'(u)| \cdot |rb|$.*

(K2) $T^{-1/2} << b << T^{-1/4}$.

Let $\widehat{\mathbf{f}}_T(\omega_k) = \widehat{\mathbf{B}}(\omega_k)\overline{\widehat{\mathbf{B}}(\omega_k)}'$, where $\widehat{\mathbf{B}}(\omega_k)$ is the (lower-triangular) Cholesky decomposition of $\widehat{\mathbf{f}}_T(\omega_k)$ and $\widehat{\mathbf{L}}(\omega_k) := \widehat{\mathbf{B}}^{-1}(\omega_k)$. Thus $\widehat{\mathbf{B}}(\omega_k)$ and $\widehat{\mathbf{L}}(\omega_k)$ are estimators of $\mathbf{B}(\omega_k)$ and $\mathbf{L}(\omega_k)$ respectively.

Using the above spectral density matrix estimator, we now define the weighted DFT covariance matrix at lags r and ℓ

$$\widehat{\mathbf{C}}_T(r,\ell) = \frac{1}{T}\sum_{k=1}^{T} \widehat{\mathbf{L}}(\omega_k)\underline{J}_T(\omega_k)\overline{\underline{J}_T(\omega_{k+r})}'\overline{\widehat{\mathbf{L}}(\omega_{k+r})}'\exp(i\ell\omega_k), \quad r > 0 \text{ and } \ell \in \mathbb{Z}. \tag{2.5}$$

We observe that due to the periodicity of the DFT, $\widehat{\mathbf{C}}_T(r,\ell)$ is also periodic in r, where $\widehat{\mathbf{C}}_T(r,\ell) = \widehat{\mathbf{C}}_T(r+T,\ell)$ for all $r \in \mathbb{Z}$. To understand the motivation behind this definition, we recall that frequency domain methods for stationary time series use similar statistics. For example, allowing $r = 0$ we observe that in the univariate case $\widehat{\mathbf{C}}_T(0,0)$ corresponds to the classical Whittle likelihood (where $\widehat{L}(\omega_k)$ is replaced with the square-root inverse of a spectral density function with a parametric form, see for example, Whittle (1953), Walker (1963) and Eichler (2012)). Likewise, by removing $\widehat{L}(\omega_k)$ from the definition, we find that $\widehat{\mathbf{C}}_T(0,\ell)$ corresponds to the sample Yule-Walker autocovariance of $\{\underline{X}_t\}$ at lag ℓ. The fundamental difference between the DFT covariance and frequency domain ratio statistics methods for stationary time series is that the periodogram $\underline{J}_T(\omega_k)\overline{\underline{J}_T(\omega_k)}'$ has been replaced with $\underline{J}_T(\omega_k)\overline{\underline{J}_T(\omega_{k+r})}'$, and it is this that facilitates the detection of second order nonstationary behavior.

Example 2.1 *We illustrate the above for the univariate case ($d = 1$). If the time series is second order stationary, then* $\mathrm{E}|J_T(\omega)|^2 \to f(\omega)$, *which means* $\mathrm{E}|f(\omega)^{-1/2}J_T(\omega)|^2 \to 1$. *The corresponding weighted DFT covariance is*

$$\widehat{\mathbf{C}}_T(r,\ell) = \frac{1}{T}\sum_{k=1}^{T}\frac{J_T(\omega_k)\overline{J_T(\omega_{k+r})}}{\widehat{f}_T(\omega_k)^{1/2}\widehat{f}_T(\omega_{k+r})^{1/2}}\exp(i\ell\omega_k) \quad r > 0 \text{ and } \ell \in \mathbb{Z}.$$

We will show later in this section that under Gaussianity, the asymptotic variance of $\widehat{\mathbf{C}}_T(r,\ell)$ *does not depend on any nuisance parameters. One can also define the DFT covariance without standardizing with* $f(\omega)^{-1/2}$. *However, the variance of the non-standardized DFT covariance is a function of the spectral density function and only detects changes in the autocovariance function at lag* ℓ.

In later sections, we derive the asymptotic distribution properties of $\widehat{\mathbf{C}}_T(r,\ell)$. In particular, we show that under second order stationarity (and some additional technical conditions)

$$\sqrt{T}\begin{pmatrix} \Re\widehat{\mathbf{K}}_n(1) \\ \Im\widehat{\mathbf{K}}_n(1) \\ \vdots \\ \Re\widehat{\mathbf{K}}_n(m) \\ \Im\widehat{\mathbf{K}}_n(m) \end{pmatrix} \overset{D}{\to} \mathcal{N}\left(\underline{0}_{2mn}, \begin{pmatrix} \mathbf{W}_n & 0 & 0 & \ldots & 0 \\ 0 & \mathbf{W}_n & 0 & \ldots & 0 \\ \ldots & \ddots & \vdots & \ldots & \\ 0 & 0 & \ldots & \mathbf{W}_n & 0 \\ 0 & 0 & \ldots & 0 & \mathbf{W}_n \end{pmatrix}\right), \tag{2.6}$$

as $T \to \infty$, where

$$\widehat{\mathbf{K}}_n(r) = \Big(\mathrm{vech}(\widehat{\mathbf{C}}_T(r,0))', \mathrm{vech}(\widehat{\mathbf{C}}_T(r,1))', \ldots, \mathrm{vech}(\widehat{\mathbf{C}}_T(r,n-1))'\Big)'. \tag{2.7}$$

This result is used to define the test statistic in Section 2.3. However, in order to construct the test statistic, we need to understand $\mathbf{W}_n$. Therefore, for the remainder of this section, we will discuss (2.6) and the form that $\mathbf{W}_n$ takes for various stationary time series (the remainder of this section can be skipped on first reading).

The DFT covariance of univariate stationary time series

We first consider the case that $\{X_t\}$ is a univariate, fourth order stationary (to be precisely defined in Assumption 3.1) time series. To detect nonstationarity, we will consider the DFT covariance over various lags of ℓ and define the vector

$$\widehat{\mathbf{K}}_n(r)' = (\widehat{\mathbf{C}}_T(r,0),\ldots,\widehat{\mathbf{C}}_T(r,n-1)).$$

Since $\widehat{\mathbf{K}}_n(r)$ is a complex random vector we consider separately the real and imaginary parts denoted by $\Re\widehat{\mathbf{K}}_n(r)$ and $\Im\widehat{\mathbf{K}}_n(r)$, respectively. In the simple case that $\{X_t\}$ is a univariate stationary Gaussian time series, it can be shown that the asymptotic normality result in (2.6) holds, where

$$\mathbf{W}_n = \frac{1}{2}\mathrm{diag}(2,\underbrace{1,1,\ldots,1}_{n-1}) \tag{2.8}$$

and $\underline{0}_d$ denotes the d-dimensional zero vector. Therefore, for stationary Gaussian time series, the distribution of $\widehat{\mathbf{K}}_n(r)$ is asymptotically pivotal (does not depend on any unknown parameters). However, if we were to relax the assumption of Gaussianity, then a similar result holds but $\mathbf{W}_n$ is more complex

$$\mathbf{W}_n = \frac{1}{2}\mathrm{diag}(2,\underbrace{1,1,\ldots,1}_{n-1}) + \mathbf{W}_n^{(2)},$$

where the (ℓ_1+1,ℓ_2+1)th element of $\mathbf{W}^{(2)}$ is $\mathbf{W}^{(2)}_{\ell_1+1,\ell_2+1} = \frac{1}{2}\kappa^{(\ell_1,\ell_2)}$ with

$$\kappa^{(\ell_1,\ell_2)} = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{2\pi}\frac{f_4(\lambda_1,-\lambda_1,-\lambda_2)}{f(\lambda_1)f(\lambda_2)}\exp(i\ell_1\lambda_1 - i\ell_2\lambda_2)d\lambda_1 d\lambda_2 \tag{2.9}$$

and f_4 is the tri-spectral density $f_4(\lambda_1,\lambda_2,\lambda_3) = \frac{1}{(2\pi)^3}\sum_{t_1,t_2,t_3=-\infty}^{\infty}\kappa_4(t_1,t_2,t_3)exp(-i(t_1\lambda_1 + t_2\lambda_2 + t_3\lambda_3))$ and $\kappa_4(t_1,t_2,t_3) = cum(X_0,X_{t_1},X_{t_2},X_{t_3})$ (for statistical properties of the tri-spectral density see Brillinger (1981), Subba Rao and Gabr (1984) and Terdik (1999)). $\kappa^{(\ell_1,\ell_2)}$ can be rewritten in terms of fourth order cumulants by observing that if we define the pre-whitened time series $\{Z_t\}$ (where $\{Z_t\}$ is a linear transformation of $\{X_t\}$ which is uncorrelated) then

$$\kappa^{(\ell_1,\ell_2)} = \sum_{h\in\mathbb{Z}}\mathrm{cum}(Z_0,Z_h,Z_{h+\ell_1},Z_{\ell_2}). \tag{2.10}$$

The expression for $\mathbf{W}_n^{(2)}$ is unwieldy, but in certain situations (besides the Gaussian case) it has a simple form. For example, in the case that the time series $\{X_t\}$ is non-Gaussian, but linear with transfer function $A(\lambda)$, and innovations with variance σ^2 and fourth order cumulant κ_4, respectively, then the above reduces to

$$\kappa^{(\ell_1,\ell_2)} = \int \frac{\kappa_4 |A(\lambda_1)A(\lambda_2)|^2}{\sigma^4 |A(\lambda_1)^2| A(\lambda_2)|^2} \exp(i\ell_1\lambda_1 - i\ell_2\lambda_2) d\lambda_1 d\lambda_2 = \frac{\kappa_4}{\sigma^4}\delta_{\ell_1,0}\delta_{\ell_2,0},$$

where δ_{jk} is the Kronecker delta. Therefore, for (univariate) linear time series, we have $\mathbf{W}_n^{(2)} = \frac{\kappa_4}{2\sigma^4} I_n$ and $\mathbf{W}_n$ is still a diagonal matrix. This example illustrates that even in the univariate case the complexity of the variance of the DFT covariance $\widehat{\mathbf{K}}_n(r)$ increases the more we relax the assumptions on the distribution. Regardless of the distribution of $\{X_t\}$, so long as it satisfies (2.1) (and some mixing-type assumptions), then $\widehat{\mathbf{K}}_n(r)$ is asymptotically normal and centered about zero.

The DFT covariance of multivariate stationary time series

We now consider the distribution of $\widehat{\mathbf{C}}_T(r,\ell)$ in the multivariate case. We will show in Lemma A.11 (in the appendix) that the covariance of $\widehat{\mathbf{C}}_T(r,0)$ is singular. To avoid the singularity we will only consider the lower triangular vectorized version of $\widehat{\mathbf{C}}_T(r,\ell)$, i.e.

$$\text{vech}(\widehat{\mathbf{C}}_T(r,\ell)) = (\widehat{c}_{1,1}(r,\ell), \widehat{c}_{2,1}(r,\ell), \ldots, \widehat{c}_{d,1}(r,\ell), \widehat{c}_{2,2}(r,\ell), \ldots, \widehat{c}_{d,2}(r,\ell), \ldots, \widehat{c}_{d,d}(r,\ell))',$$

where $\widehat{c}_{j_1,j_2}(r,\ell)$ is the (j_1,j_2)th element of $\widehat{\mathbf{C}}_T(r,\ell)$, and we use this to define the $nd(d+1)/2$-dimensional vector $\widehat{\mathbf{K}}_n(r)$ (given in (2.7)). In the case that $\{\underline{X}_t\}$ is a Gaussian stationary time series then we obtain an analogous result to (2.8) where similar to the univariate case $\mathbf{W}_n$ is a diagonal matrix with $\mathbf{W}_n = \text{diag}(\mathbf{W}_0^{(1)}, \ldots, \mathbf{W}_{n-1}^{(1)})$, where

$$\mathbf{W}_\ell^{(1)} = \begin{cases} \frac{1}{2} I_{d(d+1)/2} & \ell \neq 0 \\ \text{diag}(\lambda_1, \ldots, \lambda_{d(d+1)/2}) & \ell = 0 \end{cases} \tag{2.11}$$

with

$$\lambda_j = \begin{cases} 1, & j \in \left\{1 + \sum_{n=m+1}^{d} n \text{ for } m \in \{1,2,\ldots,d\}\right\} \\ \frac{1}{2}, & \text{otherwise} \end{cases}.$$

However, in the non-Gaussian case $\mathbf{W}_n$ is equal to the above diagonal matrix plus an additional (not necessarily diagonal) matrix consisting of the fourth order spectral densities, ie. $\mathbf{W}_n$ consists of $n^2\ d(d+1)/2$ square blocks, where the (ℓ_1+1, ℓ_2+1) block is

$$(\mathbf{W}_n)_{\ell_1+1,\ell_1+1} = \mathbf{W}_{\ell_1}^{(1)}\delta_{\ell_1,\ell_2} + \mathbf{W}_{\ell_1,\ell_2}^{(2)}, \tag{2.12}$$

where $\mathbf{W}_\ell^{(1)}$ and $\mathbf{W}_{\ell_1,\ell_2}^{(2)}$ are defined in (2.11) and in (2.15) below. In order to appreciate the structure of $\mathbf{W}_{\ell_1,\ell_2}^{(2)}$, we first consider some examples. We start by defining the multivariate version of (2.9)

$$\begin{aligned}&\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)\\ =\ &\frac{1}{2\pi}\int_0^{2\pi}\int_0^{2\pi}\sum_{s_1,s_2,s_3,s_4=1}^{d} L_{j_1s_1}(\lambda_1)\overline{L_{j_2s_2}(\lambda_1)L_{j_3s_3}(\lambda_2)}L_{j_4s_4}(\lambda_2)\exp(i\ell_1\lambda_1-i\ell_2\lambda_2)\\ &\times f_{4;s_1,s_2,s_3,s_4}(\lambda_1,-\lambda_1,-\lambda_2)d\lambda_1 d\lambda_2,\end{aligned} \tag{2.13}$$

where

$$f_{4;s_1,s_2,s_3,s_4}(\lambda_1,\lambda_2,\lambda_3)=\frac{1}{(2\pi)^3}\sum_{t_1,t_2,t_3=-\infty}^{\infty}\kappa_{4;s_1,s_2,s_3,s_4}(t_1,t_2,t_3)\exp(i(-t_1\lambda_1-t_2\lambda_2-t_3\lambda_3))$$

is the joint tri-spectral density of $\{\underline{X}_t\}$ and

$$\kappa_{4;s_1,s_2,s_3,s_4}(t_1,t_2,t_3)=\mathrm{cum}(X_{0,s_1},X_{t_1,s_2},X_{t_2,s_3},X_{t_3,s_4}). \tag{2.14}$$

We note that a similar expression to (2.10) can be derived for $\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)$, with

$$\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)=\sum_{h\in\mathbb{Z}}\mathrm{cum}(Z_{j_1,0},Z_{j_2,h},Z_{j_3,h+\ell_1},Z_{j_4,\ell_2}).$$

where $\{\underline{Z}_t=(Z_{1,t},\ldots,Z_{d,t})'\}$ is the decorrelated (or prewhitened) version of $\{\underline{X}_t\}$.

Example 2.2 (Structure of $\mathbf{W}_n$) *For $n\in\mathbb{N}$ and $\ell_1,\ell_2\in\{0,\ldots,n-1\}$, we have $(\mathbf{W}_n)_{\ell_1+1,\ell_2+1}=\mathbf{W}_{\ell_1}^{(1)}\delta_{\ell_1,\ell_2}+\mathbf{W}_{\ell_1,\ell_2}^{(2)}$, where:*

(i) For $d=2$, we have $\mathbf{W}_\ell^{(1)}=\frac{1}{2}diag(2,1,2)$ and for $\ell\geq 1$ $\mathbf{W}_\ell^{(1)}=\frac{1}{2}I_3$ and

$$\mathbf{W}_{\ell_1,\ell_2}^{(2)}=\frac{1}{2}\begin{pmatrix}\kappa^{(\ell_1,\ell_2)}(1,1,1,1) & \kappa^{(\ell_1,\ell_2)}(1,1,2,1) & \kappa^{(\ell_1,\ell_2)}(1,1,2,2)\\ \kappa^{(\ell_1,\ell_2)}(2,1,1,1) & \kappa^{(\ell_1,\ell_2)}(2,1,2,1) & \kappa^{(\ell_1,\ell_2)}(2,1,2,2)\\ \kappa^{(\ell_1,\ell_2)}(2,2,1,1) & \kappa^{(\ell_1,\ell_2)}(2,2,2,1) & \kappa^{(\ell_1,\ell_2)}(2,2,2,2)\end{pmatrix}$$

(ii) For $d=3$, we have $\mathbf{W}_{\ell_1}^{(1)}=\frac{1}{2}diag(2,1,1,2,1,2)$ for $\ell\geq 1$ $\mathbf{W}_\ell^{(1)}=I_6$ and $\mathbf{W}_{\ell_1,\ell_2}^{(2)}$ is analogous to (i).

(iv) For general d and $n=1$, we have $\mathbf{W}_n=\mathbf{W}_0^{(1)}+\mathbf{W}^{(2)}$, where $\mathbf{W}_0^{(1)}$ is the diagonal matrix defined in (2.11) and $\mathbf{W}^{(2)}=\mathbf{W}_{0,0}^{(2)}$ (which is defined in (2.15).

We now define the general form of $\mathbf{W}^{(2)}$

$$\mathbf{W}_{\ell_1,\ell_2}^{(2)}=\mathbf{E}_d\mathbf{V}_{\ell_1,\ell_2}^{(2)}\mathbf{E}_d, \tag{2.15}$$

where $\mathbf{E}_d$ with $\mathbf{E}_d\text{vec}(\mathbf{A}) = \text{vech}(\mathbf{A})$ is the $(d(d+1)/2 \times d^2)$ elimination matrix [cf. Lütkepohl (2006), p.662] that transforms the vec-version of a $(d \times d)$ matrix $\mathbf{A}$ to its vech-version and the entry (j_1, j_2) of the $(d^2 \times d^2)$ matrix $\mathbf{V}^{(2)}_{\ell_1,\ell_2}$ is such that

$$(V^{(2)}_{\ell_1,\ell_2})_{j_1,j_2} = \kappa^{(\ell_1,\ell_2)}\left((j_1-1)\text{mod } d+1, \left\lceil \frac{j_1}{d} \right\rceil, (j_2-1)\text{mod } d+1, \left\lceil \frac{j_2}{d} \right\rceil\right), \tag{2.16}$$

respectively, where $\lceil x \rceil$ is the smallest integer greater than or equal to x.

Example 2.3 ($\kappa^{(\ell_1,\ell_2)}(j_1, j_2, j_3, j_4)$ under linearity of $\{\underline{X}_t\}$) *Suppose the additional assumption of linearity of the process $\{\underline{X}_t\}$ is satisfied, that is, $\{\underline{X}_t\}$ satisfies a representation*

$$\underline{X}_t = \sum_{\nu=-\infty}^{\infty} \mathbf{\Gamma}_\nu \underline{e}_{t-\nu}, \quad t \in \mathbb{Z}, \tag{2.17}$$

where $\sum_{\nu=-\infty}^{\infty} |\mathbf{\Gamma}_\nu|_1 < \infty$, $\mathbf{\Gamma}_0 = \mathbf{I}_d$ and $\{\underline{e}_t, t \in \mathbb{Z}\}$ are zero mean, i.i.d. random vectors with $\text{E}(\underline{e}_t\underline{e}_t') = \mathbf{\Sigma}_e$ *positive definite and whose fourth moments exist. By plugging-in* (2.17) *in* (2.14) *and then evaluating the integrals in* (2.13), *the quantity $\kappa^{(\ell_1,\ell_2)}(j_1, j_2, j_3, j_4)$ becomes*

$$\begin{aligned}
&\kappa^{(\ell_1,\ell_2)}(j_1, j_2, j_3, j_4) \\
= &\sum_{s_1,s_2,s_3,s_4=1}^{d} \kappa_{4,s_1,s_2,s_3,s_4} \left\{\frac{1}{2\pi}\int_0^{2\pi} (\mathbf{L}(\lambda_1)\mathbf{\Gamma}(\lambda_1))_{j_1s_1} \overline{(\mathbf{L}(\lambda_1)\mathbf{\Gamma}(\lambda_1))}_{j_2s_2} \exp(i\ell_1\lambda_1)d\lambda_1\right\} \\
&\times \left\{\frac{1}{2\pi}\int_0^{2\pi} \overline{(\mathbf{L}(\lambda_2)\mathbf{\Gamma}(\lambda_2))}_{j_3s_3} (\mathbf{L}(\lambda_2)\mathbf{\Gamma}(\lambda_2))_{j_4s_4} \exp(-i\ell_2\lambda_2)d\lambda_2\right\},
\end{aligned}$$

where $\mathbf{\Gamma}(\omega) = \frac{1}{\sqrt{2\pi}}\sum_{\nu=-\infty}^{\infty} \mathbf{\Gamma}_\nu e^{-i\nu\omega}$ is the transfer function of $\{\underline{X}_t\}$ and $\kappa_{4,s_1,s_2,s_3,s_4} = \text{cum}(\underline{e}_{0,s_1}, \underline{e}_{0,s_2}, \underline{e}_{0,s_3}, \underline{e}_{0,s_4})$. The shape of $\kappa^{(\ell_1,\ell_2)}(j_1, j_2, j_3, j_4)$ is now discussed for two special cases of linearity.

(i) If $\mathbf{\Gamma}_\nu = \mathbf{0}$ for $\nu \neq 0$, we have $\underline{X}_t = \underline{e}_t$ and $\kappa^{(\ell_1,\ell_2)}(j_1, j_2, j_3, j_4)$ simplifies to

$$\kappa^{(\ell_1,\ell_2)}(j_1, j_2, j_3, j_4) = \widetilde{\kappa}_{4,j_1,j_2,j_3,j_4}\delta_{\ell_1 0}\delta_{\ell_2 0},$$

where $\mathbf{\Sigma}_e^{-1/2}\underline{e}_t = (\tilde{e}_{t,1}, \ldots, \tilde{e}_{t,d})'$ and $\widetilde{\kappa}_{4,s_1,s_2,s_3,s_4} = \text{cum}(\tilde{e}_{0,s_1}, \tilde{e}_{0,s_2}, \tilde{e}_{0,s_3}, \tilde{e}_{0,s_4})$.

(ii) The univariate time series $\{X_{t,k}\}$ are independent for $k = 1, \ldots, d$ (the components of $\underline{X}_t$ are independent), then we have

$$\kappa^{(\ell_1,\ell_2)}(j_1, j_2, j_3, j_4) = \kappa_{4,j}\delta_{\ell_1 0}\delta_{\ell_2 0}1(j_1 = j_2 = j_3 = j_4 = j),$$

where $\kappa_{4,j} = \text{cum}_4(\underline{\varepsilon}_{0,j})/\sigma_j^4$ and $\mathbf{\Sigma}_e = \text{diag}(\sigma_1^2, \ldots, \sigma_d^2)$.

2.3 The test statistic

We now use the results in the previous section to motivate the test statistic. We have shown that $\{\widehat{\mathbf{K}}_n(r)\}_r$ (and also $\Re\widehat{\mathbf{K}}_n(r)$ and $\Im\widehat{\mathbf{K}}_n(r)$) are asymptotically uncorrelated. Therefore, we simply standardize $\{\widehat{\mathbf{K}}_n(r)\}$ and define the test statistic

$$\mathcal{T}_{m,n,d} = T\sum_{r=1}^{m}\left(|\mathbf{W}_n^{-1/2}\Re\widehat{\mathbf{K}}_n(r)|_2^2 + |\mathbf{W}_n^{-1/2}\Im\widehat{\mathbf{K}}_n(r)|_2^2\right), \tag{2.18}$$

where $\widehat{\mathbf{K}}_n(r)$ and $\mathbf{W}_n$ are defined in (2.7) and (2.12) respectively. By using (2.6), it is clear that

$$\mathcal{T}_{m,n,d} \overset{D}{\to} \chi^2_{mnd(d+1)}, \tag{2.19}$$

where $\chi^2_{mnd(d+1)}$ is a χ^2-distribution with $mnd(d+1)$ degrees of freedom.

Therefore, using the above result, we reject the null of second order stationarity at the $\alpha \times 100\%$ level if $\mathcal{T}_{m,n,d} > \chi^2_{mnd(d+1)}(1-\alpha)$, where $\chi^2_{mnd(d+1)}(1-\alpha)$ is the $(1-\alpha)$-quantile of the χ^2-distribution with $mnd(d+1)$ degrees of freedom.

Example 2.4 *(i) In the univariate case using $n = 1$, the test statistic reduces to*

$$\mathcal{T}_{m,1,1} = \sum_{r=1}^{m}\frac{|\widehat{\mathbf{C}}_T(r,0)|^2}{1+\frac{1}{2}\kappa^{(0,0)}},$$

where $\kappa^{(0,0)}$ is defined in (2.9).

(ii) In most situations, it is probably enough to use $n = 1$. In this case the test statistic reduces to

$$\mathcal{T}_{m,1,d} = T\sum_{r=1}^{m}\left(|\mathbf{W}_1^{-1/2}vech(\Re\widehat{\mathbf{C}}_n(r,0))|_2^2 + |\mathbf{W}_1^{-1/2}vech(\Im\widehat{\mathbf{C}}_n(r,0))|_2^2\right).$$

(iii) If we can assume that $\{\underline{X}_t\}$ is Gaussian, then $\mathcal{T}_{m,n,d}$ has the simple form

$$\begin{aligned}\mathcal{T}_{m,n,d,G} &= T\sum_{r=1}^{m}\left(|(\mathbf{W}_0^{(1)})^{-1/2}vech(\Re\widehat{\mathbf{C}}_T(r,0))|_2^2 + |(\mathbf{W}_0^{(1)})^{-1/2}vech(\Im\widehat{\mathbf{C}}_T(r,0))|_2^2\right)\\ &\quad +2T\sum_{r=1}^{m}\sum_{\ell=1}^{n-1}\left(|vech(\Re\widehat{\mathbf{C}}_T(r,\ell))|_2^2 + |vech(\Im\widehat{\mathbf{C}}_T(r,\ell))|_2^2\right),\end{aligned} \tag{2.20}$$

where $\mathbf{W}_0^{(1)}$ is a diagonal matrix composed of ones and halves defined in (2.11).

The above test statistic was constructed as if the standardization matrix $\mathbf{W}_n$ were known. However, only in the case of Gaussianity this matrix will be known, for non-Gaussian time series we need to estimate it. In the following section, we propose a bootstrap method for estimating $\mathbf{W}_n$.

2.4 A bootstrap estimator of the variance $\mathbf{W}_n$

The proposed test does not make any model assumptions on the underlying time series. This level of generality means that the test statistic involves unknown parameters which, in practice, can be extremely difficult to directly estimate. The objective of this section is to construct a consistent estimator of these unknown parameters. We propose an estimator of the asymptotic variance matrix $\mathbf{W}_n$ using a block bootstrap procedure. There exists several well known block bootstrap methods, (cf. Lahiri (2003) for a review), but the majority of these sampling schemes, are nonstationary when conditioned on the original time series. An exception is the stationary bootstrap, proposed in Politis and Romano (1994) (see also Parker, Paparoditis, and Politis (2006)), which is designed such that the bootstrap distribution is stationary. As we are testing for stationarity, we use the stationary bootstrap to estimate the variance.

The bootstrap testing scheme

Step 1. Given the d-variate observations $\underline{X}_1, \ldots, \underline{X}_T$, evaluate $\mathrm{vech}(\Re\widehat{\mathbf{C}}_T(r,\ell))$ and $\mathrm{vech}(\Im\widehat{\mathbf{C}}_T(r,\ell))$ for $r = 1, \ldots, m$ and $\ell = 0, \ldots, n-1$.

Step 2. Define the blocks

$$B_{I,L} = \{\underline{Y}_I, \ldots, \underline{Y}_{I+L-1}\},$$

where $\underline{Y}_j = \underline{X}_{j \bmod T} - \overline{\underline{X}}$ (hence there is wrapping on a torus if $j > T$) and $\overline{\underline{X}} = \frac{1}{T}\sum_{t=1}^{T}\underline{X}_t$. We will suppose that the points on the time series $\{I_i\}$ and the block length $\{L_i\}$ are iid random variables, where $P(I_i = s) = T^{-1}$ for $1 \leq s \leq T$ (discrete uniform distribution) and $P(L_i = s) = p(1-p)^{s-1}$ for $s \geq 1$ (geometric distribution).

Step 3. We draw blocks $\{B_{I_i,L_i}\}_i$ until the total length of the blocks $(B_{I_1,L_1}, \ldots, B_{I_r,L_r})$ satisfies $\sum_{i=1}^{r} L_i \geq T$ and we discard the last $\sum_{i=1}^{r} L_i - T$ values to get a bootstrap sample $\underline{X}_1^*, \ldots, \underline{X}_T^*$.

Step 4. Define the bootstrap spectral density estimator

$$\widehat{\mathbf{f}}_T^*(\omega_k) = \frac{1}{T}\sum_{j=-\lfloor\frac{T-1}{2}\rfloor}^{\lfloor\frac{T}{2}\rfloor} K_b(\omega_k - \omega_j)\underline{J}_T^*(\omega_j)\overline{\underline{J}_T^*(\omega_j)}', \tag{2.21}$$

where $K_b(\omega_j) = \sum_r \lambda_b(r)\exp(ir\omega_j)$, its lower-triangular Cholesky matrix $\widehat{\mathbf{B}}^*(\omega)$, its inverse $\widehat{\mathbf{L}}^*(\omega) = (\widehat{\mathbf{B}}^*(\omega))^{-1}$ and the bootstrap DFT covariances

$$\widehat{\mathbf{C}}_T^*(r,\ell) = \frac{1}{T}\sum_{k=1}^{T}\widehat{\mathbf{L}}^*(\omega_k)\underline{J}_T^*(\omega_k)\overline{\underline{J}_T^*(\omega_{k+r})}'\overline{\widehat{\mathbf{L}}^*(\omega_{k+r})}'\exp(i\ell\omega_k), \tag{2.22}$$

where $\underline{J}^*_T(\omega_k) = \frac{1}{\sqrt{2\pi T}}\sum_{t=1}^{T}\underline{X}^*_t e^{-it\omega_k}$ is the bootstrap DFT.

Step 5. Repeat Steps 1-4 N times (where N is large), to obtain $\mathrm{vech}(\Re\widehat{\mathbf{C}}^*_T(r,\ell))^{(j)}$ and $\mathrm{vech}(\Im\widehat{\mathbf{C}}^*_T(r,\ell))^{(j)}$, $j=1,\ldots,N$. For $r=1,\ldots,m$ and $\ell_1,\ell_2 = 0,\ldots,n-1$, we compute the bootstrap covariance estimators of the real parts that is

$$\begin{aligned}\{\widehat{\mathbf{W}}^*_{\Re}(r)\}_{\ell_1+1,\ell_2+1} &= T\Bigg(\frac{1}{N}\sum_{j=1}^{N}\mathrm{vech}(\Re\widehat{\mathbf{C}}^*_T(r,\ell_1))^{(j)}\mathrm{vech}(\Re\widehat{\mathbf{C}}^*_T(r,\ell_2))^{(j)\prime} \\ &\quad -\left(\frac{1}{N}\sum_{j=1}^{N}\mathrm{vech}(\Re\widehat{\mathbf{C}}^*_T(r,\ell_1))^{(j)}\right)\left(\frac{1}{N}\sum_{j=1}^{N}\mathrm{vech}(\Re\widehat{\mathbf{C}}^*_T(r,\ell_2))^{(j)}\right)^{\prime}\Bigg)\end{aligned} \tag{2.23}$$

and, similarly, we define its analogues $\{\widehat{\mathbf{W}}^*_{\Im}(r)\}_{\ell_1+1,\ell_2+1}$ using the imaginary parts.

Step 6. Define the bootstrap covariance estimator $\{\widehat{\mathbf{W}}^*(r)\}_{\ell_1+1,\ell_2+1}$ as

$$\{\widehat{\mathbf{W}}^*(r)\}_{\ell_1+1,\ell_2+1} = \frac{1}{2}\left[\{\widehat{\mathbf{W}}^*_{\Re}(r)\}_{\ell_1+1,\ell_2+1} + \{\widehat{\mathbf{W}}^*_{\Im}(r)\}_{\ell_1+1,\ell_2+1}\right],$$

$\widehat{\mathbf{W}}^*(r)$ is the bootstrap estimator of the rth block of $\mathbf{W}_{m,n}$ defined in (3.6).

Step 7. Finally, define the bootstrap test statistic $\mathcal{T}^*_{m,n,d}$ as

$$\mathcal{T}^*_{m,n,d} = T\sum_{r=1}^{m}\left(|[\widehat{\mathbf{W}}^*(r)]^{-1/2}\mathrm{vech}(\Re\widehat{\mathbf{K}}_n(r))|_2^2 + |[\widehat{\mathbf{W}}^*(r)]^{-1/2}\mathrm{vech}(\Im\widehat{\mathbf{K}}_n(r))|_2^2\right) \tag{2.24}$$

and reject H_0 if $\mathcal{T}^*_{m,n,d} > \chi^2_{mnd(d+1)}(1-\alpha)$, where $\chi^2_{mnd(d+1)}(1-\alpha)$ is the $(1-\alpha)$-quantile of the χ^2-distribution with $mnd(d+1)$ degrees of freedom to obtain a test of asymptotic level $\alpha\in(0,1)$.

Remark 2.1 (Step 4^*) *A simple variant of the above bootstrap, is to use the spectral density estimator $\widehat{\mathbf{f}}_T(\omega)$ rather than bootstrap spectral density estimator $\widehat{\mathbf{f}}^*_T(\omega)$ ie.*

$$\acute{\mathbf{C}}^*_T(r,\ell) = \frac{1}{T}\sum_{k=1}^{T}\widehat{\mathbf{L}}(\omega_k)\underline{J}^*_T(\omega_k)\overline{\underline{J}^*_T(\omega_{k+r})}^{\prime}\overline{\widehat{\mathbf{L}}(\omega_{k+r})}^{\prime}\exp(i\ell\omega_k). \tag{2.25}$$

Using the above bootstrap covariance greatly simplifies the speed of the bootstrap procedure and the theoretical analysis of the bootstrap (in particular the assumptions required). However, empirical evidence suggests that estimating the spectral density matrix at each bootstrap sample gives a better finite sample approximation of the variance (though we cannot theoretically prove that using $\widehat{\mathbf{C}}^*_T(r,\ell)$ gives a better variance approximation than $\acute{\mathbf{C}}^*_T(r,\ell)$).

We observe that because the blocks are random and their length is determined by a geometric distribution, their lengths vary. However, the mean length of a block is approximately $1/p$ (only approximately since only block lengths less than length T are used in the scheme). As it has to be assumed that $p \to 0$ and $Tp \to \infty$ as $T \to \infty$, the mean block length increases as the sample size T grows. However, we will show in Section 5 that a sufficient condition for consistency of the stationary bootstrap estimator is that $Tp^4 \to \infty$ as $T \to \infty$. This condition constrains the mean length of the block and prevents it growing too fast.

3 Analysis of the DFT covariance under stationarity and non-stationarity of the time series

3.1 The DFT covariance $\widehat{\mathbf{C}}_T(r,\ell)$ under stationarity

Directly deriving the sampling properties of $\widehat{\mathbf{C}}_T(r,\ell)$ is not possible as it involves the estimators $\widehat{\mathbf{L}}(\omega)$. Instead, in the analysis below, we replace $\widehat{\mathbf{L}}(\omega)$ by its deterministic limit $\mathbf{L}(\omega)$, and consider the quantity

$$\widetilde{\mathbf{C}}_T(r,\ell) = \frac{1}{T}\sum_{k=1}^{T} \mathbf{L}(\omega_k)\underline{J}_T(\omega_k)\overline{\underline{J}_T(\omega_{k+r})}'\,\overline{\mathbf{L}(\omega_{k+r})}'\exp(i\ell\omega_k). \tag{3.1}$$

Below, we show that $\widehat{\mathbf{C}}_T(r,\ell)$ and $\widetilde{\mathbf{C}}_T(r,\ell)$ are asymptotically equivalent. This allows us to analyze $\widetilde{\mathbf{C}}_T(r,\ell)$ without any loss in generality. We will require the following assumptions.

3.1.1 Assumptions

Let $|\cdot|_p$ denote the ℓ_p-norm of a vector or matrix, i.e. $|A|_p = (\sum_{i,j}|A_{ij}|^p)^{1/p}$ for some matrix $A = (a_{ij})$ and let $\|\underline{X}\|_p = (\mathrm{E}|\underline{X}|^p)^{1/p}$.

Assumption 3.1 (The process $\{\underline{X}_t\}$) *(P1) Let us suppose that $\{\underline{X}_t, t \in \mathbb{Z}\}$ is a d-variate constant mean, fourth order stationary (ie. the first, second, third and fourth order moments of the time series are invariant to shift), α-mixing time series which satisfies*

$$\sup_{k\in\mathbb{Z}} \sup_{\substack{A\in\sigma(\underline{X}_{t+k},\underline{X}_{t+k+1},\ldots)\\ B\in\sigma(\underline{X}_k,\underline{X}_{k-1},\ldots)}} |P(A\cap B) - P(A)P(B)| \le Ct^{-\alpha}, \quad t > 0, \tag{3.2}$$

where C is a constant and $\alpha > 0$.

(P2) For some $s > \frac{4\alpha}{\alpha-6} > 0$ with α such that (3.2) holds, we have $\sup_{t\in\mathbb{Z}}\|\underline{X}_t\|_s < \infty$.

(P3) The spectral density matrix $\mathbf{f}(\omega)$ is non-singular on $[0, 2\pi]$.

(P4) For some $s > \frac{8\alpha}{\alpha-7} > 0$ *with* α *such that* (3.2) *holds, we have* $\sup_{t\in\mathbb{Z}} \|\underline{X}_t\|_s < \infty$.

(P5) For a given lag order n*, let* $\mathbf{W}_n$ *be the variance matrix defined in (2.12), then* $\mathbf{W}_n$ *is assumed to be non-singular.*

Some comments on the assumptions are in order. The α-mixing assumption is satisfied by a wide range of processes, including, under certain assumptions on the innovations, the vector AR models (see Pham and Tran (1985)) and other Markov models which are irreducible (cf. Feigin and Tweedie (1985), Mokkadem (1990), Meyn and Tweedie (1993), Bousamma (1998), Franke, Stockis, and Tadjuidje-Kamgaing (2010)). We show in Corollary A.1 that Assumption (P2) implies $\sum_{h=-\infty}^{\infty} |h| \cdot |Cov(X_{h,j_1}, X_{0,j_2})| < \infty$ for all $j_1, j_2 = 1, \ldots, d$ and absolute summability of the fourth order cumulants. In addition, Assumption (P2) is required to show asymptotic normality of $\widetilde{\mathbf{C}}_T(r,\ell)$ (using a Mixingale proof). Assumption (P4) is slightly stronger than (P2) and it is used to show the asymptotic equivalence of $\sqrt{T}\widehat{\mathbf{C}}_T(r,\ell)$ and $\sqrt{T}\widetilde{\mathbf{C}}_T(r,\ell)$. In the case that the multivariate time series $\{\underline{X}_t\}$ is geometric mixing, Assumption (P4) implies that for some $\delta > 0$, $(8+\delta)$-moments of $\{\underline{X}_t\}$ should exist. Assumption (P5) is immediately satisfied in the case that $\{\underline{X}_t\}$ is a Gaussian time series, in this case $\mathbf{W}_n$ is a diagonal matrix (see (2.12)).

Remark 3.1 (The fourth order stationarity assumption) *Although the purpose of this paper is to derive a test for second order stationarity, we derive the proposed test statistic under the assumption of fourth order stationarity of* $\{\underline{X}_t\}$ *(see Theorem 3.3). The main advantage of this slightly stronger assumption is that it guarantees that the DFT covariances* $\widehat{\mathbf{C}}_T(r_1,\ell)$ *and* $\widehat{\mathbf{C}}_T(r_2,\ell)$ *are asymptotically uncorrelated at different lags* $r_1 \neq r_2$*. For details see the end of the proof of Theorem 3.2, on the bounds of the fourth order cumulant term).*

3.2 The sampling properties of $\widehat{\mathbf{C}}_T(r,\ell)$ under the assumption of fourth order stationarity

Using the above assumptions we have the following result.

Theorem 3.1 (Asymptotic equivalence of $\widehat{\mathbf{C}}_T(r,\ell)$ and $\widetilde{\mathbf{C}}_T(r,\ell)$ under the null) *Suppose Assumption 3.1 is satisfied and let* $\widehat{\mathbf{C}}_T(r,\ell)$ *and* $\widetilde{\mathbf{C}}_T(r,\ell)$ *be defined as in* (2.5) *and* (3.1)*, respectively. Then we have*

$$\sqrt{T}\big|\widehat{\mathbf{C}}_T(r,\ell) - \widetilde{\mathbf{C}}_T(r,\ell)\big|_1 = O_p\left(\frac{1}{b\sqrt{T}} + b + b^2\sqrt{T}\right).$$

We now obtain the mean and variance of $\widetilde{\mathbf{C}}_T(r,\ell)$ under the stated assumptions. Let $\widetilde{c}_{j_1,j_2}(r,\ell) = \widetilde{\mathbf{C}}_T(r,\ell)_{j_1,j_2}$ denote entry (j_1, j_2) of the unobserved $(d \times d)$ DFT covariance matrix $\widetilde{\mathbf{C}}_T(r,\ell)$.

Theorem 3.2 (First and second order structure of $\{\widetilde{\mathbf{C}}_T(r,\ell)\}$) *Suppose that* $\sup_{j_1,j_2} \sum_h |h| \cdot |\mathrm{cov}(X_{0,j_1}, X_{h,j_2})| < \infty$ *and* $\sup_{j_1,\ldots,j_4} \sum_{h_1,h_2,h_3} |h_i| \cdot |\mathrm{cum}(X_{0,j_1}, X_{h_1,j_2}, X_{h_2}, j_3, X_{h_3,j_4})| < \infty$ *(satisfied by Assumption 3.1(P1,P2)). Then, the following assertions are true*

(i) For all fixed $r \in \mathbb{N}$ and $\ell \in \mathbb{Z}$, we have $E(\widetilde{\mathbf{C}}_T(r,\ell)) = O(\frac{1}{T})$.

(ii) Let $\Re Z$ and $\Im Z$ be the real and the imaginary parts of a random variable Z, respectively. Then, for fixed $r_1, r_2 \in \mathbb{N}$ and $\ell_1, \ell_2 \in \mathbb{Z}$ and all $j_1, j_2, j_3, j_4 \in \{1, \ldots, d\}$, we have

$$\begin{aligned} TCov\left(\Re\widetilde{c}_{j_1,j_2}(r_1,\ell_1), \Re\widetilde{c}_{j_3,j_4}(r_2,\ell_2)\right) &= \frac{1}{2}\left\{\delta_{j_1j_3}\delta_{j_2j_4}\delta_{\ell_1\ell_2} + \delta_{j_1j_4}\delta_{j_2j_3}\delta_{\ell_1,-\ell_2}\right\}\delta_{r_1,r_2} \\ &\quad + \frac{1}{2}\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)\delta_{r_1,r_2} + O\left(\frac{1}{T}\right) \end{aligned} \tag{3.3}$$

and

$$\begin{aligned} TCov\left(\Im\widetilde{c}_{j_1,j_2}(r_1,\ell_1), \Im\widetilde{c}_{j_3,j_4}(r_2,\ell_2)\right) &= \frac{1}{2}\left\{\delta_{j_1j_3}\delta_{j_2j_4}\delta_{\ell_1\ell_2} + \delta_{j_1j_4}\delta_{j_2j_3}\delta_{\ell_1,-\ell_2}\right\}\delta_{r_1,r_2} \\ &\quad + \frac{1}{2}\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)\delta_{r_1,r_2} + O\left(\frac{1}{T}\right), \end{aligned} \tag{3.4}$$

where $\delta_{jk} = 1$ if $j = k$ and $\delta_{jk} = 0$ otherwise.

Below we state the asymptotic normality result, which forms the basis of the test statistic.

Theorem 3.3 (Asymptotic distribution of vech($\widehat{\mathbf{C}}_T(r,\ell)$) under the null) *Suppose Assumptions 3.1 and 2.1 hold. Let the $nd(d+1)/2$-dimensional vector $\widehat{\mathbf{K}}_n(r)$ be defined as in (2.7). Then, for fixed $m, n \in \mathbb{N}$, we have*

$$\sqrt{T}\begin{pmatrix} \Re\widehat{\mathbf{K}}_n(1) \\ \Im\widehat{\mathbf{K}}_n(1) \\ \vdots \\ \Re\widehat{\mathbf{K}}_n(m) \\ \Im\widehat{\mathbf{K}}_n(m) \end{pmatrix} \overset{D}{\to} \mathcal{N}(\underline{0}_{mnd(d+1)}, \mathbf{W}_{m,n}), \tag{3.5}$$

where $\mathbf{W}_{m,n}$ is a $(mnd(d+1) \times mnd(d+1))$ block diagonal matrix

$$\mathbf{W}_{m,n} = diag(\underbrace{\mathbf{W}_n, \ldots, \mathbf{W}_n}_{2m \ times}), \tag{3.6}$$

and $\mathbf{W}_n$ is defined in (2.15).

The above theorem immediately gives the asymptotic distribution of the test statistic.

Theorem 3.4 (Limiting distribution of $\mathcal{T}_{m,n,d}$ under the null) *Let us suppose that Assumptions 3.1 and 2.1 are satisfied. Then we have*

$$\mathcal{T}_{m,n,d} \stackrel{D}{\to} \chi^2_{mnd(d+1)}, \tag{3.7}$$

where $\chi^2_{mnd(d+1)}$ is a χ^2-distribution with $mnd(d+1)$ degrees of freedom.

3.3 Behaviour of $\widehat{\mathbf{C}}_T(r,\ell)$ for locally stationary time series

We now consider the behavior of the DFT covariance $\widehat{\mathbf{C}}_T(r,\ell)$ when the underlying process is second order nonstationary. There are several different alternatives one can consider, including unit root processes, periodically stationary time series, time series with change points etc. However, here we shall focus on time series whose correlation structure changes slowly over time (early work on time-varying time series include Priestley (1965), Subba Rao (1970) and Hallin (1984)). As in nonparametric regression and other work on nonparametric statistics we use the rescaling device to develop the asymptotic theory. The same tool has been used, for example, in nonparametric time series by Robinson (1989) and by Dahlhaus (1997) in his definition of local stationarity. We use rescaling to define a locally stationary process as a time series whose second order structure can be 'locally' approximated by the covariance function of a stationary time series (see Dahlhaus (1997), Dahlhaus and Polonik (2006) and for a recent overview of the current state-of-the-art Dahlhaus (2012)).

3.3.1 Assumptions

In order to prove the results in this paper for the case of local stationarity, we require the following assumptions.

Assumption 3.2 (Locally stationary vector processes) *Let us suppose that the locally stationary process $\{\underline{X}_{t,T}, t \in \mathbb{Z}\}$ is a d-variate, constant mean time series that satisfies the following assumptions:*

(L1) $\{\underline{X}_{t,T}, t \in \mathbb{Z}\}$ is an α-mixing time series with the rate

$$\sup_{k,T\in\mathbb{Z}} \sup_{\substack{A\in\sigma(\underline{X}_{t+k,T},\underline{X}_{t+k+1,T},\ldots) \\ B\in\sigma(\underline{X}_{k,T},\underline{X}_{k-1,T},\ldots)}} |P(A\cap B) - P(A)P(B)| \leq Ct^{-\alpha}, \qquad t>0 \tag{3.8}$$

where C is a constant and $\alpha > 0$.

(L2) *There exists a covariance function* $\{\boldsymbol{\kappa}(u,h)\}_h$ *and function* $\boldsymbol{\kappa}_2(h)$ *such that* $|\mathrm{cov}(\underline{X}_{t_1,T},\underline{X}_{t_2,T})-\boldsymbol{\kappa}(t_1-t_2,\frac{t_1}{T})|_1 \leq \frac{1}{T}\kappa_2(t_1-t_2)$. *We assume the function* $\{\boldsymbol{\kappa}(u,h)\}_h$ *satisfies the following conditions:* $\sum_h h^2 \cdot |\boldsymbol{\kappa}(u,h)|_1 < \infty$ *and* $\sup_{u\in[0,1]}\sum_h |\frac{\partial \boldsymbol{\kappa}(u,h)}{\partial u}| < \infty$, *where on the boundary* 0 *and* 1 *we consider the right and left derivative respective (this assumption can be relaxed to* $\boldsymbol{\kappa}(u,h)$ *being piecewise continuous, where within each piece the function has a bounded derivative). The function* $\boldsymbol{\kappa}_2(h)$ *satisfies* $\sum_h |\kappa_2(h)| < \infty$.

(L3) *For some* $s > \frac{4\alpha}{\alpha-6} > 0$ *with* α *such that* (3.8) *holds, we have* $\sup_{t,T}\|\underline{X}_{t,T}\|_s < \infty$.

(L4) *Let* $\mathbf{f}(u;\omega) = \sum_{r=-\infty}^{\infty}\mathrm{cov}(\underline{X}_0(u),\underline{X}_r(u))\exp(-ir\omega)$. *Then the integrated spectral density matrix* $\mathbf{f}(\omega) = \int_0^1 \mathbf{f}(u,\omega)du$ *is non-singular on* $[0,2\pi]$.

Note that (L2) implies that $\sup_u |\frac{\partial \mathbf{f}(u;\omega)}{\partial u}| < \infty$.

(L5) *For some* $s > \frac{8\alpha}{\alpha-7} > 0$ *with* α *such that* (3.8) *holds, we have* $\sup_{t,T}\|\underline{X}_{t,T}\|_s < \infty$.

As in the stationary case, it can be shown that several nonlinear time series satisfy Assumption 3.2 (L1) (cf. Fryzlewicz and Subba Rao (2011) and Vogt (2012) who derive sufficient conditions for α-mixing of a general class of nonstationary time series). Assumption 3.2(L2) is used to show that the covariance changes slowly over time (these assumptions are used in order to derive the limit of the DFT covariance under local stationarity). The stronger Assumption (L5) is required to replace $\widehat{\mathbf{L}}(\omega)$ with its deterministic limit (see below for the limit).

3.3.2 Sampling properties of $\widehat{\mathbf{C}}_T(r,\ell)$ under local stationarity

As in the stationary case, it is difficult to directly analyze $\widehat{\mathbf{C}}_T(r,\ell)$. Therefore, we show that it can be replaced by $\widetilde{\mathbf{C}}_T(r,\ell)$ (defined in (3.1)), where in the locally stationary case $\mathbf{L}(\omega)$ are lower-triangular Cholesky matrices which satisfy $\overline{\mathbf{L}(\omega)}'\mathbf{L}(\omega) = \mathbf{f}^{-1}(\omega)$ and $\mathbf{f}(\omega) = \int_0^1 \mathbf{f}(u;\omega)du$.

Theorem 3.5 (Asymptotic equivalence of $\widehat{\mathbf{C}}_T(r,\ell)$ and $\widetilde{\mathbf{C}}_T(r,\ell)$ under local stationarity) *Suppose Assumption 3.2 is satisfied and let* $\widehat{\mathbf{C}}_T(r,\ell)$ *and* $\widetilde{\mathbf{C}}_T(r,\ell)$ *be defined as in* (2.5) *and* (3.1), *respectively. Then we have*

$$\sqrt{T}\widehat{\mathbf{C}}_T(r,\ell) = \sqrt{T}\Big(\widetilde{\mathbf{C}}_T(r,\ell) + \mathbf{S}_T(r,\ell) + \mathbf{B}_T(r,\ell)\Big) + O_P\left(\frac{\log T}{b\sqrt{T}} + b\log T + b^2\sqrt{T}\right) \quad (3.9)$$

and

$$\widehat{\mathbf{C}}_T(r,\ell) = \mathrm{E}(\widetilde{\mathbf{C}}_T(r,\ell)) + o_p(1),$$

where $\mathbf{B}_T(r,\ell) = O(b)$ *and* $\mathbf{S}_T(r,\ell)$ *are a deterministic bias and stochastic term, respectively, which are defined in Appendix A.2, equation (A.7).*

Remark 3.2 *There are some subtle differences between Theorems 3.1 and 3.5. In particular, the inclusion of the additional terms* $\mathbf{B}_T(r,\ell)$ *and* $\mathbf{S}_T(r,\ell)$*. We give a rough justification for this difference in the univariate case. Taking differences, it can be shown that*

$$\begin{aligned}\widehat{\mathbf{C}}_T(r,\ell)-\widetilde{\mathbf{C}}_T(r,\ell) &\approx \frac{1}{T}\sum_{k=1}^{T}\mathrm{E}(J_T(\omega_k)\overline{J_T(\omega_{k+r})})(\hat{g}_k-g_k)\\ &= \underbrace{\frac{1}{T}\sum_{k=1}^{T}\mathrm{E}(J_T(\omega_k)\overline{J_T(\omega_{k+r})})(\hat{g}_k-\mathrm{E}(\hat{g}_k))}_{\mathbf{B}_T(r,\ell)}+\underbrace{\frac{1}{T}\sum_{k=1}^{T}\mathrm{E}(J_T(\omega_k)\overline{J_T(\omega_{k+r})})(\mathrm{E}(\hat{g}_k)-g_k)}_{\mathbf{S}_T(r,\ell)},\end{aligned}$$

where g_k *is a function of the spectral density (see Appendix A.2 for details). In the case of second order stationarity, since* $\mathrm{E}(J_T(\omega_k)\overline{J_T(\omega_{k+r})})=O(T^{-1})$ *(for* $r\neq 0$*), the above terms are negligible, whereas in the case that the time series is nonstationary,* $\mathrm{E}(J_T(\omega_k)\overline{J_T(\omega_{k+r})})$ *is no longer negligible. In the nonstationary univariate case, the* $\mathbf{S}_T(r,\ell)$ *and* $\mathbf{B}_T(r,\ell)$ *become*

$$\begin{aligned}\mathbf{S}_T(r,\ell) &= \frac{-1}{2T}\sum_{t,\tau}\lambda_b(t-\tau)(X_tX_\tau-\mathrm{E}(X_tX_\tau)\\ &\quad\times\frac{1}{T}\sum_{k=1}^{T}h(\omega_k;r)e^{i\ell\omega_k}\left(\frac{\exp(i(t-\tau)\omega_k)}{\sqrt{f(\omega_k)^3f(\omega_{k+r})}}+\frac{\exp(i(t-\tau)\omega_{k+r})}{\sqrt{f(\omega_k)f(\omega_{k+r})^3}}\right)+O(\frac{1}{T})\\ \mathbf{B}_T(r,\ell) &= \frac{-1}{2T}\sum_{k=1}^{T}h(\omega_k,r)\begin{pmatrix}\mathrm{E}[\hat{f}_T(\omega_k)]-f(\omega_k)\\ \mathrm{E}[\hat{f}_T(\omega_{k+r})]-f(\omega_{k+r})\end{pmatrix}'\underline{A}(\omega_k,\omega_r),\end{aligned}$$

where

$$\underline{A}(\omega_k,\omega_r) = \begin{pmatrix}\frac{1}{(f(\omega_k)^3f(\omega_{k+r}))^{1/2}}\\ \frac{1}{((f(\omega_k)f(\omega_{k+r})^3)^{1/2}}\end{pmatrix}$$

and $h(\omega;r)=\int_0^1 f(u;\omega)\exp(2\pi iur)du$ *(see Lemma A.7 for details). A careful analysis will show that* $\mathbf{S}_T(r,\ell)$ *and* $\widetilde{\mathbf{C}}_T(r,\ell)$ *are both quadratic forms of the same order, this allows us to show asymptotic normality of* $\widehat{\mathbf{C}}_T(r,\ell)$ *under local stationarity.*

Lemma 3.1 *Suppose Assumption 3.2 is satisfied. Then for all* $r\in\mathbb{Z}$ *and* $\ell\in\mathbb{Z}$ *we have*

$$\mathrm{E}\big(\widetilde{\mathbf{C}}_T(r,\ell)\big)\to\mathbf{A}(r,\ell),\quad \text{and}\quad \widehat{\mathbf{C}}_T(r,\ell)\stackrel{P}{\to}\mathbf{A}(r,\ell)$$

as $T\to\infty$ *where*

$$\mathbf{A}(r,\ell)=\frac{1}{2\pi}\int_0^{2\pi}\int_0^1\mathbf{L}(\omega)\mathbf{f}(u;\omega)\overline{\mathbf{L}(\omega)}'\exp(i2\pi ru)\exp(i\ell\omega)dud\omega. \tag{3.10}$$

Since $\widehat{\mathbf{C}}_T(r,\ell)$ is an estimator of $\mathbf{A}(r,\ell)$, we now discuss how to interpret this.

Lemma 3.2 *Let* $\mathbf{A}(r,\ell)$ *be defined as in (3.10). Then under Assumption 3.2(L2) and (L4) we have that*

(i) $\mathbf{L}(\omega)\mathbf{f}(u,\omega)\overline{\mathbf{L}(\omega)}'$ *satisfies the representation*

$$\mathbf{L}(\omega)\mathbf{f}(u,\omega)\overline{\mathbf{L}(\omega)}' = \sum_{r,\ell\in\mathbb{Z}} \mathbf{A}(r,\ell)\exp(-i2\pi ru)\exp(-i\ell\omega).$$

and, consequently, $\mathbf{f}(u,\omega) = \mathbf{B}(\omega)\Big(\sum_{r,\ell\in\mathbb{Z}} \mathbf{A}(r,\ell)\exp(-i2\pi ru)\exp(-i\ell\omega)\Big)\overline{\mathbf{B}(\omega)}'$.

(ii) $\mathbf{A}(r,\ell)$ *is zero for all* $r \neq 0$ *and* $\ell \in \mathbb{Z}$ *iff* $\{X_t\}$ *is second order stationary.*

(iii) For all $\ell \neq 0$ *and* $r \neq 0$, $|\mathbf{A}(r,\ell)|_1 \leq K|r|^{-1}|\ell|^{-2}$ *(for some finite constant* K*).*

(iv) $\overline{\mathbf{A}(r,\ell)} = \mathbf{A}(-r,\ell)'$.

We see from part (ii) of the the above lemma that for $r \neq 0$, the coefficients $\{\mathbf{A}(r,\ell)\}$ characterize the nonstationarity. One consequence of Lemma 3.2 is that only for second order stationary time series, we do find that

$$\sum_{r=1}^{m}\sum_{\ell=0}^{n-1}\left(|\mathbf{S}_{r,\ell}\text{vech}(\Re\mathbf{A}(r,\ell))|_2^2 + |\mathbf{S}_{r,\ell}\text{vech}(\Im\mathbf{A}(r,\ell))|_2^2\right) = 0 \tag{3.11}$$

for any non-singular matrices $\{\mathbf{S}_{r,\ell}\}$ and all $n,m \in \mathbb{N}$. Therefore, under the alternative of local stationarity, the purpose of the test statistic is to detect the coefficients $A(r,\ell)$. Lemma 3.2 highlights another crucial point, that is, under local stationarity the absolute value of $|\mathbf{A}(r,\ell)|_1$ decays at the rate $C|r|^{-1}|\ell|^{-2}|$. Thus, the test will loose power if a large number of lags are used.

Theorem 3.6 (Limiting distributions of vech($\widehat{\mathbf{K}}_n(r)$)) *Let us assume that Assumption 3.2 holds and let* $\widehat{\mathbf{K}}_n(r)$ *be defined as in (2.7). Then, for fixed* $m,n \in \mathbb{N}$*, we have*

$$\sqrt{T}\begin{pmatrix} \Re\widehat{\mathbf{K}}_n(1) - \Re\mathbf{A}_n(1) - \Re\mathbf{B}_n(1) \\ \Im\widehat{\mathbf{K}}_n(1) - \Im\mathbf{A}_n(1) - \Im\mathbf{B}_n(1) \\ \vdots \\ \Re\widehat{\mathbf{K}}_n(m) - \Re\mathbf{A}_n(m) - \Re\mathbf{B}_n(m) \\ \Im\widehat{\mathbf{K}}_n(m) - \Im\mathbf{A}_n(m) - \Im\mathbf{B}_n(m) \end{pmatrix} \overset{D}{\to} \mathcal{N}(\underline{0}_{mnd(d+1)}, \widetilde{\mathbf{W}}_{m,n}),$$

where $\widetilde{\mathbf{W}}_{m,n}$ *is an* $(mnd(d+1) \times mnd(d+1))$ *variance matrix (which is not necessarily block diagonal),* $\mathbf{A}_n(r) = (vech(\mathbf{A}(r,0))', \ldots, vech(\mathbf{A}(r,n-1))')'$ *are the vectorized Fourier coefficients and* $\mathbf{B}_n(r) = (vech(\mathbf{B}(r,0))', \ldots, vech(\mathbf{B}(r,n-1))')' = O(b)$.

4 Properties of the stationary bootstrap applied to stationary and nonstationary time series

In this section, we consider the moments and cumulants of observations sampled using the stationarity bootstrap and its corresponding discrete Fourier transform. We use these results to analyze the bootstrap procedure proposed in Section 2.4. In order to reduce unnecessary notation, we state the results in this section for the univariate case only (all these results easily generalize to the multivariate case). The results in this section may also be of independent interest as they compare the differing characteristics of the stationary bootstrap when the underlying process is stationary and nonstationary. For this reason, this section is self-contained, where the main assumptions are mixing and moment conditions. The justification for the use of these mixing and moment conditions can be found in the proof of Lemma 4.1, which can be found in Appendix A.5.

We start by defining the ordinary and the cyclical sample covariances

$$\widehat{\kappa}(h) = \frac{1}{T}\sum_{t=1}^{T-|r|} X_t X_{t+h} - (\bar{X})^2 \qquad \widehat{\kappa}^C(h) = \frac{1}{T}\sum_{t=1}^{T} Y_t Y_{t+h} - (\bar{X})^2,$$

where $Y_t = X_{(t-1)\text{mod } T+1}$ and $\bar{X} = \frac{1}{T}\sum_{t=1}^{T} X_t$. We will also consider the higher order cumulants. Therefore we define the sample moments

$$\widehat{\mu}_n(h_1,\ldots,h_{n-1}) = \frac{1}{T}\sum_{t=1}^{T-\max|h_i|} X_t \prod_{i=1}^{n-1} X_{t+h_i}, \qquad \widehat{\mu}_n^C(h_1,\ldots,h_{n-1}) = \frac{1}{T}\sum_{t=1}^{T} Y_t \prod_{i=1}^{n-1} Y_{t+h_i} \quad (4.1)$$

(we set $r_1 = 0$) and the nth order cumulants corresponding to these moments

$$\begin{aligned} \widehat{\kappa}_n^C(h_1,\ldots,h_{n-1}) &= \sum_{\pi}(|\pi|-1)!(-1)^{|\pi|-1}\prod_{B\in\pi}\widehat{\mu}_n^C(\pi_{i\in B}), \qquad (4.2)\\ \widehat{\kappa}_n(h_1,\ldots,h_{n-1}) &= \sum_{\pi}(|\pi|-1)!(-1)^{|\pi|-1}\prod_{B\in\pi}\widehat{\mu}_n(\pi_{i\in B}), \end{aligned}$$

where π runs through all partitions of $\{0, h_1, \ldots, h_{n-1}\}$ and B are all blocks of the partition π. In order to obtain an expression for the cumulant of the DFT, we require the following lemma. We note that E^*, cov^*, cum^* and P^* denote the expectation, covariance, cumulant and probability with respect to the stationary bootstrap measure defined in Step 2 of Section 2.4.

Lemma 4.1 *Let $\{X_t\}$ be a time series with constant mean. We define the following expected quantities*

$$\widetilde{\kappa}_n(h_1,\ldots,h_{n-1}) = \sum_{\pi}(|\pi|-1)!(-1)^{|\pi|-1}\prod_{B\in\pi}\mathrm{E}[\widehat{\mu}_{|B|}(\pi_{i\in B})], \qquad (4.3)$$

where π runs through all partitions of $\{0, h_1 \ldots, h_{n-1}\}$, $\widehat{\mu}_n(h_1, \ldots, h_{n-1})$ is defined in (4.1),

$$\overline{\kappa}_n(h_1, \ldots, h_{n-1}) = \sum_{\pi}(|\pi|-1)!(-1)^{|\pi|-1} \prod_{B\in\pi}\left(\frac{1}{T}\sum_{t=1}^{T} \underbrace{\mathrm{E}[\mu_{|B|}(\pi_{i\in B+t})]}_{=\mathrm{E}(X_{t+i_1}X_{t+i_2}\ldots X_{t+i_B})}\right), \tag{4.4}$$

$\pi_{i\in B+t} = (i_1+t, \ldots, i_B+t)$ and $i_1, \ldots, i_B$ is a subset of $\{0, h_1, \ldots, h_{n-1}\}$.

(i) Suppose that $0 \le t_2 \le t_3 \ldots \le t_n$, then

$$\mathrm{cum}^*(X_0^*, \ldots, X_{t_n}^*) = (1-p)^{\max(t_n,0)-\min(t_1,0)}\widehat{\kappa}_n^C(t_2, \ldots, t_n).$$

To prove the assertions (ii-iv) below, we require the additional assumption that the time series $\{X_t\}$ is α-mixing, where for a given $q \ge 2n$ we have $\alpha > q$ and for some $s > q\alpha/(\alpha - q/n)$ we have $\sup_t \|X_t\|_s < \infty$. Note that this is a technical assumption that is used to give the following moment bounds, the exact details for their use can be found in the proof.

(ii) Approximation of circulant cumulant $\widehat{\kappa}^C$ by regular sample cumulant

$$\left\|\widehat{\kappa}_n^C(h_1, \ldots, h_{n-1}) - \widehat{\kappa}_n(h_1, \ldots, h_{n-1})\right\|_{q/n} \le C\frac{|h_{n-1}|}{T}\sup_{t,T}\|X_{t,T}\|_q^n, \tag{4.5}$$

where C is a finite constant which only depends on the order of the cumulant.

(iii) Approximation of sample regular cumulant by ' the cumulant of averages':

$$\|\widehat{\kappa}_n(h_1, \ldots, h_{n-1}) - \widetilde{\kappa}_n(h_1, \ldots, h_{n-1})\|_{q/n} = O(T^{-1/2}) \tag{4.6}$$

and

$$|\widetilde{\kappa}_n(h_1, \ldots, h_{n-1}) - \overline{\kappa}_n(h_1, \ldots, h_{n-1})| \le C\frac{\max(h_i,0)-\min(h_i,0)}{T}. \tag{4.7}$$

(iv) In the case of nth order stationarity, it holds $\overline{\kappa}_n = \mathrm{cum}(X_0, X_{t+h_1}, \ldots, X_{t+h_{n-1}})$.

However, if the time series is nonstationary, then

(a) $\overline{\kappa}_2(h) = \frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_t, X_{t+h})$

(b) $\overline{\kappa}_3(h_1, h_2) = \frac{1}{T}\sum_{t=1}^{T}\mathrm{cum}(X_t, X_{t+h_1}, X_{t+h_2})$

(c) The situation is different for the fourth order cumulant and we have

$$
\begin{aligned}
&\bar{\kappa}_4(h_1,h_2,h_3) = \frac{1}{T}\sum_{t=1}^{T}\mathrm{cum}(X_t, X_{t+h_1}, X_{t+h_2}, X_{t+h_3}) \\
&+\frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_t,X_{t+h_1})\mathrm{cov}(X_{t+h_2},X_{t+h_3}) - \left(\frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_t,X_{t+h_1})\right)\left(\frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_{t+h_2},X_{t+h_3})\right) \\
&+\frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_t,X_{t+h_2})\mathrm{cov}(X_{t+h_1},X_{t+h_3}) - \left(\frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_t,X_{t+h_2})\right)\left(\frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_{t+h_1},X_{t+h_3})\right) \\
&+\frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_t,X_{t+h_3})\mathrm{cov}(X_{t+h_1},X_{t+h_2}) - \left(\frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_t,X_{t+h_3})\right)\left(\frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_{t+h_1},X_{t+h_2})\right)
\end{aligned} \tag{4.8}
$$

(d) A similar expansion holds for $\bar{\kappa}_n(h_1,\ldots,h_{n-1})$ $(n > 4)$, ie. $\bar{\kappa}_n(\cdot)$ can be written as the average nth order cumulants plus additional lower order average cumulants terms.

In the above lemma we have shown that for stationary time series, the bootstrap cumulant is an approximation of the corresponding cumulant of the time series, which is not surprising. However, in the nonstationary case the bootstrap cumulant behaves differently. Under the assumption that the mean of the nonstationary time series is constant, the bootstrap cumulant of both second and third orders are the averages of the corresponding local cumulants. In other words, the second and third order bootstrap cumulants of a nonstationary time series behave like a stationary cumulant, ie. there is a decay in the cumulant the further apart the time lag. However, the bootstrap cumulants of higher orders (fourth and above) is not the average of the local cumulants, there are additional terms (see (4.8)). This means that the cumulants do not have the same decay as regular cumulants have. For example, from equation (4.8) we see that as the difference $|h_1 - h_2| \to \infty$, the function $\bar{\kappa}_4(h_1,h_2,h_3)$ does not converge to zero, whereas $\mathrm{cum}(X_t, X_{t+h_1}, X_{t+h_2}, X_{t+h_3})$ does (see Lemma A.9, in the Appendix).

We use Lemma 4.1 to derive results analogous to Brillinger (1981), Theorem 4.3.2, where an expression for the cumulant of DFTs in terms of the higher order spectral densities was derived. However, to prove this result we first need to derive the limit of the Fourier transform of the cumulant estimators. We define the sample higher order spectral density function as

$$
\begin{aligned}
&\widehat{h}_n(\omega_{k_1},\ldots,\omega_{k_{n-1}}) \\
= \; &\frac{1}{(2\pi)^{n-1}}\sum_{h_1,\ldots,h_{n-1}=-T}^{T}(1-p)^{\max(h_i,0)-\min(h_i,0)}\widehat{\kappa}_n(h_1,\ldots,h_{n-1})e^{-ih_1\omega_{k_1}-\ldots-ih_{n-1}\omega_{k_{n-1}}},
\end{aligned} \tag{4.9}
$$

where $\widehat{\kappa}_n(\cdot)$ are the sample cumulants defined in (4.3). In the following lemma, we show that $\widehat{h}_n(\cdot)$ approximates the 'pseudo' higher order spectral density

$$\begin{aligned} & f_{n,T}(\omega_{k_1},\ldots,\omega_{k_{n-1}}) \qquad (4.10) \\ = & \frac{1}{(2\pi)^{n-1}} \sum_{h_1,\ldots,h_{n-1}=-T}^{T} (1-p)^{\max(h_i,0)-\min(h_i,0)} \overline{\kappa}_n(h_1,\ldots,h_{n-1}) e^{-ih_1\omega_{k_1}-ih_2\omega_{k_2}-\ldots-ih_{n-1}\omega_{k_{n-1}}}, \end{aligned}$$

where $\overline{\kappa}_n(\cdot)$ is defined in (4.4)

We now show that under certain conditions $\widehat{h}_n(\cdot)$ is an estimator of the higher spectral density function.

Lemma 4.2 *Suppose the time series $\{X_t\}$ (where $\mathrm{E}(X_t) = \mu$ for all t) is α-mixing and $\sup_t \|X_t\|_s < \infty$ where $\alpha > q$ and $s > q\alpha/(\alpha - q/n)$.*

(i) Let $\widehat{h}_n(\cdot)$ and $\overline{f}_n(\cdot)$ be defined in (4.9) and (4.10), respectively. Then we have

$$\left\| \widehat{h}_n(\omega_1,\ldots,\omega_{n-1}) - f_{n,T}(\omega_1,\ldots,\omega_{n-1}) \right\|_{q/n} = O\left(\frac{1}{Tp^n} + \frac{1}{T^{1/2}p^{(n-1)}} + p \right), \qquad (4.11)$$

(ii) If the time series is nth order stationary, then we have

$$\left\| \widehat{h}_n(\omega_{k_1},\ldots,\omega_{k_{n-1}}) - f_n(\omega_{k_1},\ldots,\omega_{k_{n-1}}) \right\|_{q/n} = O\left(\frac{1}{Tp^n} + \frac{1}{T^{1/2}p^{(n-1)}} + p \right) \qquad (4.12)$$

and $\sup_{\omega_1,\ldots,\omega_{n-1}} |f_n(\omega_1,\ldots,\omega_{n-1})| < \infty$, where f_n is the nth order spectral density function defined as

$$f_n(\omega_{k_1},\ldots,\omega_{k_{n-1}}) = \frac{1}{(2\pi)^{n-1}} \sum_{h_1,\ldots,h_{n-1}=-\infty}^{\infty} \kappa_n(h_1,\ldots,h_{n-1}) e^{-ih_1\omega_{k_1}-\ldots-ih_{n-1}\omega_{k_{n-1}}}$$

and $\kappa_n(h_1,\ldots,h_{n-1})$ denotes the nth order joint cumulant of the stationary time series $\{X_t\}$.

(iii) On the other hand, if the time series is nonstationary:

(a) For $n \in \{2,3\}$, we have

$$\begin{aligned} \left\| \widehat{h}_2(\omega_{k_1}) - f_{2,T}(\omega_{k_1}) \right\|_{q/n} &= O\left(\frac{1}{Tp^2} + \frac{1}{T^{1/2}p} + p \right), \\ \left\| \widehat{h}_3(\omega_{k_1},\omega_{k_2}) - f_{3,T}(\omega_{k_1},\omega_{k_2}) \right\|_{q/n} &= O\left(\frac{1}{Tp^3} + \frac{1}{T^{1/2}p^2} + p \right), \\ \text{where } f_{2,T}(\omega) &= \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \bar{\kappa}_2(h) \exp(ih\omega) \end{aligned}$$

with $\bar{\kappa}_2(h)$ defined as in Lemma 4.1(iva) and $f_{3,T}$ is defined similarly. Since the average covariances and cumulants are absolutely summable, we have $\sup_{T,\omega} |f_{2,T}(\omega)| < \infty$ and $\sup_{T,\omega_1,\omega_2} |f_{3,T}(\omega_1,\omega_2)| < \infty$.

(b) *For* $n = 4$, *we have* $\sup_{\omega_1,\omega_2,\omega_3} |f_{4,T}(\omega_1,\omega_2,\omega_3)| = O(p^{-1})$.

(c) *For* $n \geq 4$, *we have* $\sup_{\omega_1,\ldots,\omega_{n-1}} |f_{n,T}(\omega_1,\ldots,\omega_{n-1})| = O(p^{-(n-3)})$.

The following result is the bootstrap analogue of (Brillinger, 1981), Theorem 4.3.2.

Theorem 4.1 *Let* $J_T^*(\omega)$ *denote the DFT of the stationary bootstrap observations. Under the assumptions that* $\|X_t\|_n < \infty$, *we have*

$$\left\|\mathrm{cum}^*(J_T^*(\omega_{k_1}),\ldots,J_T^*(\omega_{k_n}))\right\|_1 = O\left(\frac{1}{T^{n/2-1}p^{n-1}}\right). \tag{4.13}$$

By imposing the additional condition that $\{X_t\}$ *is an* α*-mixing time series with a constant mean,* $q/n \geq 2$, *the mixing rate* $\alpha > q$ *and* $\|X_t\|_s < \infty$ *for some* $s > q\alpha/(\alpha - q/n)$, *we obtain*

$$\begin{aligned}
&\mathrm{cum}^*(J_T^*(\omega_{k_1}),\ldots,J_T^*(\omega_{k_n})) \\
= &\frac{(2\pi)^{n/2-1}}{T^{n/2-1}}\widehat{h}_n(\omega_{k_2},\ldots,\omega_{k_n})\frac{1}{T}\sum_{t=1}^{T}\exp(-it(\omega_{k_1}+\ldots+\omega_{k_n})) + R_{T,n}^{(1)},
\end{aligned} \tag{4.14}$$

where $\|R_{T,n}^{(1)}\|_{q/n} = O(\frac{1}{T^{n/2}p^n})$.

(a) *If* $\{X_t\}_t$ *is nth order stationary then*

$$\begin{aligned}
&\left\|\mathrm{cum}^*(J_T^*(\omega_{k_1}),\ldots,J_T^*(\omega_{k_n}))\right\|_{q/n} \\
= &\frac{(2\pi)^{n/2-1}}{T^{n/2-1}}f_n(\omega_{k_2},\ldots,\omega_{k_n})\frac{1}{T}\sum_{t=1}^{T}\exp(-it(\omega_{k_1}+\ldots+\omega_{k_n})) + R_{T,n}^{(2)} \\
= &\begin{cases} O\left(\frac{1}{T^{n/2-1}} + \frac{1}{(T^{1/2}p)^{n-1}}\right), & \sum_{l=1}^{n}\omega_{k_l} \in \mathbb{Z} \\ O\left(\frac{1}{(T^{1/2}p)^n}\right), & \sum_{l=1}^{n}\omega_{k_l} \notin \mathbb{Z} \end{cases},
\end{aligned} \tag{4.15}$$

which is uniform over $\{\omega_k\}$ *and* $\|R_{T,n}^{(2)}\|_{q/n} = O(\frac{1}{(T^{1/2}p)^{1/2}})$.

(b) *If* $\{X_t\}$ *is nonstationary (with constant mean) then for* $n \in \{2,3\}$, *we replace* f_n *with* $f_{2,T}$ *and* $f_{3,T}$, *respectively, and obtain the same as above.*

For $n \geq 4$, *we have*

$$\left\|\mathrm{cum}^*(J_T^*(\omega_{k_1}),\ldots,J_T^*(\omega_{k_n}))\right\|_{q/n} = \begin{cases} O\left(\frac{1}{T^{n/2-1}p^{n-3}} + \frac{1}{(T^{1/2}p)^{n-1}}\right), & \sum_{l=1}^{n}\omega_{k_l} \in \mathbb{Z} \\ O\left(\frac{1}{(T^{1/2}p)^n}\right), & \sum_{l=1}^{n}\omega_{k_l} \notin \mathbb{Z} \end{cases} \tag{4.16}$$

An immediately implication of the theorem above is that the bootstrap variance of the DFT can be used as an estimator of spectral density function.

5 Analysis of the test statistic

In Section 3.1, we derived the properties of the DFT covariance in the case of stationarity. These results show that the distribution of the test statistic, in the unlikely event that $\mathbf{W}^{(2)}$ is known, is a chi-square (see Theorem 3.4). In the case that $\mathbf{W}^{(2)}$ is unknown as in Section 2.4 we proposed a method to estimate $\mathbf{W}^{(2)}$ and thus the bootstrap statistic. In this section we show that under fourth order stationarity of the time series, the bootstrap variance defined in Step 6 of the algorithm is a consistent estimator of $\mathbf{W}^{(2)}$. Thus, the bootstrap statistic $T^{*}_{m,n,d}$ asymptotically converges to a chi-squared distribution. We also investigate the power of the test under the alternative of local stationarity. To derive the power, we use the results in Section 3.3, where we show that for at least some values of r and ℓ (usually the low orders), $\widehat{\mathbf{C}}_T(r,\ell)$ has a non-centralized normal distribution. However, the test statistic also involves $\mathbf{W}^{(2)}$, which is estimated as if the underlying time series is stationary (using the stationary bootstrap procedure). Therefore, in this section, we derive an expression for the quantity that $\mathbf{W}^{(2)}$ is estimating under the assumption of second order nonstationarity, and explain how this influences the power of the test.

We use the following assumption in Lemma 5.1, where we show that the variance of the bootstrap cumulants converge to zero as the sample size increases.

Assumption 5.1 *Suppose that* $\{\underline{X}_t\}$ *is* α*-mixing with* $\alpha > 8$ *and the moments satisfy* $\|\underline{X}\|_s < \infty$, *where* $s > 8\alpha/(\alpha-2)$.

Lemma 5.1 *Suppose that the time series* $\{\underline{X}_t\}$ *satisfies Assumption 5.1.*

(i) If $\{\underline{X}_t\}$ *is fourth order stationary, then we have*

(a) $\mathrm{cum}^{*}(J^{*}_{T,j_1}(\omega_{k_1}), J^{*}_{T,j_2}(\omega_{k_2})) = f_{j_1,j_2}(\omega_{k_1}) I(k_1 = -k_2) + R_1$.

(b) $\mathrm{cum}^{*}(J^{*}_{T,j_1}(\omega_{k_1}), J^{*}_{T,j_2}(\omega_{k_2}), J^{*}_{T,j_3}(\omega_{k_3}), J^{*}_{T,j_4}(\omega_{k_4})) = \frac{(2\pi)}{T} f_{4;j_1,\ldots,j_4}(\omega_{k_1},\omega_{k_2},\omega_{k_3}) I(k_4 = -k_1 - k_2 - k_3) + R_2$,

where $\|R_1\|_4 = O(\frac{1}{Tp^2})$ *and* $\|R_2\|_2 = O(\frac{1}{T^2p^4})$

(ii) If $\{\underline{X}_t\}$ *has a constant mean, then we have*

(a) $\mathrm{cum}^{*}(J^{*}_{T,j_1}(\omega_{k_1}), J^{*}_{T,j_2}(\omega_{k_2})) = f_{2,T;j_1,j_2}(\omega_{k_1}) I(k_1 = -k_2) + R_1^{*}$

(b) $\mathrm{cum}^{*}(J^{*}_{T,j_1}(\omega_{k_1}), J^{*}_{T,j_2}(\omega_{k_2}), J^{*}_{T,j_3}(\omega_{k_3}), J^{*}_{T,j_4}(\omega_{k_4})) = \frac{(2\pi)}{T} f_{4,T;j_1,\ldots,j_4}(\omega_{k_1},\omega_{k_2},\omega_{k_3}) I(k_4 = -k_1 - k_2 - k_3) + R_2^{*}$,

where $\|R_1^*\|_4 = O(\frac{1}{Tp^2})$ *and* $\|R_2^*\|_2 = O(\frac{1}{T^2p^4})$.

In order to obtain the limit of the bootstrap variance estimator, we define

$$\widetilde{\mathbf{C}}_T^*(r,\ell) = \frac{1}{T}\sum_{k=1}^{T} \mathbf{L}(\omega_k)\underline{J}_T^*(\omega_k)\overline{\underline{J}_T^*(\omega_{k+r})}'\overline{\mathbf{L}(\omega_{k+r})}'\exp(i\ell\omega_k).$$

We observe that this is almost identical to the bootstrap DFT $\widehat{\mathbf{C}}_T^*(r,\ell)$ and $\acute{\mathbf{C}}_T^*(r,\ell)$, except that $\widehat{\mathbf{L}}^*(\cdot)$ and $\widehat{\mathbf{L}}(\cdot)$ have been replaced with their limit $\mathbf{L}(\cdot)$. We first obtain the variance of $\widetilde{\mathbf{C}}_T^*(r,\ell)$, which is simply a consequence of Lemma 5.1. Later, we show that it is equivalent to the bootstrap variance of $\widehat{\mathbf{C}}_T^*(r,\ell)$ and $\acute{\mathbf{C}}_T^*(r,\ell)$.

Theorem 5.1 (Consistency of the variance estimator based on $\widetilde{\mathbf{C}}_T^*(r,\ell)$) *Suppose that* $\{\underline{X}_t\}$ *is an* α*-mixing time series which satisfies Assumption 5.1 and let*

$$\widetilde{\mathbf{K}}_n^*(r) = \Big(vech(\widetilde{\mathbf{C}}_T^*(r,0))', vech(\widetilde{\mathbf{C}}_T^*(r,1))', \ldots, vech(\widetilde{\mathbf{C}}_T^*(r,n-1))'\Big)'.$$

Suppose $Tp^4 \to \infty$, $bTp^2 \to \infty$, $b \to 0$ *and* $p \to 0$ *as* $T \to \infty$,

(i) In addition suppose that $\{\underline{X}_t\}$ *is a fourth order stationary time series. Let* $\mathbf{W}_n$ *be defined as in (2.12).*

Then for fixed $m, n \in \mathbb{N}$ *we have* $T\mathrm{var}^*(\Re\widetilde{\mathbf{K}}_n(r)) = \mathbf{W}_n + o_p(1)$ *and* $T\mathrm{var}^*(\Im\widetilde{\mathbf{K}}_n(r)) = \mathbf{W}_n + o_p(1)$.

(ii) On the other hand, suppose $\{\underline{X}_t\}$ *is a locally stationary time series which satisfies Assumption 3.2(L2). Let*

$$\begin{aligned}
&\kappa_T^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)\\
=\;& \sum_{k_1,k_2=1}^{T}\sum_{s_1,s_2,s_3,s_4=1}^{d} L_{j_1s_1}(\lambda_{k_1})\overline{L_{j_2s_2}(\lambda_{k_1})L_{j_3s_3}(\lambda_{k_2})}L_{j_4s_4}(\lambda_{k_2})\exp(i\ell_1\lambda_{k_1} - i\ell_2\lambda_{k_2})\\
&\times f_{4,T;s_1,s_2,s_3,s_4}(\lambda_{k_1}, -\lambda_{k_1}, -\lambda_{k_2})d\lambda_1 d\lambda_2, \qquad (5.1)
\end{aligned}$$

where $\mathbf{L}(\omega)\overline{\mathbf{L}(\omega)}' = \mathbf{f}(\omega)^{-1}$, $\mathbf{f}(\omega) = \int_0^1 \mathbf{f}(u;\omega)du$ *and*

$$\begin{aligned}
&f_{4,T;s_1,s_2,s_3,s_4}(\lambda_{k_1}, -\lambda_{k_1}, -\lambda_{k_2})\\
=\;& \frac{1}{(2\pi)^3}\sum_{h_1,h_2,h_3=-T}^{T}(1-p)^{\max(h_i,0)-\min(h_i,0)}\overline{\kappa}_{4,s_1,s_2,s_3,s_4}(h_1,h_2,h_3)e^{-ih_1\omega_{k_1}+ih_2\omega_{k_1}-ih_3\omega_{k_2}}.
\end{aligned}$$

Using the above we define

$$(\mathbf{W}_{T,n})_{\ell_1+1,\ell_1+1} = \mathbf{W}_{\ell_1}^{(1)}\delta_{\ell_1,\ell_2} + \mathbf{W}_{\ell_1,\ell_2}^{(2)}, \qquad (5.2)$$

where $\mathbf{W}_{\ell}^{(1)}$ *and* $\mathbf{W}_{\ell_1,\ell_2}^{(2)}$ *are defined as in (2.11) and (2.15) but with* $\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)$ *replaced with* $\kappa_T^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)$.

Then, for fixed $m,n \in \mathbb{N}$, *we have* $T\text{var}^*(\Re\widetilde{\mathbf{K}}_n(r)) = \mathbf{W}_{T,n} + o_p(1)$ *and* $T\text{var}^*(\Im\widetilde{\mathbf{K}}_n(r)) = \mathbf{W}_{T,n} + o_p(1)$. *Furthermore,* $|\kappa_T^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)| = O(p^{-1})$ *and* $|\mathbf{W}_{T,n}|_1 = O(p^{-1})$.

The above theorem shows that if $\widetilde{\mathbf{C}}_T^*(r,\ell)$ were known, then the bootstrap variance estimator is consistent under fourth order stationarity. Now we show that both the asymptotic bootstrap variance of $\widehat{\mathbf{C}}_T^*(r,\ell)$ and $\acute{\mathbf{C}}_T^*(r,\ell)$ are equivalent to the variance of $\widetilde{\mathbf{C}}_T^*(r,\ell)$.

Assumption 5.2 (Variance equivalence) *(B1) Let* $\bar{\mathbf{f}}_{\alpha,T}(\omega) = \alpha(\omega)\hat{\mathbf{f}}_T^*(\omega) + (1-\alpha(\omega))\hat{\mathbf{f}}_T(\omega)$, *where* $\alpha : [0,2\pi] \to [0,1]$ *and* $L_{j_1,j_2}(\bar{\mathbf{f}}_T(\omega))$ *denote the* (j_1,j_2)*th element of the matrix* $\mathbf{L}(\bar{\mathbf{f}}_T(\omega))$. *Let* $\nabla^i L_{j_1,j_2}(\mathbf{f}(\omega))$ *denote the* i*th derivative with respect to the vector* $\mathbf{f}(\omega)$. *We assume that for every* $\varepsilon > 0$ *there exists a* $0 < M_\varepsilon < \infty$ *such that*

$$P\left(\sup_{\alpha,\omega}(\mathrm{E}^*|\nabla^i L_{j_1,j_2}(\bar{\mathbf{f}}_{\alpha,T}(\omega))|^8)^{1/8} > M_\varepsilon\right) < \varepsilon,$$

for $i = 0,1,2$. *In other words the sequence* $\{\sup_{\omega,\alpha}\mathrm{E}^*|\nabla^i L_{j_1,j_2}(\bar{\mathbf{f}}_{\alpha,T}(\omega))|^8)^{1/8}\}_T$ *is bounded in probability.*

(B2) The time series $\{\underline{X}_t\}$ *is* α*-mixing with* $\alpha > 16$ *and has a finite* s*th moment* ($\sup_t \|X_t\|_s < \infty$) *such that* $s > 16\alpha/(\alpha-2)$).

Remark 5.1 (On Assumption 5.2(B1)) *(i) This is a technical assumption that is required when showing equivalence of the bootstrap variance estimator using* $\widehat{\mathbf{C}}_T^*(r,\ell)$ *to the bootstrap variance using* $\widetilde{\mathbf{C}}_T^*(r\ell)$. *In the case we use* $\acute{\mathbf{C}}_T^*(r,\ell)$ *to construct the bootstrap variance (defined in (2.25)) we do not require this assumption.*

(ii) Let ∇^2 *denote the second derivative with respect to the vector* $(\bar{\mathbf{f}}_{\alpha,T}(\omega_1), \bar{\mathbf{f}}_{\alpha,T}(\omega_2))$. *Assumption 5.2(B1) implies that the sequence* $\sup_{\omega_1,\omega_2,\alpha}\mathrm{E}^*|\nabla^2 L_{j_1,j_2}(\bar{\mathbf{f}}_{\alpha,T}(\omega_1))L_{j_1,j_2}(\bar{\mathbf{f}}_{\alpha,T}(\omega_2))|^4)^{1/4}$ *is bounded in probability. We use this result in the proof of Lemma A.16.*

(ii) In the case $d = 1$ $L(\omega) = f^{-1/2}(\omega)$ *and Assumption 5.2(B1) corresponds to the condition that for* $i = 0,1,2$ *the sequence* $\{\sup_{\alpha,\omega}\left[\mathrm{E}^*\left(\bar{f}_{\alpha,T}^*(\omega)^{-4(2i+1)}\right)\right]^{1/8}\}_T$ *is bounded in probability.*

Using the above results, we now derive a bound for the difference between the covariances $\widehat{\mathbf{C}}_T^*(r,\ell_1)$ and $\widetilde{\mathbf{C}}_T^*(r,\ell_2)$.

Lemma 5.2 *Suppose that $\{\underline{X}_t\}$ is a fourth order stationary time series or a constant mean locally stationary time series which satisfies Assumption 3.2(L2)), Assumption 5.2(B2) holds and $Tp^4 \to \infty$ $bTp^2 \to \infty$, $b \to 0$ and $p \to 0$ as $T \to \infty$. Then, we have*

(i)

$$|T \left(\mathrm{cov}^*(\Re \acute{\mathbf{C}}^*_T(r,\ell_1), \Re \acute{\mathbf{C}}^*_T(r,\ell_2)) - \mathrm{cov}^*(\Re \widetilde{\mathbf{C}}^*_T(r,\ell_1), \Re \widetilde{\mathbf{C}}^*_T(r,\ell_2)) \right) | = o_p(1).$$

and

$$|T \left(\mathrm{cov}^*(\Im \acute{\mathbf{C}}^*_T(r,\ell_1), \Im \acute{\mathbf{C}}^*_T(r,\ell_2)) - \mathrm{cov}^*(\Im \widetilde{\mathbf{C}}^*_T(r,\ell_1), \Im \widetilde{\mathbf{C}}^*_T(r,\ell_2)) \right) | = o_p(1)$$

(ii) If in addition Assumption 5.2(B1) holds, then we have

$$|T \left(\mathrm{cov}^*(\Re \widehat{\mathbf{C}}^*_T(r,\ell_1), \Re \widehat{\mathbf{C}}^*_T(r,\ell_2)) - \mathrm{cov}^*(\Re \widetilde{\mathbf{C}}^*_T(r,\ell_1), \Re \widetilde{\mathbf{C}}^*_T(r,\ell_2)) \right) | = o_p(1)$$

and

$$|T \left(\mathrm{cov}^*(\Im \widehat{\mathbf{C}}^*_T(r,\ell_1), \Im \widehat{\mathbf{C}}^*_T(r,\ell_2)) - \mathrm{cov}^*(\Im \widetilde{\mathbf{C}}^*_T(r,\ell_1), \Im \widetilde{\mathbf{C}}^*_T(r,\ell_2)) \right) | = o_p(1).$$

Finally, by using the above, we obtain the following result.

Theorem 5.2 *Suppose 5.2(B2) holds. Let the test statistic $\mathcal{T}^*_{m,n,d}$ be defined as in (2.24), where the bootstrap variance is constructed using either $\widehat{\mathbf{C}}^*_T(r,\ell)$ or $\acute{\mathbf{C}}^*_T(r,\ell)$ (if $\widehat{\mathbf{C}}^*_T(r,\ell)$ is used to construct the test statistic, then Assumption 5.2(B1) also holds).*

(i) Suppose Assumption 3.1 holds. Then we have

$$\mathcal{T}^*_{m,n,d} \stackrel{P}{\to} \chi^2_{mnd(d+1)}.$$

(ii) Suppose Assumption 3.2 and $A(r,\ell) \neq 0$ for some $0 < r \leq m$ and $0 \leq \ell \leq n$ hold, then we have

$$\mathcal{T}^*_{m,n,d} = O_p(Tp).$$

The above theorem shows that under fourth order stationarity the asymptotic distribution of $\mathcal{T}^*_{m,n,d}$ (where we use the bootstrap variance as an estimator of $\mathbf{W}_n$) is asymptotically equivalent to the test statistic as if $\mathbf{W}_n$ were known. We observe that the mean length of the bootstrap block $1/p$ does not play a role in the asymptotic distribution under stationarity. This is in sharp contrast to the locally stationarity case. If we did not use a bootstrap scheme to estimate $\mathbf{W}_n^{-1/2}$ (ie. we were to use $\mathbf{W}_n = \mathbf{W}_n^{(1)}$, which is the variance in the case of Gaussianity), then under local stationarity $\mathcal{T}_{m,n,d} = O_p(T)$. However, by using the bootstrap scheme we incur a slight loss in power since $\mathcal{T}^*_{m,n,d} = O_p(pT)$.

6 Practical Issues

In this section, we consider the implementation issues related to the test statistic. We will be considering both the test statistic $\mathcal{T}^*_{m,n,d}$, where we use the stationary bootstrap to estimate the variance, and compare it to the test statistic $\mathcal{T}_{m,n,d,G}$ (defined in (2.20)) that is constructed as if the observations are Gaussian.

6.1 Selection of the tuning parameters

We recall from the definition of the test statistic that there are four different tuning parameters that need to be selected in order to construct the test statistic, to recap these are b the bandwidth for spectral density matrix estimation, m the number of DFT covariances $\widehat{C}_T(r,\ell)$ (where $r = 1,\ldots,m$), n the number of DFT covariances $\widehat{C}_T(r,\ell)$ (where $\ell = 0,\ldots,n-1$) and p which determines the average block length (which is p^{-1}) in the bootstrap scheme. For the simulations below and the real data example, we use $n = 1$. This is because (a) in most situations it is likely that the nonstationarity is 'seen' in $\widehat{C}_T(r,0)$ and (b) we have shown that under the alternative of local stationarity $\widehat{C}_T(r,\ell) \stackrel{P}{\to} \mathbf{A}(r,\ell)$ and for $\ell \neq 0$ $\mathbf{A}(r,\ell)$ so using a large n can result in a loss of power. However, we do recommend that a plot of $\widehat{C}_T(r,\ell)$ (or a standardized $\widehat{C}_T(r,\ell)$) is made against r and ℓ (similar to Figures 1–3) to see if there are any large coefficients which may be statistically significant. We now discuss how to select b, m and p. These procedures will be used in the simulations below.

Choice of the bandwidth b

To estimate the spectral density matrix we need to select the bandwidth b. We use the cross-validation criterion, suggested in Beltrao and Bloomfield (1987) (see also Robinson (1991)).

Choice of the number of lags m

We select the m by adapting the data driven rule suggested by Escanciano and Lobato (2009) (who propose a method for selecting the number of lags in a Portmanteau test for testing uncorrelatedness of a time series). We summarize their procedure and then discuss how we use it to select m in our test for stationarity. For univariate time series $\{X_t\}$, Escanciano and Lobato (2009) suggest selecting the number of lags in a Portmanteau test using the criterion

$$\widetilde{m}_P = \min\{m : 1 \leq m \leq D : L_m \geq L_h, h = 1,2,\ldots,D\},$$

where $L_m = Q_m - \pi(m,T,q)$, $Q_m = T\sum_{j=1}^{m}|\widehat{R}(j)/\widehat{R}(0)|^2$, D is a fixed upper bound and $\pi(m,T,q)$ is a penalty term that takes the form

$$\pi(m,T,q) = \begin{cases} m\log(T), & \max_{1\le k\le D}\sqrt{T}|\widehat{R}(k)/\widehat{R}(0)| \le \sqrt{q\log(T)} \\ 2m, & \max_{1\le k\le D}\sqrt{T}|\widehat{R}(k)/\widehat{R}(0)| > \sqrt{q\log(T)} \end{cases},$$

where $\widehat{R}(k) = \frac{1}{T-k}\sum_{j=1}^{T-|k|}(X_j - \overline{X})(X_{j+|k|} - \overline{X})$. We now propose to adapt this rule to select m. More precisely, depending on whether we use $\mathcal{T}^*_{m,n,d}$ or $\mathcal{T}_{m,n,d;G}$ we define the sequences of bootstrap covariances $\{\widehat{\gamma}^*(r), r \in \mathbb{N}\}$ and non-bootstrap covariances $\{\widehat{\gamma}(r), r\in\mathbb{N}\}$, where

$$\widehat{\gamma}^*(r) = \frac{1}{d(d+1)}\sum_{j=1}^{d(d+1)/2}\left\{\widehat{\mathbf{S}}^*(r)\mathrm{vech}(\Re\widehat{\mathbf{K}}_n(r))_j + \widehat{\mathbf{S}}^*(r)\mathrm{vech}(\Im\widehat{\mathbf{K}}_n(r))_j\right\}$$

and $\widehat{\gamma}(r)$ is defined similarly with $\mathbf{S}^*(r)$ replaced by $\mathbf{W}_0^{(1)}$ as in (2.20). We select m by using

$$\widehat{m} = \min\{m : 1 \le m \le D : L_m \ge L_h, h = 1,2,\ldots,D\},$$

where $L_m = \mathcal{T}^*_{m,n,d} - \pi^*(m,T,q)$ (or $\mathcal{T}_{m,n,d;G} - \pi(m,T,q)$ if Gaussianity is assumed) and

$$\pi^*(m,T,q) = \begin{cases} m\log(T), & \max_{1\le r\le D}\sqrt{T}|\widehat{\gamma}^*(r)| \le \sqrt{q\log(T)} \\ 2m, & \max_{1\le r\le D}\sqrt{T}|\widehat{\gamma}^*(r)| > \sqrt{q\log(T)} \end{cases},$$

and $\pi(m,T,q)$ is defined similarly but using $\gamma(r)$ instead of $\gamma^*(r)$.

Choice of the average block size $1/p$

For the bootstrap test, the tuning parameter $1/p$ is chosen by adapting the rule suggested by Politis and White (2004) (and later corrected in Patton, Politis, and White (2009)) that was originally proposed in order to estimate the finite sample distribution of the univariate sample mean (using the stationary bootstrap). More precisely, to bootstrap the sample mean for dependent univariate time series $\{X_t\}$, they suggest to select the tuning parameter for the stationary bootstrap as

$$\frac{1}{p} = \left(\frac{\widehat{G}^2}{\widehat{g}^2(0)}\right)^{1/3} T^{1/3}, \tag{6.1}$$

where $\widehat{G} = \sum_{k=-M}^{M}\lambda(k/M)|k|\widehat{R}(k)$, $\widehat{g}(0) = \sum_{k=-M}^{M}\lambda(k/M)\widehat{R}(k)$, $\widehat{R}(k) = \frac{1}{T}\sum_{j=1}^{T-|k|}(Y_j - \overline{Y})(Y_{j+|k|} - \overline{Y})$ and

$$\lambda(t) = \begin{cases} 1, |t| \in [0,1/2] \\ 2(1-|t|), |t| \in [1/2,1] \\ 0, \text{ otherwise} \end{cases}$$

is a trapezoidal shape symmetric flat-top taper. We have to adapt the rule (6.1) in two ways for our purposes. First, the theory established in Section 5 requires $Tp^4 \to \infty$ for the stationary bootstrap to be consistent. Hence, we suggest to use the same (estimated) constant as in (6.1), but we multiply it with $T^{1/5}$ instead of $T^{1/3}$ to meet these requirements. Second, as (6.1) is tailor-made for univariate data, we propose to apply it separately to all components of multivariate data and to define $1/p$ as the average value. We mention that proper selection of a p (and in general the block length in any bootstrap procedure) is an extremely difficult problem and requires further investigation (see, for example, Paparoditis and Politis (2004) and Parker et al. (2006)).

6.2 Simulations

We now illustrate the performance of the test for stationarity of a multivariate time series through simulations. We will compare the test statistics $\mathcal{T}^*_{m,n,d}$ and $\mathcal{T}_{m,n,d;G}$, which are defined in (2.24) and (2.20), respectively. In the following, we refer to $\mathcal{T}^*_{m,n,d}$ and $\mathcal{T}_{m,n,d;G}$ as the bootstrap and the non-bootstrap test, respectively. Observe that the non-bootstrap test is asymptotically a test of level α only in the case that the fourth order cumulants are zero (which includes the Gaussian case). We reject the null of stationarity at the nominal level $\alpha \in (0, 1)$ if

$$\mathcal{T}^*_{m,n,d} > \chi^2_{mnd(d+1)}(1-\alpha) \quad \text{and} \qquad \mathcal{T}_{m,n,d;G} > \chi^2_{mnd(d+1)}(1-\alpha). \tag{6.2}$$

6.2.1 Simulation setup

In the simulations below, we consider several stationary and nonstationary bivariate ($d = 2$) time series models. For each model we have generated $M = 400$ replications of the bivariate time series ($\underline{X}_t = (X_{t,1}, X_{t,2})', t = 1, \ldots, T$) with sample size $T = 500$. As described above, the bandwidth b for estimating the spectral density matrices is chosen by cross-validation. To select m, we set $q = 2.4$ (as recommended in Escanciano and Lobato (2009)) and $D = 10$. To compute the quantities $\widehat{G}$ and $\widehat{g}(0)$ for the selection procedure of $1/p$ (see (6.1)), we set $M = 1/b$. Further, we have used $N = 400$ bootstrap replications for each time series.

6.2.2 Models under the null

To investigate the behavior of the tests under the null of (second order) stationarity of the process $\{\underline{X}_t\}$, we consider realizations from two vector autoregressive models (VAR), two GARCH-type models and one Markov switching model. Throughout this section, let

$$A = \begin{pmatrix} 0.6 & 0.2 \\ 0 & 0.3 \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} 1 & 0.3 \\ 0.3 & 1 \end{pmatrix}.$$

To cover linear time series, we consider data $\underline{X}_1, \ldots, \underline{X}_T$ from the bivariate VAR(1) models

$$\text{Model S(I) \& S(II)} \quad \underline{X}_t = A\underline{X}_{t-1} + \underline{e}_t, \tag{6.3}$$

where $\{\underline{e}_t, t \in \mathbb{Z}\}$ is a bivariate i.i.d. white noise process. For Model S(I), we let $\underline{e}_t \sim \mathcal{N}(0, \Sigma)$. For Model S(II), the first component of $\{\underline{e}_t, t \in \mathbb{Z}\}$ consists of i.i.d. uniformly distributed random variables, $e_{t,1} \sim \mathcal{R}(-\sqrt{3}, \sqrt{3})$ and the second component $\{e_{t,2}\}$ of t-distributed random variables with 5 degrees of freedom that are suitably multiplied such that $E(\underline{e}_t\underline{e}_t') = \Sigma$ holds. Observe that the excess kurtosis for these two innovation distributions are $-6/5$ and 6, respectively.

The two GARCH-type Models S(III) and S(IV) are based on two independent, but identically distributed univariate GARCH(1,1) processes $\{Y_{t,i}, t \in \mathbb{Z}\}$, $i = 1, 2$, each with

$$\text{Model S(III) \& S(IV)} \quad Y_{t,i} = \sigma_{t,i} e_{t,i}, \quad \sigma_{t,i}^2 = 0.01 + 0.3Y_{t-1,i}^2 + 0.5\sigma_{t-1,i}^2, \tag{6.4}$$

where $\{e_{t,i}, t \in \mathbb{Z}\}$, $i = 1, 2$, are two independent i.i.d. standard normal white noise processes. Now, Model S(III) and S(IV) correspond to the processes $\{\underline{X}_t = \Sigma^{1/2}(Y_{t,1}, Y_{t,2})', t \in \mathbb{Z}\}$ and $\{\underline{X}_t = \Sigma^{1/2}\{(|Y_{t,1}|, |Y_{t,2}|)' - E[(|Y_{t,1}|, |Y_{t,2}|)']\}, t \in \mathbb{Z}\}$, respectively (the first is the GARCH process, the second are the absolute values of the GARCH). Both these models are nonlinear and their fourth order cumulant structure is complex. Finally, we consider a VAR(1) regime switching model

$$\text{Model S(V)} \quad \underline{X}_t = \begin{cases} A\underline{X}_{t-1} + \underline{e}_t, & s_t = 0, \\ \underline{e}_t, & s_t = 1, \end{cases} \tag{6.5}$$

where $\{s_t\}$ is a (hidden) Markov process with two regimes such that $P(s_t \in \{0, 1\}) = 1$ and $P(s_t = s_{t-1}) = 0.95$ and $\{\underline{e}_t, t \in \mathbb{Z}\}$ is a bivariate i.i.d. white noise process with $\underline{e}_t \sim \mathcal{N}(0, \Sigma)$).

Realizations of stationary Models S(I)–S(V) are shown in Figure 1 together with the corresponding DFT covariances $T|\widehat{C}_{11}(r, 0)|^2$, $T|\sqrt{2}\widehat{C}_{21}(r, 0)|^2$ and $T|\widehat{C}_{22}(r, 0)|^2$, $r = 1, \ldots, 10$. The performance under the null of both tests $\mathcal{T}_{m,n,d}^*$ and $\mathcal{T}_{m,n,d;G}$ are reported in Table 1.

Discussion of the simulations under the null

For the stationary Models S(I)–S(V), the DFT covariances for lags $r = 1, \ldots, 10$ are shown in

Figure 1. These plots illustrate their different behaviors under Gaussianity and non-Gaussianity. In particular, for the Gaussian Model S(I), it can be seen that the DFT covariances seem to fit to the theoretical χ^2-distribution. Contrary to that, for the corresponding non-Gaussian Model S(II), they appear to have larger variances. Hence, in this case, it is necessary to use the bootstrap to estimate the proper variance in order to standardize the DFT covariances before constructing the test statistic. For the non-linear GARCH-type Models S(III) and S(IV), this effect becomes even more apparent and here it is absolutely necessary to use the bootstrap to correct for the larger variance (due to the fourth order cumulants). For the Markov switching Model S(V), this effect is also present, but not that strong in comparison to the GARCH-type models S(III) and S(IV).

In Table 1, the performance in terms of actual size of the bootstrap test $\mathcal{T}^*_{m,n,d}$ and of the non-bootstrap test $\mathcal{T}_{m,n,d;G}$ are presented. For Model S(I), where the underlying time series is Gaussian, the test $\mathcal{T}_{m,n,d;G}$ performs superior to $\mathcal{T}^*_{m,n,d}$, which tends to be conservative and underrejects the null. However, if we leave the Gaussian world, the corresponding non-Gaussian Model S(II) shows a different picture. In this case, the non-bootstrap test $\mathcal{T}_{m,n,d;G}$ clearly over-rejects the null significantly, where the bootstrap test $\mathcal{T}^*_{m,n,d}$ still remains conservative, but holds the prescribed level. For the GARCH-type Model S(III), both tests do not succeed in attaining the nominal level (over rejecting the null). However, there are two important factors which explain this. On the one hand, the non-bootstrap test $\mathcal{T}_{m,n,d;G}$ just does not take the fourth order structure contained in the process dynamics into account, which leads to a test that significantly overrejects the null, because in this case the DFT covariances are not properly standardized. On the other hand, the bootstrap procedure used for constructing $\mathcal{T}^*_{m,n,d}$ relies to a large extent on the choice of the tuning parameter p, which controls the average block length of the stationary bootstrap and, hence, for the dependence captured by the bootstrap samples. However, the data-driven rule (defined in Section 6.1) for selecting $1/p$ is based on the correlation structure of the data and the GARCH process is uncorrelated. This leads the rule to selecting a very small $1/p$ (typically it chooses a mean block length of 1 or 2). With such a small block length the fourth order cumulant in the variance cannot be estimated properly, indeed it underestimates it. For Model S(IV), we take the absolute values of GARCH processes, which induces serial correlation in the data. Hence, a the data-drive rule selects a larger tuning parameter $1/p$ in comparison to Model S(III). Therefore, a relatively accurate estimate of the (large) variance of the DFT covariance obtained, leading to the bootstrap test $\mathcal{T}^*_{m,n,d}$ attaining an accurate nominal level. However, as expected, the non-bootstrap test $\mathcal{T}_{m,n,d;G}$ fails to attain the nominal level (since the kurtosis of the GARCH model is large kurtosis, thus it is highly 'non-Gaussian'). Finally, the bootstrap test performs well for the VAR(1) switching Model S(V), whereas the non-bootstrap test $\mathcal{T}_{m,n,d;G}$ tends to slightly overreject the null.

6.2.3 Models under the alternative

To illustrate the behavior of the tests under the alternative of (second order) nonstationarity, we consider realizations from three models fulfilling different types of nonstationary behavior. As we focus on locally stationary alternatives, where nonstationarity is caused by smoothly changing dynamics, we consider first the time-varying VAR(1) model (tvVAR(1))

$$\text{Model NS(I)} \quad \underline{X}_t = A\underline{X}_{t-1} + \sigma(\frac{t}{T})\underline{e}_t, \quad t = 1, \ldots, T, \tag{6.6}$$

where $\sigma(u) = 2\sin(2\pi u)$. Further, we include a second tvVAR(1) model, where the dynamics are not present in the innovation variance, but in the coefficient matrix. More precisely, we consider the tvVAR(1) model

$$\text{Model NS(II)} \quad \underline{X}_t = A(\frac{t}{T})\underline{X}_{t-1} + \underline{e}_t, \quad t = 1, \ldots, T, \tag{6.7}$$

where $A(u) = \sin(2\pi u)\, A$. Finally, we consider the unit root case (noting that several authors have considered tests for stochastic trend, including Pelagatti and Sen (2013)), though this case has not been treated in our asymptotic theory. In particular, we consider observations from a bivariate random walk

$$\text{Model NS(III)} \quad \underline{X}_t = \underline{X}_{t-1} + \underline{e}_t, \quad t = 1, \ldots, T, \quad X_0 = 0. \tag{6.8}$$

In all Models NS(I)–NS(III) above, $\{\underline{e}_t, t \in \mathbb{Z}\}$ is a bivariate i.i.d. white noise process with $\underline{e}_t \sim \mathcal{N}(0, \Sigma)$.

In Figure 2 we show realizations of nonstationary Models NS(I)–NS(III) together with DFT covariances $T|\widehat{C}_{11}(r,0)|^2$, $T|\sqrt{2}\widehat{C}_{21}(r,0)|^2$ and $T|\widehat{C}_{22}(r,0)|^2$, $r = 1, \ldots, 10$ to illustrate how the type of nonstationarity is encoded. The performance under nonstationarity of both tests $\mathcal{T}^*_{m,n,d}$ and $\mathcal{T}_{m,n,d;G}$ are reported in Table 2 for sample size $T = 500$.

Discussion of the simulations under the alternative

The DFT covariances for the nonstationary Models NS(I)–NS(III) as displayed in Figures 2 illustrate how and why the proposed testing procedure is able to detect nonstationarity in the data. For both locally stationary Models NS(I) and NS(II), it can be seen that the nonstationarity is encoded mainly in the DFT covariances at lag two, where the peak is significantly more pronounced for Model NS(I) in comparison to Model NS(II). Contrary to that behavior, for the random walk Model NS(III), the DFT covariances are large for all lags.

In Table 2 we report the results for the tests, where the power for the bootstrap test $\mathcal{T}^*_{m,n,d}$ and for the non-bootstrap test $\mathcal{T}_{m,n,d;G}$ are given. It can be seen that both tests have good power properties for the tvVAR(1) Model NS(I), where the non-bootstrap test $\mathcal{T}_{m,n,d;G}$ is slightly superior to the bootstrap test $\mathcal{T}^*_{m,n,d}$. Here, it is interesting to note that the time-varying spectral density for Model NS(I) is $\mathbf{f}(u,\omega) = \frac{1}{2}(1-\cos(4\pi u))\mathbf{f}_Y(\omega)$, where $\mathbf{f}_Y(\omega)$ is the spectral density matrix corresponding to the stationary time series $\underline{Y}_t = A\underline{Y}_{t-1} + 2\underline{e}_t$. Comparing this to the coefficients Fourier $\mathbf{A}(r,0)$ (defined in (3.10)), we see that for this example $\mathbf{A}(2,0) \neq 0$ whereas $\mathbf{A}(r,0) \neq 0$ for $r \neq 2$ (which can be seen in Figure 2). In contrast, neither the bootstrap nor non-bootstrap test performs well for Model NS(II) (here the rejection rate is less than 40% even in the Gaussian case when using the 10% level). However, from Figure 2 of the DFT covariance we do see a clear peak at lag two, but this peak is substantially smaller than the corresponding peak in Model NS(1). A plausible explanation for the poor performance of the test in this case is that even when $m = 2$ the test we use a chi-square with $d(d+1)\times m = 2\times 3\times 2 = 12$ degrees of freedom which pushes the rejection region to the right, thus making it extremely difficult to reject the null unless the sample size or $\mathbf{A}(r,\ell)$ are extremely large. Since a visual inspection of the covariance shows clear signs of nonstationarity, this suggests that further work is needed in selecting which DFT covariances should be used in the testing procedure (especially in the multivariate setting where using a component wise scheme may be useful).

Finally, both tests have good power properties for the random walk Model NS(III). As the theory suggests (see Theorem 5.2), for all three nonstationary models the non-bootstrap procedure has better power than the bootstrap procedure.

6.3 Real data application

We now consider a real data example, in particular the log-returns over $T = 513$ trading days of the FTSE 100 and the DAX 30 stock price indexes between January 1st 2011-December 31st, 2012. A plot of both indexes is given in Figure 3. Typically, a stationary GARCH-type model is fitted to the log returns of stock index data. Therefore, in this section we investigate whether it is reasonable to assume that this time series is stationary. We first make a plot of the DFT covariances $T|\widehat{C}_{11}(r,0)|^2$, $T|\sqrt{2}\widehat{C}_{21}(r,0)|^2$ and $T|\widehat{C}_{22}(r,0)|^2$ (see Figure 3). We observe that most of the covariances are above the 5% level (however we note that $\widehat{C}_T(r,0)$ has not been standardized). We then apply the bootstrap test $\mathcal{T}^*_{m,n,d}$ and the non-bootstrap test $\mathcal{T}_{m,n,d;G}$ to the raw log-returns. In this case, both tests reject the null of second-order stationarity at the $\alpha = 1\%$ level. However, we recall from the simulation study in Section 6.2 (Models S(III) and S(IV)) that the tests tends to falsely reject the null for a GARCH model. Therefore, to make sure that the small p-value is not a mistake in the testing procedure, we consider the absolute values of log returns. A plot of the corresponding DFT covariances $T|\widehat{C}_{11}(r,0)|^2$, $T|\sqrt{2}\widehat{C}_{21}(r,0)|^2$ and $T|\widehat{C}_{22}(r,0)|^2$ is given in Figure 3. Applying the non-bootstrap test gives a p-value of less than 0.1% and the bootstrap test gives a p-value of 3.9%. Therefore, an analysis of both the log-returns and the absolute log-returns of the FTSE 100 and DAX 20 stock price indexes strongly suggest that this time series is nonstationary and fitting a stationary model to this data may not be appropriate.

Acknowledgement

The research of Carsten Jentsch was supported by the travel program of the German Academic Exchange Service (DAAD). Suhasini Subba Rao has been supported by the NSF grants DMS-0806096 and DMS-1106518. The authors are extremely grateful to the co-editor, associate editor and two anonymous referees whose valuable comments and suggestions substantially improved the quality and content of the paper.

A Proofs

A.1 Preliminaries

In order to derive its properties, we use that $\widetilde{c}_{j_1,j_2}(r,\ell)$ can be written as

$$\begin{aligned}
\widetilde{c}_{j_1,j_2}(r,\ell) &= \frac{1}{T}\sum_{k=1}^{T} \mathbf{L}_{j_1,\bullet}(\omega_k)\underline{J}_T(\omega_k)\overline{\underline{J}_T(\omega_{k+r})}'\,\overline{\mathbf{L}_{j_2,\bullet}(\omega_{k+r})}'\exp(i\ell\omega_k) \\
&= \frac{1}{T}\sum_{k=1}^{T}\sum_{s_1,s_2=1}^{d} L_{j_1,s_1}(\omega_k)J_{T,s_1}(\omega_k)\overline{J_{T,s_2}(\omega_{k+r})L_{j_2,s_2}(\omega_{k+r})}\exp(i\ell\omega_k),
\end{aligned}$$

where $L_{j,s}(\omega_k)$ is entry (j,s) of $\mathbf{L}(\omega_k)$ and $\mathbf{L}_{j,\bullet}(\omega_k)$ denotes its jth row.

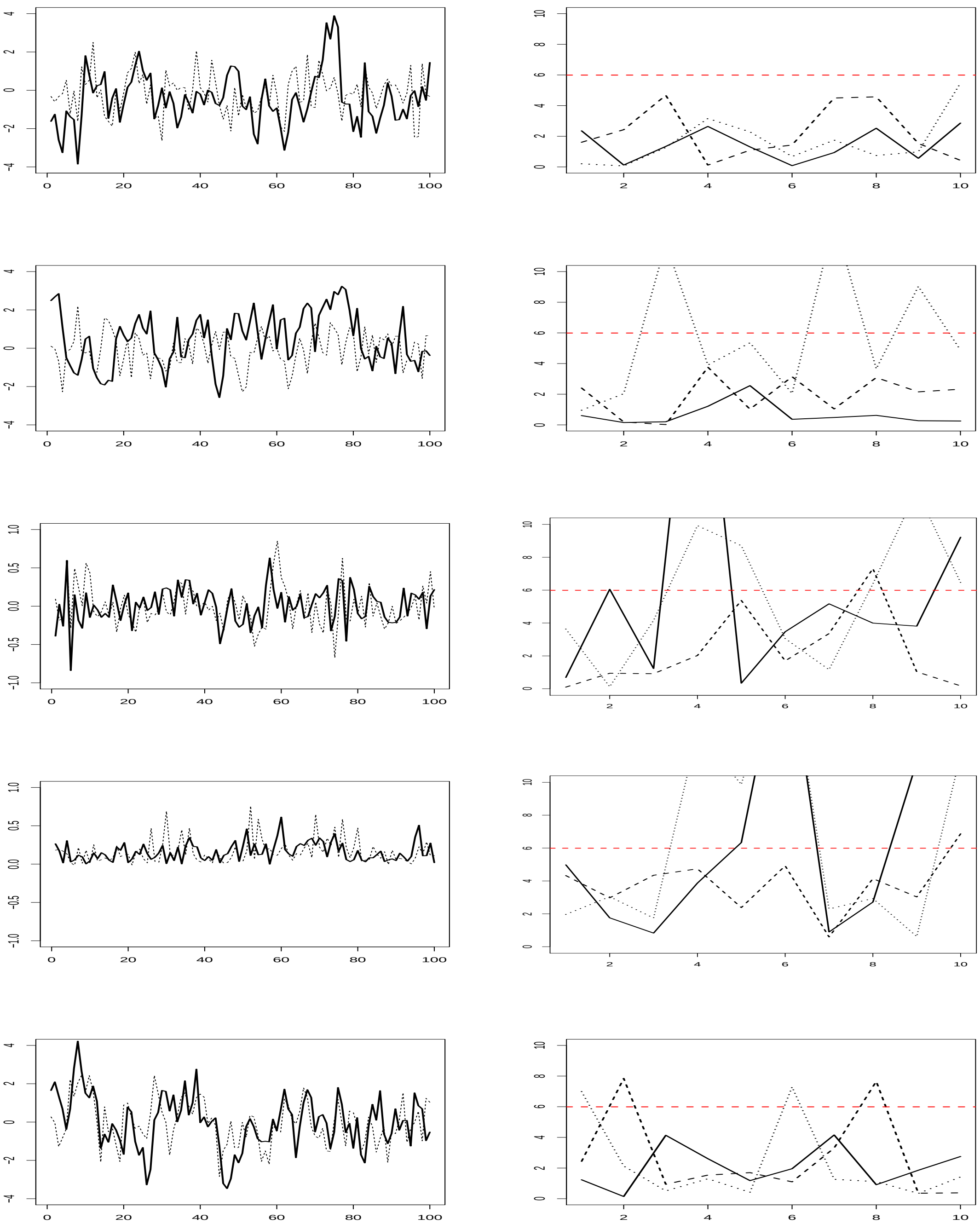

Figure 1: Stationary case: Bivariate realizations (left panels) and DFT covariances (right panels) $T|\widehat{C}_{11}(r,0)|^2$ (solid), $T|\sqrt{2}\widehat{C}_{21}(r,0)|^2$ (dashed) and $T|\widehat{C}_{22}(r,0)|^2$ (dotted) for stationary models S(I)–S(V) (top to bottom). The dashed red line is the 0.95-quantile of the χ^2 distribution with two degrees of freedom and DFT covariances are reported for sample size $T = 500$.

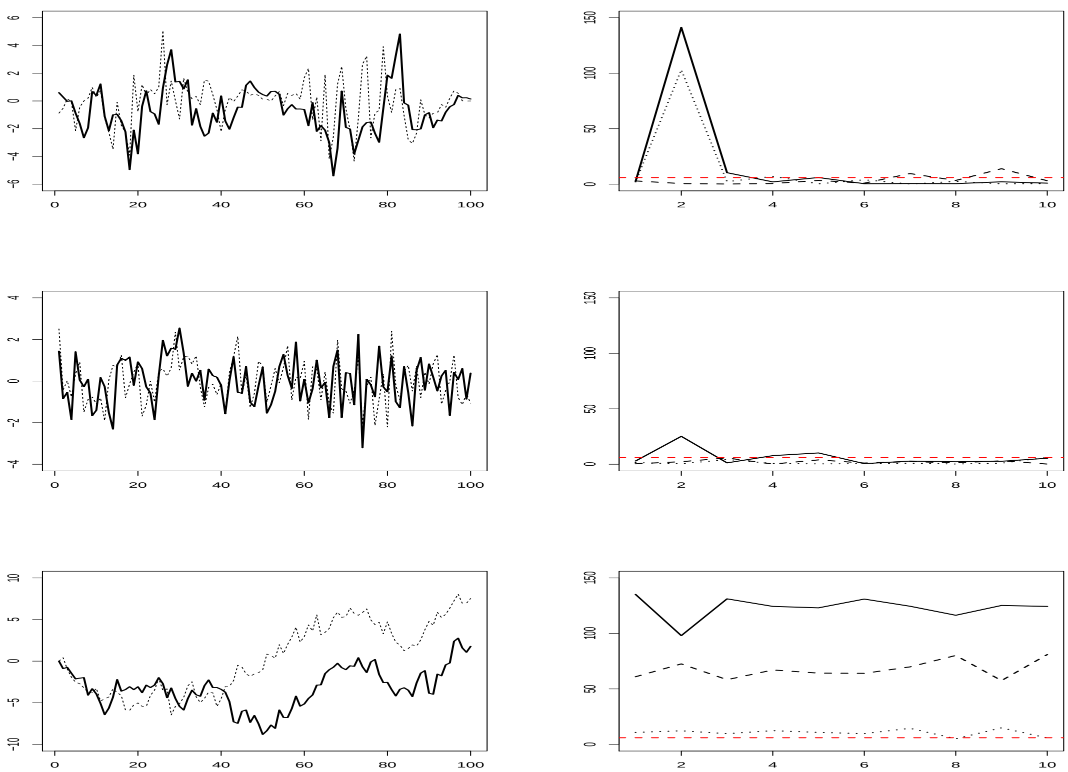

Figure 2: Nonstationary case: Bivariate realizations (left panels) and DFT covariances (right panels) $T|\widehat{C}_{11}(r,0)|^2$ (solid), $T|\sqrt{2}\widehat{C}_{21}(r,0)|^2$ (dashed) and $T|\widehat{C}_{22}(r,0)|^2$ (dotted) for nonstationary models S(I)–S(III) (top to bottom). The dashed red line is the 0.95-quantile of the χ^2 distribution with two degrees of freedom and DFT covariances are reported for sample size $T = 500$.

Model	α	$\mathcal{T}^*_{m,n,d}$	$\mathcal{T}_{m,n,d;G}$
S(I)	1%	0.00	0.00
	5%	0.50	3.00
	10%	1.25	6.00
S(II)	1%	0.00	21.25
	5%	0.25	32.25
	10%	1.00	40.25
S(III)	1%	55.00	89.75
	5%	69.00	93.50
	10%	76.50	96.50
S(IV)	1%	0.50	88.75
	5%	3.50	93.75
	10%	6.75	95.25
S(V)	1%	0.00	1.75
	5%	2.50	7.50
	10%	5.00	13.00

Table 1: Stationary case: Actual size of $\mathcal{T}^*_{m,n,d}$ and of $\mathcal{T}_{m,n,d;G}$ for $d = 2$, $n = 1$ for sample size $T = 500$ and stationary Models S(I)–S(V).

Model	α	$\mathcal{T}^*_{m,n,d}$	$\mathcal{T}_{m,n,d;G}$
NS(I)	1%	87.00	100.00
	5%	94.50	100.00
	10%	96.75	100.00
NS(II)	1%	2.75	10.75
	5%	9.75	24.25
	10%	16.50	35.25
NS(III)	1%	61.00	94.75
	5%	66.00	95.50
	10%	68.50	95.75

Table 2: Nonstationary case: Power of $\mathcal{T}^*_{m,n,d}$ and of $\mathcal{T}_{m,n,d;G}$ for $d = 2$, $n = 1$ for sample size $T = 500$ and nonstationary Models NS(I)–NS(II).

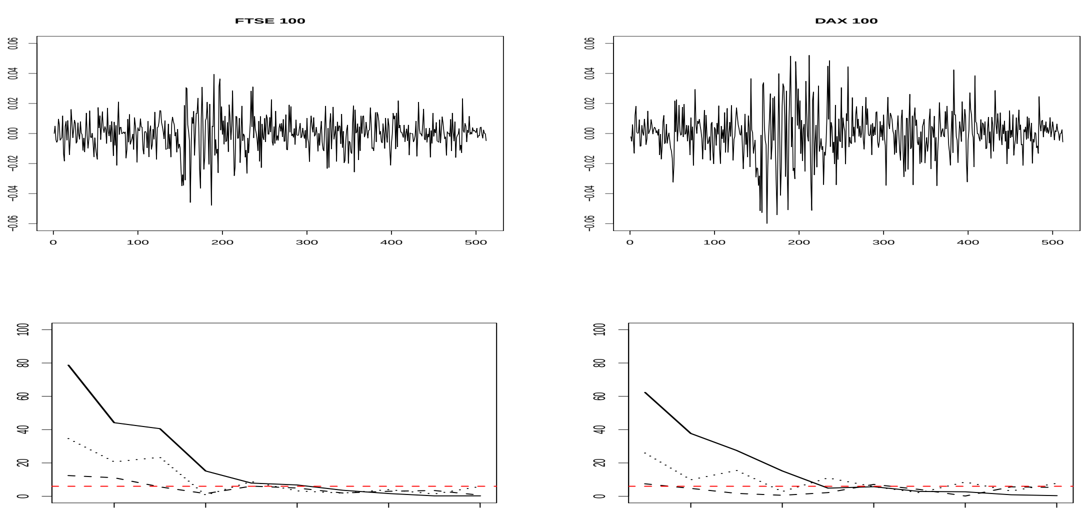

Figure 3: Log-returns of the FTSE 100 (top left panel) and of the DAX 30 (top right panel) stock price indexes over $T = 513$ trading days from January 1st, 2011 to December 31, 2012. Corresponding DFT covariances $T|\widehat{C}_{11}(r,0)|^2$ (solid, FTSE), $T|\sqrt{2}\widehat{C}_{21}(r,0)|^2$ (dashes) and $T|\widehat{C}_{22}(r,0)|^2$ (dotted, DAX) based on log-returns (bottom left panel) and on absolute values of log-returns (bottom right panel). The dashed red line is the 0.95-quantile of the χ^2 distribution with two degrees of freedom.

We will assume throughout the appendix that the lag window satisfies Assumption 2.1 and we will use the notation $\underline{f}(\omega) = \text{vec}(\mathbf{f}(\omega))$, $\widehat{\underline{f}}(\omega) = \text{vec}(\widehat{\mathbf{f}}(\omega))$, $J_{k,s} = J_{T,s}(\omega_k)$, $\underline{f}_k = \underline{f}(\omega_k)$, $\widehat{\underline{f}}_k = \widehat{\underline{f}}(\omega_k)$, $\underline{f}_{k,r} = (\text{vec}(\widehat{\mathbf{f}}(\omega_k)), \text{vec}(\widehat{\mathbf{f}}(\omega_{k+r})))$,

$$A_{j_1,s_1,j_2,s_2}(\underline{f}(\omega_1), \underline{f}(\omega_2)) = L_{j_1 s_1}(\underline{f}(\omega_1))\overline{L_{j_2 s_2}(\underline{f}(\omega_2))} \tag{A.1}$$

and $A_{j_1,s_1,j_2,s_2}(\underline{f}_{k,r}) = L_{j_1 s_1}(\underline{f}_k)\overline{L_{j_2 s_2}(\underline{f}_{k+r})}$. Furthermore, let us suppose that $\mathbf{G}$ is a positive definite matrix, $\underline{G} = vec(\mathbf{G})$ and define the lower-triangular matrix $\mathbf{L}(\underline{G})$ such that $\mathbf{L}(\underline{G})\mathbf{G}\overline{\mathbf{L}(\underline{G})}' = \mathbf{I}$ (hence $\mathbf{L}(\underline{G})$ is the inverse of the Cholesky decomposition of $\mathbf{G}$). Let $L_{js}(\underline{G})$ denote the (j,s)th element of the Cholesky matrix $\mathbf{L}(\underline{G})$. Let $\nabla L_{js}(\underline{G}) = (\frac{\partial L_{js}(\underline{G})}{\partial \mathbf{G}_{11}}, \ldots, \frac{\partial L_{js}(\underline{G})}{\partial \mathbf{G}_{dd}})'$ and $\nabla^n L_{js}(\underline{G})$ denote the vector of all partial nth order derivatives wrt $\underline{G}$. Furthermore, to reduce notation let $\widehat{L}_{js}(\omega) = L_{js}(\widehat{\underline{f}}(\omega))$ and $L_{js}(\omega) = L_{js}(\underline{f}(\omega))$. In the stationary case, let $\boldsymbol{\kappa}(h) = \text{cov}(\underline{X}_0, \underline{X}_h)$ and in the locally stationary case let $\kappa(u;h) = \text{cov}(\underline{X}_0(u), \underline{X}_h(u))$.

Before proving Theorems 3.1 and 3.5 we first state some preliminary results.

Lemma A.1 *(i) Let $\mathbf{G} = (g_{kl})$ be a positive definite $(d \times d)$ matrix. Then, for all $1 \leq j,s \leq d$ and all $r \in \mathbb{N}_0$, there exists an $\epsilon > 0$ and a set $\mathcal{M}_\epsilon = \{\mathbf{M} : |\mathbf{G} - \mathbf{M}|_1 < \epsilon$ and $\mathbf{M}$ is positive definite$\}$ such that*

$$\sup_{\mathbf{M} \in \mathcal{M}_\epsilon} |\nabla^r L_{js}(\mathbf{M})|_1 < \infty.$$

(ii) Let $\mathbf{G}(\omega)$ be a $(d \times d)$ uniformly continuous spectral density matrix function such that $\inf_\omega \lambda_{min}(\mathbf{G}(\omega)) > 0$. Then, for all $1 \leq j,s \leq d$ and all $r \in \mathbb{N}_0$, there exists an $\epsilon > 0$ and a set $\mathcal{M}_{\epsilon,\omega} = \{\mathbf{M}(\cdot) : |\mathbf{G}(\omega) - \mathbf{M}(\omega)|_1 < \epsilon$ and $\mathbf{M}(\omega)$ is positive definite for all $\omega\}$ such that

$$\sup_{\omega} \sup_{\mathbf{M}(\cdot) \in \mathcal{M}_\epsilon} |\nabla^r L_{js}(\mathbf{M}(\omega))|_1 < \infty.$$

PROOF. (i) For a positive definite matrix $\mathbf{M}$, let $\mathbf{M} = \mathbf{B}\overline{\mathbf{B}}'$, where $\mathbf{B}$ denotes the lower-triangular Cholesky decomposition of $\mathbf{M}$ and we set $\mathbf{C} = \mathbf{B}^{-1}$. Further, let $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$ be defined by $\mathbf{B} = \boldsymbol{\Psi}(\mathbf{G})$ and $\mathbf{C} = \boldsymbol{\Phi}(\mathbf{B})$, i.e. $\boldsymbol{\Psi}$ maps a positive definite matrix to its Cholesky matrix and $\boldsymbol{\Phi}$ maps an invertible matrix to its inverse. Further suppose $\lambda_{min}(\mathbf{G}) =: \underline{\eta}$ and $\lambda_{max}(\mathbf{G}) =: \overline{\eta}$ for some positive constants $\underline{\eta} \leq \overline{\eta}$ and let $\epsilon > 0$ be sufficiently small such that $0 < \underline{\eta} - \delta = \lambda_{min}(\mathbf{M}) \leq \lambda_{max}(\mathbf{M}) = \overline{\eta} + \delta < \infty$ for all $M \in \mathcal{M}_\epsilon$ and some $\delta > 0$. The latter is possible because the eigenvalues are continuous functions in the matrix entries. Now, due to

$$L_{kl}(\mathbf{M}) = c_{kl} = \Phi_{kl}(\mathbf{B}) = \Phi_{kl}(\boldsymbol{\Psi}(\mathbf{M}))$$

and the chain rule, it suffices to show that (a) all entries of $\mathbf{\Psi}$ have partial derivatives of all orders on the set of all positive definite matrices $\mathbf{M} = (m_{kl})$ with $0 < \underline{\eta} - \delta = \lambda_{min}(\mathbf{M}) \leq \lambda_{max}(\mathbf{M}) = \overline{\eta} + \delta < \infty$ for some $\delta > 0$ and (b) all entries of $\mathbf{\Phi}$ have partial derivatives of all orders on the set $\mathcal{L}_\epsilon$ of all lower triangular matrices with diagonal elements lying in $[\underline{\zeta}, \overline{\zeta}]$ for some suitable $0 < \underline{\zeta} \leq \overline{\zeta} < \infty$ depending on δ above such that $\mathbf{\Psi}(\mathcal{M}_\epsilon) \subset \mathcal{L}_\epsilon$. In particular, the diagonal entries (the eigenvalues) of $\mathbf{B}$ are bounded from above and are also bounded away from zero. As there are no explicit formulas for $\mathbf{B} = \mathbf{\Psi}(\mathbf{M})$ and $\mathbf{C} = \mathbf{\Phi}(\mathbf{B})$, their entries have to be calculated recursively by

$$b_{kl} = \begin{cases} \frac{1}{b_{ll}}(m_{kl} - \sum_{j=1}^{l-1} b_{kj}\bar{b}_{lj}), & k > l \\ (m_{kk} - \sum_{j=1}^{k-1} b_{kj}\bar{b}_{kj})^{1/2}, & k = l \\ 0, & k < l \end{cases} \quad \text{and} \quad c_{kl} = \begin{cases} -\frac{1}{b_{kk}} \sum_{j=l}^{k-1} b_{kj}c_{jl}, & k > l \\ \frac{1}{b_{kk}}, & k = l \\ 0, & k < l \end{cases},$$

where the recursion is done row by row (top first), starting from the left hand side of each row to the right. To prove (a), we order the non-zero entries of $\mathbf{B}$ row-wise and get for the first entry $\Psi_{11}(\mathbf{M}) = b_{11} = \sqrt{m_{11}}$, which is arbitrarily often partially differentiable as $m_{11} > 0$ is bounded away from zero on $\mathcal{M}_\epsilon$. Now we proceed recursively by induction. Suppose that $b_{kl} = \Psi_{kl}(\mathbf{M})$ is arbitrarily often partially differentiable for the first p non-zero elements of $\mathbf{B}$ on $\mathcal{M}_\epsilon$. The $(p+1)$th non-zero element is b_{st}, say. For $s = t$, we get

$$\Psi_{ss}(\mathbf{M}) = b_{ss} = \left(m_{ss} - \sum_{j=1}^{s-1} b_{sj}\bar{b}_{sj}\right)^{1/2} = \left(m_{ss} - \sum_{j=1}^{s-1} \Psi_{sj}(\mathbf{M})\overline{\Psi_{sj}(\mathbf{M})}\right)^{1/2},$$

and for $s > t$, we have

$$\Psi_{st}(\mathbf{M}) = b_{st} = \frac{1}{\Psi_{tt}(\mathbf{M})}\left(m_{st} - \sum_{j=1}^{t-1} \Psi_{sj}(\mathbf{M})\overline{\Psi_{tj}(\mathbf{M})}\right),$$

such that all partial derivatives of $\Psi_{st}(\mathbf{M})$ exist on $\mathcal{M}_\epsilon$ as $\Psi_{st}(\mathbf{M})$ is composed of such functions and due to $m_{ss} - \sum_{j=1}^{s-1} b_{sj}\bar{b}_{sj}$ and $\Psi_{tt}(\mathbf{M})$ uniformly bounded away from zero on $\mathcal{M}_\epsilon$. This proves part (a). To prove part (b), we get immediately that $\Phi_{kk}(\mathbf{B}) = c_{kk} = 1/b_{kk}$ has all partial derivatives on $\mathcal{L}_\epsilon$ as b_{kk} is bounded way from zero for all k. Now, we order the non-zero off-diagonal elements of $\mathbf{C}$ row-wise and for the first such entry we get $\Phi_{21}(\mathbf{B}) = c_{21} = -b_{21}c_{11}/b_{22}$ which is arbitrarily often partially differentiable again as b_{22} is bounded way from zero. Now we proceed again recursively by induction. Suppose that $c_{kl} = \Phi_{kl}(\mathbf{B})$ is arbitrarily often partially differentiable for the first p non-zero off-diagonal elements of $\mathbf{C}$. The $(p+1)$th non-zero element equals c_{st}, say, and we have

$$\Phi_{st}(\mathbf{B}) = c_{st} = -\frac{1}{b_{ss}}\sum_{j=l}^{s-1} b_{sj}c_{jt} = -\frac{1}{b_{ss}}\sum_{j=l}^{s-1} b_{sj}\Phi_{jt}(\mathbf{B})$$

and all partial derivatives of $\Phi_{st}(\mathbf{B})$ exist on $\mathcal{L}_\epsilon$ as $\Phi_{st}(\mathbf{B})$ is composed of such functions and due to $b_{ss} > 0$ uniformly bounded away from zero on $\mathcal{L}_\epsilon$. This proves part (b) and concludes part (i) of this proof.

(ii) As in part (i), we get with an analogue notation (depending on ω) the relation

$$L_{kl}(\mathbf{M}(\omega)) = c_{kl}(\omega) = \Phi_{kl}(\mathbf{B}(\omega)) = \Phi_{kl}(\mathbf{\Psi}(\mathbf{M}(\omega)))$$

and again by the chain rule, it suffices to show that (a) all entries of $\mathbf{\Psi}$ have partial derivatives of all orders on the set of all uniformly positive definite matrix functions $\mathbf{M}(\cdot)$ with $0 < \underline{\eta} - \delta = \inf_\omega \lambda_{min}(\mathbf{M}(\omega)) \leq \sup_\omega \lambda_{max}(\mathbf{M}(\omega)) = \overline{\eta} + \delta < \infty$ for some $\delta > 0$ and (b) all entries of $\mathbf{\Phi}$ have partial derivatives of all orders on the set $\mathcal{L}_{\epsilon,\omega}$ of all lower triangular matrix functions with diagonal elements lying in $[\underline{\zeta}, \overline{\zeta}]$ for some suitable $0 < \underline{\zeta} \leq \overline{\zeta} < \infty$ depending on δ such that $\mathbf{\Psi}(\mathcal{M}_{\epsilon,\omega}) \subset \mathcal{L}_{\epsilon,\omega}$. The rest of the proof of part (ii) is analogue to the proof of (i) above. $\square$

Lemma A.2 (Spectral density matrix estimator) *Suppose that $\{\underline{X}_t\}$ is a second order stationary or locally stationary time series (which satisfies Assumption 3.2(L2)) where for $h \neq 0$ either the covariance of local covariance satisfies $|\boldsymbol{\kappa}(h)|_1 \leq C|h|^{-(2+\varepsilon)}$ or $|\boldsymbol{\kappa}(u;h)|_1 \leq C|h|^{-(2+\varepsilon)}$ and $\sup_t \sum_{h_1,h_2,h_3} |\mathrm{cum}(X_{t,j_1}, X_{t+h_1,j_2}, X_{t+h_2,j_1}, X_{t+h_3,j_2})| < \infty$. Let $\widehat{\mathbf{f}}_T$ be defined as in (2.4).*

(a) $\mathrm{var}(\widehat{\mathbf{f}}_T(\omega_k)) = O((bT)^{-1})$ *and* $\sup_\omega |\mathrm{E}(\widehat{\mathbf{f}}_T(\omega)) - \mathbf{f}(\omega)| = O(b + (bT)^{-1})$.

(b) If in addition, we have $\sum_t |\mathrm{cum}(X_{t,j_1}, X_{t+r_1,j_2}, \ldots, X_{t+r_s,j_{s+1}})| < \infty$ for $s = 1, \ldots, 7$, then it holds
$\|\widehat{\mathbf{f}}_T(\omega) - \mathrm{E}(\widehat{\mathbf{f}}_T(\omega))\|_4 = O(\frac{1}{(bT)^{1/2}})$

(c) If in addition, $b^2T \to \infty$ then we have

  (i) $\sup_\omega |\widehat{\mathbf{f}}_T(\omega) - \mathbf{f}(\omega)|_1 \stackrel{P}{\to} 0$,

  (ii) Further, if $\mathbf{f}(\omega)$ is nonsingular on $[0, 2\pi]$, then we have $\sup_\omega |L_{js}(\underline{\widehat{f}}_T(\omega)) - L_{js}(\underline{f}(\omega))| \stackrel{P}{\to} 0$ as $T \to \infty$ for all $1 \leq j, s \leq d$.

PROOF. To simplify the proof most parts will be proven for the univariate case - the proof of the multivariate case is identical. By making a simple expansions it is straightforward to show that

$$\widehat{f}_T(\omega) = \frac{1}{2\pi T}\sum_{t,\tau=1}^{T}\lambda_b(t-\tau)(X_t-\mu)(X_\tau-\mu)\exp(i(t-\tau)\omega) + R_T(\omega),$$

where

$$R_T(\omega) = \frac{1}{2\pi T}\sum_{t,\tau=1}^{T}\lambda_b(t-\tau)(\bar{X}-\mu)\big((X_t-\mu)+(\mu-\bar{X})\big)$$

and under absolute summability of the second and fourth order cumulants we have $\mathrm{E}|\sup_\omega R_T(\omega)|^2 = O(\frac{1}{(Tb)^{3/2}} + \frac{1}{T})$ (similar bounds can also be obtained for higher moments if the corresponding cumulants are absolutely summable). We will show later on in the proof that this term is dominated by the leading term. Therefore, to simplify notation, as the mean estimator is insignificant, for the remainder of the proof we will assume that the mean is known and it is $\mathrm{E}(X_t) = 0$. Consequently, the mean is not estimated and the spectral density estimator is

$$\widehat{f}_T(\omega) = \frac{1}{2\pi T}\sum_{t,\tau=1}^{T}\lambda_b(t-\tau)X_tX_\tau\exp(i(t-\tau)\omega).$$

To prove (a) we evaluate the variance of $\widehat{f}_T(\omega)$

$$\mathrm{var}(\widehat{f}_T(\omega)) = \frac{1}{(2\pi)^2T^2}\sum_{t_1,\tau_1=1}^{T}\sum_{t_2,\tau_2=1}^{T}\lambda_b(t_1-\tau_1)\lambda_b(t_2-\tau_2)\mathrm{cov}(X_{t_1}X_{\tau_1},X_{t_2}X_{\tau_2})\exp(i(t_1-\tau_1-t_2+\tau_2)\omega).$$

By using indecomposable partitions on the covariances in the sum to partition it into covariances and cumulants of X_t and under the absolute summable covariance and cumulant assumptions, we have that $\mathrm{var}(\widehat{f}_T(\omega)) = O(\frac{1}{bT})$.

Next we obtain a bound for the bias. We do so, under the assumption of local stationarity, in particular the smooth assumptions in Assumptions 3.2(L2) (in the stationary case we do require these assumptions). Taking expectations we have

$$\begin{aligned}\mathrm{E}(\widehat{f}_T(\omega)) &= \frac{1}{2\pi T}\sum_{h=-T+1}^{T-1}\lambda_b(h)\exp(ih\omega)\sum_{t=1}^{T-|h|}\mathrm{cov}(X_t,X_{t+|h|})\\ &= \frac{1}{2\pi T}\sum_{th=-T+1}^{T-1}\lambda_b(h)\exp(ih\omega)\sum_{t=1}^{T-|h|}\kappa(\frac{t}{T};t-\tau) + R_1(\omega), \qquad (A.2)\end{aligned}$$

where

$$\sup_\omega|R_1(\omega)| \le \frac{1}{2\pi T}\sum_{t,\tau=1}^{T}|\lambda_b(t-\tau)|\big|\mathrm{cov}(X_t,X_\tau)-\kappa(\frac{t}{T};t-\tau)\big| = O(\frac{1}{T}) \text{ (by Assumption 3.2(L2)).}$$

Changing the inner sum in (A.2) with an integral gives

$$\mathrm{E}(\widehat{f}_T(\omega)) = \frac{1}{2\pi}\sum_{h=-T+1}^{T-1}\lambda_b(h)\exp(ih\omega)\kappa(h) + R_1(\omega) + R_2(\omega)$$

where

$$\sup_{\omega}|R_2(\omega)| \leq \frac{1}{2\pi}\sum_{h=-T}^{T}|\lambda_b(h)\left(|\frac{1}{T}\sum_{t=T-|h|+1}^{T}|\kappa(\frac{t}{T};h)| + \left|\frac{1}{T}\sum_{t=1}^{T}\kappa(\frac{t}{T};h) - \int_0^1\kappa(u;h)du\right|\right) = O(\frac{1}{bT}).$$

Finally, we take differences between $\mathrm{E}(\widehat{f}_T(\omega))$ and $f(\omega)$ which gives

$$\mathrm{E}(\widehat{f}_T(\omega)) - f(\omega) = \underbrace{\frac{1}{2\pi}\sum_{r=-1/b}^{1/b}\big(\lambda_b(r)-1\big)\kappa(r)\exp(ir\omega)}_{R_3(\omega)} + \underbrace{\frac{1}{2\pi}\sum_{|r|\geq 1/b}\kappa(r)\exp(ir\omega)}_{R_4(\omega)=O(b)} + R_1(\omega) + R_2(\omega).$$

To bound $R_3(\omega)$, we use Assumption 2.1 to give

$$\begin{aligned} R_3(\omega) &= \frac{1}{2\pi}\sum_{r=-T}^{T}\big(\lambda_b(r)-1\big)\kappa(r)\exp(ir\omega) = \frac{1}{2\pi}\sum_{r=-1/b}^{1/b}\big(\lambda_b(r)-1\big)\kappa(r)\exp(ir\omega) \\ &= \frac{b}{2\pi}\sum_{r=-T}^{T} r\cdot\lambda'(\overline{rb})\kappa(r)\exp(ir\omega) = \frac{1}{2\pi}\sum_{r=-1/b}^{1/b}\big(\lambda_b(r)-1\big)\kappa(r)\exp(ir\omega), \end{aligned}$$

where $\overline{rb}$ lies between 0 and rb. Therefore, we have $\sup_\omega|R_3(\omega)| = O(b)$. Altogether, this gives the bias $O(b+\frac{1}{bT})$ and we have proven (a).

To evaluate $\mathrm{E}|\widehat{f}_T(\omega) - \mathrm{E}(\widehat{f}_T(\omega))|^4$, we use the expansion

$$\mathrm{E}|\widehat{f}_T(\omega) - \mathrm{E}(\widehat{f}_T(\omega))|^4 = \underbrace{\mathrm{var}(\widehat{f}_T(\omega))^2}_{O((bT)^{-2})} + \mathrm{cum}_4(\widehat{f}_T(\omega)).$$

The bound for $\mathrm{cum}_4(\widehat{f}_T(\omega))$ uses an identical method to the variance calculation in part (a). By using the cumulant summability assumption we have $\mathrm{cum}_4(\widehat{f}_T(\omega)) = O(\frac{1}{(bT)^2})$, this proves (b).

We now prove (ci). By the triangle inequality, we have

$$\sup_\omega|\widehat{f}_T(\omega) - f(\omega)| \leq \sup_\omega|\widehat{f}_T(\omega) - \mathrm{E}(\widehat{f}(\omega))| + \underbrace{\sup_\omega|\mathrm{E}(\widehat{f}_T(\omega)) - f(\omega)|}_{O(b+(bT)^{-1})\text{ by (a)}}.$$

Therefore, we need only that the first term of the above converges to zero. To prove $\sup_\omega|\widehat{f}_T(\omega) - \mathrm{E}[\widehat{f}_T(\omega)]| \xrightarrow{P} 0$, we first show

$$\mathrm{E}\left(\sup_\omega\left|\widehat{f}_T(\omega) - \mathrm{E}(\widehat{f}_T(\omega))\right|^2\right) \to 0 \qquad \text{as } T\to\infty$$

and then we apply Chebyshev's inequality. To bound $\mathrm{E}\sup_\omega \left|\widehat{f}_T(\omega)-\mathrm{E}(\widehat{f}_T(\omega))\right|^2$, we will use Theorem 3B, page 85, Parzen (1999). There it is shown that if $\{X(\omega);\omega\in[0,\pi]\}$ is a zero mean stochastic process, then

$$\mathrm{E}\big(\sup_{0\leq\omega\leq\pi}|X(\omega)|^2\big)\leq\frac{1}{2}\mathrm{E}|X(0)|^2+\frac{1}{2}\mathrm{E}|X(\pi)|^2+\int_0^\pi\left[\mathrm{var}(X(\omega))\mathrm{var}\left(\frac{\partial X(\omega)}{\partial\omega}\right)\right]^{1/2}d\omega. \quad (A.3)$$

To apply the above lemma, let $X(\omega)=\widehat{f}_T(\omega)-\mathrm{E}[\widehat{f}_T(\omega)]$ and the derivative of $\widehat{f}_T(\omega)$ is

$$\frac{d\widehat{f}_T(\omega)|}{d\omega}=\frac{1}{2\pi T}\sum_{t,\tau=1}^{T}i(t-\tau)X_tX_\tau\lambda_b(t-\tau)\exp(i(t-\tau)\omega).$$

By using the same arguments as those used in (a), we have $\mathrm{var}(\frac{\partial\widehat{f}_T(\omega)}{\partial\omega})=O(\frac{1}{b^2T})$. Therefore, by using (A.3), we have

$$\begin{aligned}&\mathrm{E}\big(\sup_{0\leq\omega\leq\pi}|\widehat{f}_T(\omega)-\mathrm{E}(\widehat{f}_T(\omega))|^2\big)\\ \leq\ &\frac{1}{2}\mathrm{var}|\widehat{f}_T(0)|+\frac{1}{2}\mathrm{var}|\widehat{f}_T(\pi)|+\int_0^\pi\left[\mathrm{var}(\widehat{f}_T(\omega))\mathrm{var}(\frac{\partial\widehat{f}_T(\omega)}{\partial\omega})\right]^{1/2}d\omega=O\left(\frac{1}{b^{3/2}T}\right).\end{aligned}$$

Thus by using the above and Chebyshev's inequality, for any $\varepsilon>0$, we have

$$P\left(\sup_\omega\left|\widehat{f}_T(\omega)-\mathrm{E}[\widehat{f}_T(\omega)]\right|>\varepsilon\right)\leq\frac{\mathrm{E}\sup_\omega\left|\widehat{f}_T(\omega)-\mathrm{E}[\widehat{f}_T(\omega)]\right|^2}{\varepsilon^2}=O\left(\frac{1}{Tb^{3/2}\varepsilon}\right)\to0$$

as $Tb^{3/2}\to\infty$, $b\to0$ and $T\to\infty$. This proves (ci)

To prove (cii), we return to the multivariate case. We recall that a sequence $\{X_T\}$ converges in probability to zero if and only if for every subsequence $\{T_k\}$ there exists a subsequence $\{T_{k_i}\}$ such that $X_{T_{k_i}}\to0$ with probabiluty one (see, for example, (Billingsley, 1995), Theorem 20.5). Now, the uniform convergence in probability result in (ci) implies that for every sequence $\{T_k\}$ there exists a subsequence $\{T_{k_i}\}$ such that $\sup_\omega|\widehat{\underline{f}}_{T_{k_i}}(\omega)-\underline{f}(\omega)|\xrightarrow{P}0$ with probability one. Therefore, by applying the mean value theorem to L_{js}, we have

$$L_{js}(\widehat{\underline{f}}_{T_{k_i}}(\omega))-L_{js}(\underline{f}(\omega))=\nabla L_{js}(\bar{\underline{f}}_{T_{k_i}}(\omega))\big(\widehat{\underline{f}}_T(\omega)-\underline{f}(\omega)\big),$$

where $\bar{\mathbf{f}}_{T_{k_i}}(\omega)=\alpha_{T_{k_i}}(\omega)\widehat{\mathbf{f}}_{T_{k_i}}(\omega))+(1-\alpha_{T_{k_i}}(\omega))\mathbf{f}(\omega)$. Clearly $\bar{\mathbf{f}}_{T_{k_i}}(\omega)$ is a positive definite matrix and for a large enough T_k we have that $\bar{\mathbf{f}}_{T_{k_i}}(\omega)$ is positive definite and $\sup_\omega|\big(\widehat{\underline{f}}_{T_{k_i}}(\omega)-\underline{f}(\omega)\big)|<\varepsilon$ for all $T_{k_i}>T_k$. Thus, the conditions of Lemma A.1(ii) are satisfied and for large enough T_k we have that

$$\sup_\omega\left|L_{js}(\widehat{\underline{f}}_{T_{k_i}}(\omega))-L_{js}(\underline{f}(\omega))\right|\leq\sup_\omega\underbrace{|\nabla L_{js}(\bar{\underline{f}}_{T_{k_i}}(\omega))}_{\text{bounded in probability}}|\sup_\omega\left|\widehat{\underline{f}}_T(\omega)-\underline{f}(\omega)\right|\to0.$$

As the above result is true for every sequence $\{T_k\}$, we have proven (cii). $\square$

Above we have shown (the well known result) that spectral density estimator with unknown mean is asymptotically equivalent to the spectral density estimator as if the mean were known. Furthermore, we observe that in the definition of the DFT, we have not subtracted the mean, this is because $\underline{J}_T(\omega_k) = \tilde{\underline{J}}_T(\omega_k)$ for all $k \neq 0, T/2, T$, where

$$\tilde{\underline{J}}_T(\omega_k) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} (\underline{X}_t - \underline{\mu}) \exp(-it\omega_k), \tag{A.4}$$

with $\underline{\mu} = \mathrm{E}(\underline{X}_t)$. Therefore

$$\begin{aligned}
\widehat{\mathbf{C}}_T(r,\ell) &= \frac{1}{T}\sum_{k=1}^{T} \widehat{\mathbf{L}}(\omega_k)\underline{J}_T(\omega_k)\overline{\underline{J}_T(\omega_{k+r})}'\overline{\widehat{\mathbf{L}}(\omega_{k+r})}' \exp(i\ell\omega_k) \\
&= \frac{1}{T}\sum_{k=1}^{T} \widehat{\mathbf{L}}(\omega_k)\tilde{\underline{J}}_T(\omega_k)\overline{\tilde{\underline{J}}_T(\omega_{k+r})}'\overline{\widehat{\mathbf{L}}(\omega_{k+r})}' \exp(i\ell\omega_k) + O_p(\frac{1}{T})
\end{aligned}$$

uniformly over all frequencies. In other words, the DFT covariance is asymptotically the same as the DFT covariance constructed as if the mean were known. Therefore, from now onwards, in order to avoid unnecessary notation in the proofs, we will assume that the mean of the time series is zero and the spectral density matrix is estimated using

$$\begin{aligned}
&\widehat{\mathbf{f}}_T(\omega_s) \\
= \;& \frac{1}{2\pi T}\sum_{t,\tau=1}^{T} \lambda_b(t-\tau)\exp(i(t-\tau)\omega)\underline{X}_t\underline{X}_\tau' = \frac{1}{T}\sum_{j=-T/2}^{T/2} K_b(\omega_s-\omega_j)\underline{J}_T(\omega_j)\overline{\underline{J}_T(\omega_j)}',
\end{aligned} \tag{A.5}$$

where $K_b(\omega_j) = \sum_r \lambda_b(r)\exp(ir\omega_j)$.

A.2 Proof of Theorems 3.1 and 3.5

The main objective of this section is to prove Theorems 3.1 and 3.5. We will show that in the stationary case the leading term of $\widehat{\mathbf{C}}_T(r,\ell)$ is $\tilde{\mathbf{C}}_T(r,\ell)$, whereas in the nonstationary case it is $\tilde{\mathbf{C}}_T(r,\ell)$ plus two additional terms which are defined below. This is achieved by making a Taylor expansion and decomposing the difference $\widehat{\mathbf{C}}_T(r,\ell) - \tilde{\mathbf{C}}_T(r,\ell)$ into several terms (see Theorem A.3). On first impression, it may seem surprising that in the stationary case the bandwidth b does not have an influence on the asymptotic distribution of $\widehat{\mathbf{C}}_T(r,\ell)$. This can be explained by the decomposition below, where each of these terms are sums of DFTs. The DFTs over their frequencies behave like stochastic process with decaying correlation, how fast correlation decays depends on whether the underlying time series is stationary or not (see Lemmas A.4 and A.8 for the details).

We start by deriving the difference between $\sqrt{T}\big(\widehat{c}_{j_1,j_2}(r,\ell) - \widetilde{c}_{j_1,j_2}(r,\ell)\big)$.

Lemma A.3 *Suppose that the assumptions in Lemma A.2(c) hold. Then we have*

$$\sqrt{T}\big(\widehat{c}_{j_1,j_2}(r,\ell) - \widetilde{c}_{j_1,j_2}(r,\ell)\big) = A_{1,1} + A_{1,2} + \sqrt{T}(\mathcal{S}_{T,j_1,j_2}(r,\ell) + \mathcal{B}_{T,j_1,j_2}(r,\ell)) + O_p(A_2) + O_p(B_2),$$

where

$$\begin{aligned} A_{1,1} &= \frac{1}{\sqrt{T}}\sum_{k=1}^{T}\sum_{s_1,s_2=1}^{d}\big[J_{k,s_1}\overline{J_{k+r,s_2}} - \mathrm{E}(J_{k,s_1}\overline{J_{k+r,s_2}})\big]\big(\underline{\widehat{f}}_{k,r} - \mathrm{E}(\underline{\widehat{f}}_{k,r})\big)'\nabla A_{j_1,s_1,j_2,s_2}(\underline{f}_{k,r})e^{i\ell\omega_k} \\ A_{1,2} &= \frac{1}{\sqrt{T}}\sum_{k=1}^{T}\sum_{s_1,s_2=1}^{d}\big[J_{k,s_1}\overline{J_{k+r,s_2}} - \mathrm{E}(J_{k,s_1}\overline{J_{k+r,s_2}})\big]\big(\mathrm{E}(\underline{\widehat{f}}_{k,r}) - \underline{f}_{k,r}\big)'\nabla A_{j_1,s_1,j_2,s_2}(\underline{f}_{k,r})e^{i\ell\omega_k} \end{aligned} \tag{A.6}$$

$$\begin{aligned} A_2 &= \frac{1}{2\sqrt{T}}\sum_{k=1}^{T}\sum_{s_1,s_2=1}^{d}\big|J_{k,s_1}\overline{J_{k+r,s_2}} - \mathrm{E}(J_{k,s_1}\overline{J_{k+r,s_2}})\big|\cdot\Big|\big(\underline{\widehat{f}}_{k,r} - \underline{f}_{k,r}\big)'\nabla^2 A_{j_1,s_1,j_2,s_2}(\underline{f}_{k,r})\big(\underline{\widehat{f}}_{k,r} - \underline{f}_{k,r}\big)\Big|, \\ B_2 &= \frac{1}{2\sqrt{T}}\sum_{k=1}^{T}\sum_{s_1,s_2=1}^{d}\big|\mathrm{E}(J_{k,s_1}\overline{J_{k+r,s_2}})\big|\cdot\Big|\big(\underline{\widehat{f}}_{k,r} - \underline{f}_{k,r}\big)'\nabla^2 A_{j_1,s_1,j_2,s_2}(\underline{f}_{k,r})\big(\underline{\widehat{f}}_{k,r} - \underline{f}_{k,r}\big)\Big| \end{aligned}$$

and

$$\mathcal{S}_{T,j_1,j_2}(r,\ell) = \frac{1}{T}\sum_{k=1}^{T}\sum_{s_1,s_2=1}^{d}\mathrm{E}(J_{k,s_1}\overline{J_{k+r,s_2}})\big(\underline{\widehat{f}}_{k,r} - \mathrm{E}(\underline{\widehat{f}}_{k,r})\big)'\nabla A_{j_1,s_1,j_2,s_2}(\underline{f}_{k,r})e^{i\ell\omega_k} \tag{A.7}$$

$$\mathcal{B}_{T,j_1,j_2}(r,\ell) = \frac{1}{T}\sum_{k=1}^{T}\sum_{s_1,s_2=1}^{d}\mathrm{E}(J_{k,s_1}\overline{J_{k+r,s_2}})\big(\mathrm{E}(\underline{\widehat{f}}_{k,r}) - \underline{f}_{k,r}\big)'\nabla A_{j_1,s_1,j_2,s_2}(\underline{f}_{k,r})e^{i\ell\omega_k}.$$

PROOF. We decompose the difference between $\widehat{c}_{j_1,j_2}(r,\ell)$ and $\widetilde{c}_{j_1,j_2}(r,\ell)$ as

$$\begin{aligned} &\sqrt{T}\big(\widehat{c}_{j_1,j_2}(r,\ell) - \widetilde{c}_{j_1,j_2}(r,\ell)\big) = \frac{1}{\sqrt{T}}\sum_{k=1}^{T}\sum_{s_1,s_2=1}^{d}J_{k,s_1}\overline{J_{k+r,s_2}}\Big(A_{j_1,s_1,j_2,s_2}(\underline{\widehat{f}}_{k,r}) - A_{j_1,s_1,j_2,s_2}(\underline{f}_{k,r})\Big)e^{i\ell\omega_k} \\ &= \frac{1}{\sqrt{T}}\sum_{k=1}^{T}\sum_{s_1,s_2=1}^{d}\big[J_{k,s_1}\overline{J_{k+r,s_2}} - \mathrm{E}(J_{k,s_1}\overline{J_{k+r,s_2}})\big]\Big(A_{j_1,s_1,j_2,s_2}(\underline{\widehat{f}}_{k,r}) - A_{j_1,s_1,j_2,s_2}(\underline{f}_{k,r})\Big)e^{i\ell\omega_k} + \\ &\frac{1}{\sqrt{T}}\sum_{k=1}^{T}\sum_{s_1,s_2=1}^{d}\mathrm{E}(J_{k,s_1}\overline{J_{k+r,s_2}})\Big(A_{j_1,s_1,j_2,s_2}(\underline{\widehat{f}}_{k,r}) - A_{j_1,s_1,j_2,s_2}(\underline{f}_{k,r})\Big)e^{i\ell\omega_k} \\ &=: \ I + II. \end{aligned}$$

We observe that the difference depends on $A_{j_1,s_1,j_2,s_2}(\underline{\widehat{f}}_{k,r}) - A_{j_1,s_1,j_2,s_2}(\underline{f}_{k,r})$, therefore we replace this with the Taylor expansion

$$\begin{aligned} & A(\underline{\widehat{f}}_{k,r}) - A(\underline{f}_{k,r}) \\ =\ & \big(\underline{\widehat{f}}_{k,r} - \underline{f}_{k,r}\big)'\nabla A(\underline{f}_{k,r}) + \frac{1}{2}\big(\underline{\widehat{f}}_{k,r} - \underline{f}_{k,r}\big)'\nabla^2 A(\underline{\breve{f}}_{k,r})\big(\underline{\widehat{f}}_{k,r} - \underline{f}_{k,r}\big) \end{aligned} \tag{A.8}$$

with $\breve{\underline{f}}_{k,r}$ lying between $\widehat{\underline{f}}_{k,r}$ and $\underline{f}_{k,r}$ and A as defined in (A.1) (for clarity, both in the above and for the remainder of the proof we let $A = A_{j_1,s_1,j_2,s_2}$). Substituting the expansion (A.8) into I and II gives $I = A_1 + \breve{A}_2$ and $II = B_1 + \breve{B}_2$, where

$$A_1 = \frac{1}{\sqrt{T}} \sum_{k=1}^{T} \sum_{s_1,s_2=1}^{d} \left[J_{k,s_1}\overline{J_{k+r,s_2}} - \mathrm{E}(J_{k,s_1}\overline{J_{k+r,s_2}})\right] \left(\widehat{\underline{f}}_{k,r} - \underline{f}_{k,r}\right)' \nabla A(\underline{f}_{k,r}) e^{i\ell\omega_k}$$

$$B_1 = \frac{1}{\sqrt{T}} \sum_{k=1}^{T} \sum_{s_1,s_2=1}^{d} \mathrm{E}(J_{k,s_1}\overline{J_{k+r,s_2}}) \left(\widehat{\underline{f}}_{k,r} - \underline{f}_{k,r}\right)' \nabla A(\underline{f}_{k,r}) e^{i\ell\omega_k}$$

$$\breve{A}_2 = \frac{1}{2\sqrt{T}} \sum_{k=1}^{T} \sum_{s_1,s_2=1}^{d} \left[J_{k,s_1}\overline{J_{k+r,s_2}} - \mathrm{E}(J_{k,s_1}\overline{J_{k+r,s_2}})\right] \left(\widehat{\underline{f}}_{k,r} - \underline{f}_{k,r}\right)' \nabla^2 A(\breve{\underline{f}}_{k,r}) \left(\widehat{\underline{f}}_{k,r} - \underline{f}_{k,r}\right) e^{i\ell\omega_k},$$

$$\breve{B}_2 = \frac{1}{2\sqrt{T}} \sum_{k=1}^{T} \sum_{s_1,s_2=1}^{d} \mathrm{E}(J_{k,s_1}\overline{J_{k+r,s_2}}) \left(\widehat{\underline{f}}_{k,r} - \underline{f}_{k,r}\right)' \nabla^2 A(\breve{\underline{f}}_{k,r}) \left(\widehat{\underline{f}}_{k,r} - \underline{f}_{k,r}\right) e^{i\ell\omega_k}.$$

Next we substitute the decomposition $\widehat{\underline{f}}_{k,r} - \underline{f}_{k,r} = \widehat{\underline{f}}_{k,r} - \mathrm{E}(\widehat{\underline{f}}_{k,r}) + \mathrm{E}(\widehat{\underline{f}}_{k,r}) - \underline{f}_{k,r}$ into A_1 and B_1 to obtain $A_1 = A_{1,1} + A_{1,2}$ and $B_1 = \sqrt{T}[\mathcal{S}_{T,j_1,j_2}(r,\ell) + \mathcal{B}_{T,j_1,j_2}(r,\ell)]$. Therefore we have $I = A_{1,1} + A_{1,2} + \breve{A}_2$ and $II = \sqrt{T}(\mathcal{S}_{T,j_1,j_2}(r,\ell) + \mathcal{B}_{T,j_1,j_2}(r,\ell)) + \breve{B}_2$.

Finally, by using Lemma A.2(c) we have

$$\sup_{\omega_1,\omega_2} |\nabla^2 A(\widehat{\underline{f}}_T(\omega_1)', \widehat{\underline{f}}_T(\omega_2)') - \nabla^2 A(\underline{f}(\omega_1)', \underline{f}(\omega_2)')| \xrightarrow{P} 0.$$

Therefore, we take the absolute values of $\breve{A}_2$ and $\breve{B}_2$, and replace $\nabla^2 A(\breve{\underline{f}}_{k,r})$ with its deterministic limit $\nabla^2 A(\underline{f}_{k,r})$, to give the result. □

To simplify the notation in the rest of this section, we will drop the multivariate suffix and assume that we are in the univariate setting (the proof is identical for the multivariate case). Therefore

$$\sqrt{T}\left[\widehat{c}(r,\ell) - \widetilde{c}(r,\ell)\right] = A_{1,1} + A_{1,2} + \sqrt{T}(\mathcal{S}_T(r,\ell) + \mathcal{B}_T(r,\ell)) + O_p(A_2) + O_p(B_2), \tag{A.9}$$

where

$$A_{1,1} = \frac{1}{\sqrt{T}} \sum_{k=1}^{T} \left[J_k\overline{J_{k+r}} - \mathrm{E}(J_k\overline{J_{k+r}})\right] \left[\widehat{\underline{f}}_{k,r} - \mathrm{E}(\widehat{\underline{f}}_{k,r})\right] G(\omega_k)$$

$$A_{1,2} = \frac{1}{\sqrt{T}} \sum_{k=1}^{T} \left[J_k\overline{J_{k+r}} - \mathrm{E}(J_k\overline{J_{k+r}})\right] \left[\mathrm{E}(\widehat{\underline{f}}_{k,r}) - \underline{f}_{k,r}\right] G(\omega_k) \tag{A.10}$$

$$\mathcal{S}_T(r,\ell) = \frac{1}{\sqrt{T}} \sum_{k=1}^{T} \mathrm{E}(J_k\overline{J_{k+r}}) \left[\widehat{\underline{f}}_{k,r} - \mathrm{E}(\widehat{\underline{f}}_{k,r})\right] G(\omega_k) \tag{A.11}$$

$$\mathcal{B}_T(r,\ell) = \frac{1}{\sqrt{T}} \sum_{k=1}^{T} \mathrm{E}(J_k\overline{J_{k+r}}) \left[\mathrm{E}(\widehat{\underline{f}}_{k,r}) - \underline{f}_{k,r}\right] G(\omega_k)$$

$$A_2 = \frac{1}{\sqrt{T}}\sum_{k=1}^{T}\left|J_k\overline{J_{k+r}} - \mathrm{E}(J_k\overline{J_{k+r}})\right| \cdot \left|\left(\widehat{\underline{f}}_{k,r} - \underline{f}_{k,r}\right)^2 H(\omega_k)\right|,$$

$$B_2 = \frac{1}{\sqrt{T}}\sum_{k=1}^{T}\left|\mathrm{E}(J_k\overline{J_{k+r}})\right| \cdot \left|\left(\widehat{\underline{f}}_{k,r} - \underline{f}_{k,r}\right)^2 H(\omega_k)\right|, \quad \text{(A.12)}$$

with $G(\omega_k) = \nabla A(\underline{f}_{k,r})e^{i\ell\omega_k}$, $H(\omega_k) = \nabla^2 A(\underline{f}_{k,r})$ and $J_k = J_T(\omega_k)$. In the following lemmas we obtain bounds for each of these terms.

In the proofs below, we will often use the result that if the cumulants are absolutely summable, in the sense that $\sup_t \sum_{j_1,\ldots,j_{n-1}} |\mathrm{cum}(X_t, X_{t+j_1}, \ldots, X_{t+j_{n-1}})| < \infty$, then

$$\sup_{\omega_1,\ldots,\omega_n} \left|\mathrm{cum}(J_T(\omega_1), \ldots, J_T(\omega_n))\right| \leq \frac{K}{T^{n/2-1}} \quad \text{(A.13)}$$

for some constant K.

In the following lemma, we bound $A_{1,1}$.

Lemma A.4 *Suppose that for $1 \leq n \leq 8$, we have $\sum_{j_1,\ldots,j_{n-1}} |\mathrm{cum}(X_t, X_{t+j_1}, \ldots, X_{t+j_{n-1}})| < \infty$.*

(i) In addition, suppose that for $r \neq 0$, we have $|\mathrm{cov}(J_T(\omega_k), J_T(\omega_{k+r}))| = O(T^{-1})$, then $\|A_{1,1}\|_2 = O(\frac{1}{\sqrt{T}} + \frac{1}{Tb})$.

(ii) On the other hand, suppose that $\sum_{k=1}^{T} |\mathrm{cov}(J_T(\omega_k), J_T(\omega_{k+r}))| \leq C \log T$, then we have $\|A_{1,1}\|_2 \leq C(\frac{\log T}{b\sqrt{T}})$.

PROOF. We prove the result in case (ii) (the proof of (i) is a simpler version on this result). By using the spectral representation of the spectral density function in (A.5), we have

$$A_{1,1} = \frac{1}{T^{3/2}}\sum_{k,l} H(\omega_k)K_b(\omega_k - \omega_l)\left(J_k\overline{J}_{k+r} - \mathrm{E}(J_k\overline{J}_{k+r})\right)\left(|J_l|^2 - \mathrm{E}|J_l|^2\right). \quad \text{(A.14)}$$

Evaluating the expectation of $A_{1,1}$ gives

$$\begin{aligned} \mathrm{E}(A_{1,1}) &= \frac{1}{T^{3/2}}\sum_{k,l} H(\omega_k)K_b(\omega_k - \omega_l)\mathrm{cov}\left(J_k\overline{J}_{k+r}, J_l\overline{J}_l\right) \\ &= \frac{1}{T^{3/2}}\sum_{k,l} H(\omega_k)K_b(\omega_k - \omega_l)\Big(\mathrm{cov}(J_k, J_l)\mathrm{cov}(\bar{J}_{k+r}, \bar{J}_l) + \mathrm{cov}(J_k, \bar{J}_l)\mathrm{cov}(\bar{J}_{k+r}, J_l) + \mathrm{cum}(J_k, \bar{J}_{k+r}, J_l, \bar{J}_l)\Big) \\ &=: I + II + III. \end{aligned}$$

By using that $\sum_r \mathrm{E}(J_k\overline{J}_{k+r}) \leq C \log T$, we can show $I, II = O(\frac{\log T}{b\sqrt{T}})$ and by using (A.13), we have $III = O(\frac{1}{\sqrt{T}})$. Altogether, this gives $\mathrm{E}(A_{1,1}) = O(\frac{\log T}{b\sqrt{T}})$ (under the conditions in (i) we have $\mathrm{E}(A_{1,1}) = O(\frac{1}{\sqrt{T}})$).

We now evaluate a bound for $\mathrm{var}(A_{1,1})$. Again using (A.14) gives

$$\mathrm{var}(A_{1,1}) = \frac{1}{T^3}\sum_{k_1,l_1}\sum_{k_2,l_2} H(\omega_{k_1})H(\omega_{k_2})K_b(\omega_{k_1}-\omega_{l_1})K_b(\omega_{k_2}-\omega_{l_2})$$

$$\times \mathrm{cov}\Big(\big(J_{k_1}\overline{J}_{k_1+r} - \mathrm{E}(J_{k_1}\overline{J}_{k_1+r})\big)\big(J_{l_1}\overline{J}_{l_1} - \mathrm{E}(J_{l_1}\overline{J}_{l_1})\big), \big(J_{k_2}\overline{J}_{k_2+r} - \mathrm{E}(J_{k_2}\overline{J}_{k_2+r})\big)\big(J_{l_2}\overline{J}_{l_2} - \mathrm{E}(J_{l_2}\overline{J}_{l_2})\big)\Big).$$

By using indecomposable partitions (see Brillinger (1981) for the definition) we can show that $\mathrm{var}(A_{1,1}) = O(\frac{(\log T)^3}{T^3b^2})$ (under (i) it will be $\mathrm{var}(A_{1,1}) = O(\frac{1}{T^3b^2})$). This gives the desired result. □

In the following lemma, we bound $A_{1,2}$.

Lemma A.5 *Suppose that* $\sup_t \sum_{t_1,t_2,t_3} |\mathrm{cum}(X_t, X_{t+t_1}, X_{t+t_2}, X_{t+t_3})| < \infty$.

(i) In addition, suppose that for $r \neq 0$ *we have* $|\mathrm{cov}(J_T(\omega_k, J_T(\omega_{k+r}))| = O(T^{-1})$, *then* $\|A_{1,2}\|_2 \leq C\sup_\omega |\mathrm{E}(\widehat{f}_T(\omega)) - f(\omega)|$.

(ii) On the other hand, suppose that $\sum_{k=1}^{T} |\mathrm{cov}(J_T(\omega_k), J_T(\omega_{k+r}))| \leq C\log T$, *then we have* $\|A_{1,2}\|_2 \leq C\log T \sup_\omega |\mathrm{E}(\widehat{f}_T(\omega)) - f(\omega)|$.

PROOF. Since the mean of $A_{1,2}$ is zero, we evaluate the variance

$$\mathrm{var}\Big(\frac{1}{\sqrt{T}}\sum_{k=1}^{T} h_k\big(J_{k,j_1}\overline{J}_{k+r,j_2} - \mathrm{E}(J_{k,j_1}\overline{J}_{k+r,j_2})\big)\Big) = \frac{1}{T}\sum_{k_1,k_2=1}^{T} h_{k_1}h_{k_2}\mathrm{cov}\big(J_{k_1}\overline{J}_{k_1+r}, J_{k_2}\overline{J}_{k_2+r}\big)$$

$$= \frac{1}{T}\sum_{k_1,k_2=1}^{T} h_{k_1}h_{k_2}\Big\{\mathrm{cov}\big(J_{k_1}, J_{k_2}\big)\mathrm{cov}\big(\overline{J}_{k_1+r}, \overline{J}_{k_2+r}\big) + \mathrm{cov}\big(J_{k_1}, \overline{J}_{k_2+r}\big)\mathrm{cov}\big(\overline{J}_{k_1+r}, J_{k_2}\big) + \mathrm{cum}\big(J_{k_1}, \overline{J}_{k_1+r}, J_{k_2}, \overline{J}_{k_2+r}\big)\Big\}.$$

Therefore, under the stated conditions, and by using (A.13), the result immediately follows. □

In the following lemma, we bound A_2 and B_2.

Lemma A.6 *Suppose* $\{X_t\}_t$ *is a time series where for* $n = 2, \ldots, 8$, *we have* $\sup_t \sum_{t_1,\ldots,t_{n-1}} |\mathrm{cum}(X_t, X_{t+j_1}, \ldots, X_{t+j_{n-1}})| < \infty$ *and the assumptions of Lemma A.2 are satisfied. Then*

$$\big\|J_k\overline{J_{k+r}} - \mathrm{E}\big(J_k\overline{J_{k+r}}\big)\big\|_2 = O(1) \qquad \text{(A.15)}$$

and

$$\|A_2\|_1 = O\left(\frac{1}{b\sqrt{T}} + b^2\sqrt{T}\right) \qquad \|B_2\|_2 = O\left(\frac{1}{b\sqrt{T}} + b^2\sqrt{T}\right). \qquad \text{(A.16)}$$

PROOF. We have

$$J_k\overline{J_{k+r}} - \mathrm{E}\big(J_k\overline{J_{k+r}}\big) = \frac{1}{2\pi T}\sum_{t,\tau=1}^{T}\rho_{t,\tau}\big(X_tX_\tau - \mathrm{E}(X_tX_\tau)\big),$$

where $\rho_{t,\tau} = \exp(i\omega_k(t-\tau))\exp(-i\omega_r\tau)$. Now, by evaluating the variance, we get

$$\mathrm{E}|J_k\overline{J_{k+r}} - \mathrm{E}\big(J_k\overline{J_{k+r}}\big)\big|^2 \leq \frac{1}{(2\pi)^2}(I + II + III), \tag{A.17}$$

where

$$\begin{aligned} I &= T^{-2}\sum_{t_1,t_2=1}^{T}\sum_{\tau_1,\tau_2=1}^{T}\rho_{t_1,\tau_1}\overline{\rho_{t_2,\tau_2}}\mathrm{cov}(X_{t_1},X_{t_2})\mathrm{cov}(X_{\tau_1},X_{\tau_2}) \\ II &= T^{-2}\sum_{t_1,t_2=1}^{T}\sum_{\tau_1,\tau_2=1}^{T}\rho_{t_1,\tau_1}\overline{\rho_{t_2,\tau_2}}\mathrm{cov}(X_{t_1},X_{\tau_2})\mathrm{cov}(X_{\tau_1},X_{t_2}) \\ III &= T^{-2}\sum_{t_1,t_2=1}^{T}\sum_{\tau_1,\tau_2=1}^{T}\rho_{t_1,\tau_1}\overline{\rho_{t_2,\tau_2}}\mathrm{cum}(X_{t_1},X_{\tau_1},X_{t_2},X_{\tau_2}). \end{aligned}$$

Therefore, by using $\sup_t\sum_\tau|\mathrm{cov}(X_t,X_\tau)| < \infty$ and $\sup_t\sum_{\tau_1,t_2,\tau_2}|\mathrm{cum}(X_t,X_{\tau_1},X_{t_2},X_{\tau_2})| < \infty$, we have (A.15).

To obtain a bound for A_2 and B_2 we simply use the Cauchy-Schwarz inequality to obtain

$$\begin{aligned} \|A_2\|_1 &\leq \frac{1}{\sqrt{T}}\sum_{k=1}^{T}\big|J_k\overline{J_{k+r}} - \mathrm{E}(J_k\overline{J_{k+r}})\big|_2\cdot\big\|\underline{\widehat{f}}_{k,r} - \underline{f}_{k,r}\big\|_4^2|H(\omega_k)|, \\ \|B_2\|_2 &\leq \frac{1}{\sqrt{T}}\sum_{k=1}^{T}\big|\mathrm{E}(J_k\overline{J_{k+r}})\big|\cdot\big\|\underline{\widehat{f}}_{k,r} - \underline{f}_{k,r}\big\|_2^2|H(\omega_k)|. \end{aligned}$$

Thus, by using (A.15) and Lemma A.2(a,b), we have (A.16). □

Finally, we obtain bounds for $\sqrt{T}\mathcal{S}_T(r,\ell)$ and $\sqrt{T}\mathcal{B}_T(r,\ell)$.

Lemma A.7 *Suppose $\{X_t\}_t$ is a time series whose cumulants satisfy $\sup_t\sum_h|\mathrm{cov}(X_t,X_{t+h})| < \infty$ and $\sup_t\sum_{t_1,t_2,t_3}|\mathrm{cum}(X_t,X_{t+t_1},X_{t+t_2},X_{t+t_3})| < \infty$.*

(i) If $|\mathrm{E}(J_T(\omega_k)\overline{J}_T(\omega_{k+r}))| = O(\frac{1}{T})$ for all k and $r \neq 0$, then

$$\|\mathcal{S}_T(r,\ell)\|_2 \leq K\left(\frac{1}{b^{1/2}T^{3/2}}\right) \quad \textit{and} \quad |\mathcal{B}_T(r,\ell)| = O\left(\frac{b}{\sqrt{T}}\right).$$

(ii) On the other hand, if for fixed r and k we have $|\mathrm{E}(J_T(\omega_k)\overline{J}_T(\omega_{k+r}))| = h(\omega_k,r) + O(\frac{1}{T})$ (where $h(\cdot,r)$ is a function with a bounded derivative over $[0,2\pi]$) and the conditions in Lemma A.2(a) hold, then we have

$$\|\mathcal{S}_T(r,\ell)\|_2 = O(T^{-1/2}) \quad \textit{and} \quad |\mathcal{B}_T(r,\ell)| = O(b).$$

PROOF. We first prove (i). Bounding $\|\mathcal{S}_T(r,\ell)\|_2$ and $|\mathcal{B}_T(r,\ell)|$ gives

$$
\begin{aligned}
\|\mathcal{S}_T(r,\ell)\|_2 &\leq \frac{1}{T}\sum_{k=1}^{T}|\mathrm{E}(J_k\overline{J_{k+r}})|\left\|\underline{\widehat{f}}_{k,r}-\mathrm{E}(\underline{\widehat{f}}_{k,r})\right\|_2|G(\omega_k)|,\\
|\mathcal{B}_T(r,\ell)| &= \frac{1}{T}\sum_{k=1}^{T}|\mathrm{E}(J_k\overline{J_{k+r}})|\left|\mathrm{E}(\underline{\widehat{f}}_{k,r})-\underline{f}_{k,r}\right|\cdot|G(\omega_k)|
\end{aligned}
$$

and by substituting the bounds in Lemma A.2(a) and $|\mathrm{E}(J_k\bar{J}_{k+r})|=O(T^{-1})$ into the above, we obtain (i).

The proof of (ii) is rather different. We don't obtain the same bounds as in (i), because we do not have $|\mathrm{E}(J_k\bar{J}_{k+r})|=O(T^{-1})$. To bound $\mathcal{S}_T(r,\ell)$, we rewrite it as a quadratic form (see Section A.4 for the details)

$$
\begin{aligned}
&\mathcal{S}_T(r,\ell)\\
=&\ \frac{-1}{2T\pi}\sum_{k=1}^{T}\mathrm{E}(J_k\overline{J}_{k+r})\exp(i\ell\omega_k)\left(\frac{\hat{f}_T(\omega_k)-\mathrm{E}(\hat{f}_T(\omega_k))}{\sqrt{f(\omega_k)^3f(\omega_{k+r})}}+\frac{\hat{f}_T(\omega_{k+r})-\mathrm{E}(\hat{f}_T(\omega_{k+r}))}{\sqrt{f(\omega_k)f(\omega_{k+r})^3}}\right)\\
=&\ \frac{-1}{2T\pi}\sum_{k=1}^{T}h(\omega_k,r)\exp(i\ell\omega_k)\left(\frac{\hat{f}_T(\omega_k)-\mathrm{E}(\hat{f}_T(\omega_k))}{\sqrt{f(\omega_k)^3f(\omega_{k+r})}}+\frac{\hat{f}_T(\omega_{k+r})-\mathrm{E}(\hat{f}_T(\omega_{k+r}))}{\sqrt{f(\omega_k)f(\omega_{k+r})^3}}\right)+O(\frac{1}{T})\\
=&\ \frac{-1}{2T\pi}\sum_{t,\tau}\lambda_b(t-\tau)(X_tX_\tau-\mathrm{E}(X_tX_\tau)\underbrace{\frac{1}{T}\sum_{k=1}^{T}h(\omega_k,r)e^{i\ell\omega_k}\left(\frac{\exp(i(t-\tau)\omega_k)}{\sqrt{f(\omega_k)^3f(\omega_{k+r})}}+\frac{\exp(i(t-\tau)\omega_{k+r})}{\sqrt{f(\omega_k)f(\omega_{k+r})^3}}\right)}_{\approx d_{\omega_r}(t-\tau+\ell)}+O(\frac{1}{T}).
\end{aligned}
$$

We show in Lemma A.12 that the coefficients satisfy $|d_{\omega_r}(s)|_1\leq C(|s|^{-2}+T^{-1})$. Using this we can show that $\mathrm{var}(\mathcal{S}_T(r,\ell))=O(T^{-1/2})$. The bound on $\mathcal{B}_T(r,\ell)$ follows from Lemma A.2(a). $\square$

Having bounds for all the terms in Lemma A.3, we now show that Assumptions 3.1 and 3.2 satisfy the conditions under which we obtain these bounds.

Lemma A.8 *(i) Suppose $\{X_t\}$ is a second order stationary time series with $\sum_j|\mathrm{cov}(X_0,X_j)|<\infty$. Then, we have $\max_{1\leq k\leq T}|\mathrm{cov}(J_T(\omega_k),J_T(\omega_{k+r}))|\leq\frac{K}{T}$ for $r\neq 0,T/2,T$.*

(ii) Suppose Assumption 3.2(L2) holds. Then, we have

$$
\mathrm{cov}\big(J_T(\omega_{k_1}),J_T(\omega_{k_2})\big)=h(\omega_{k_1},k_1-k_2)+R(\omega_{k_1},\omega_{k_2}),\qquad\text{(A.18)}
$$

where $\mathrm{E}|\sup_{\omega_1,\omega_2}|R_T(\omega_1,\omega_2)|=O(T^{-1})$ and

$$
h(\omega,k_1-k_2)=\int_0^1 f(u,\omega)\exp(-2\pi i(k_1-k_2)u)du.
$$

PROOF. (i) follows from (Brillinger, 1981), Theorem 4.3.2.

To prove (ii) under local stationarity, we expand $\text{cov}\big(J_{k_1}, J_{k_2}\big)$ to give

$$\text{cov}\big(J_{k_1}, J_{k_2}\big) = \frac{1}{2\pi T}\sum_{t,\tau=1}^{T}\text{cov}(X_{t,T}, X_{\tau,T})\exp(-i(t-\tau)\omega_{k_1} + \tau(\omega_{k_1} - \omega_{k_2})).$$

Now, using Assumption 3.2(L2), we can replace $\text{cov}(X_{t,T}, X_{\tau,T})$ with $\kappa(\frac{t}{T}; t-\tau)$ to give

$$\begin{aligned}\text{cov}\big(J_{k_1}, J_{k_2}\big) &= \frac{1}{2\pi T}\sum_{t,\tau=1}^{T}\kappa(\frac{t}{T}, t-\tau)\exp(-i(t-\tau)\omega_{k_1})\exp(-i\tau(\omega_{k_1} - \omega_{k_2})) + R_1 \\ &= \frac{1}{2\pi T}\sum_{\tau=1}^{T}\exp(-i\tau(\omega_{k_1} - \omega_{k_2}))\sum_{h=-\tau}^{T-\tau}\kappa(\frac{\tau}{T}, h)\exp(-ih\omega_{k_1}) + R_1\end{aligned}$$

and by using Assumption 3.2(L2), we can show that $R_1 \leq \frac{C}{T}\sum_h |h|\cdot\kappa_2(h) = O(T^{-1})$. Next we replace the inner sum with $\sum_{h=-\infty}^{\infty}$ to give

$$\text{cov}\big(J_{k_1}, J_{k_2}\big) = \frac{1}{T}\sum_{\tau=1}^{T} f(\frac{\tau}{T}, \omega_{k_1})\exp(-i(k_1 - k_2)\omega_\tau) + R_1 + R_2,$$

where

$$R_2 = \frac{1}{2\pi T}\sum_{\tau=1}^{T}\exp(-i\tau(\omega_{k_1} - \omega_{k_2}))\Bigg(\sum_{h=-\infty}^{-\tau} + \sum_{T-\tau}^{\infty}\Bigg)\kappa(\frac{\tau}{T}, h)\exp(-ih\omega_{k_1}).$$

By using Corollary A.1, we have that $\sup_u |\kappa(u,h)| \leq C|h|^{-(2+\varepsilon)}$, therefore $|R_2| \leq CT^{-1}$. Finally, by replacing the sum by an integral, we get

$$\text{cov}\big(J_{k_1}, J_{k_2}\big) = \int_0^1 f(u, \omega_{k_1})\exp(-i(k_1 - k_2)\omega_\tau)du + R_1 + R_2 + R_3,$$

where $|R_1 + R_2 + R_3| \leq CT^{-1}$, which gives (A.18). □

In the following lemma and corollary, we show how the α-mixing rates are related to summability of the cumulants. We state the results for the multivariate case.

Lemma A.9 *Let us suppose that $\{\underline{X}_t\}$ is an α-mixing time series with rate $K|t|^{-\alpha}$ such that there exists an r where $\|\underline{X}_t\|_r < \infty$ and $\alpha > r(k-1)/(r-k)$. If $t_1 \leq t_2 \leq \ldots \leq t_k$, then we have* $|\text{cum}(X_{t_1,j_1}, \ldots, X_{t_k,j_k})| \leq C_k \sup_{t,T}\|X_{t,T}\|_r^k \prod_{i=2}^{k}|t_i - t_{i-1}|^{-\alpha(\frac{1-k/r}{k-1})}$,

$$\sup_{t_1}\sum_{t_2,\ldots,t_k=1}^{\infty}|\text{cum}(X_{t_1,j_1}, \ldots, X_{t_k,j_k})| \leq C_k \sup_{t,j}\|X_{t,j}\|_r^k\left(\sum_t |t|^{-\alpha(\frac{1-k/r}{k-1})}\right)^{k-1} < \infty, \tag{A.19}$$

and, for all $2 \leq j \leq k$, we have

$$\sup_{t_1} \sum_{t_2,\ldots,t_k=1}^{\infty} (1+|t_j|)|\mathrm{cum}(X_{t_1,j_1},\ldots,X_{t_k,j_k})| \leq C_k \sup_{t,j} \|X_{t,j}\|_r^k \left(\sum_t |t|^{-\alpha(\frac{1-k/r}{k-1})+1}\right)^{k-1} < \infty, \quad \text{(A.20)}$$

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

PROOF. The proof is identical to the proof of Lemma 4.1 in Lee and Subba Rao (2011) (see also Statulevicius and Jakimavicius (1988) and Neumann (1996)). □

Corollary A.1 *Suppose Assumption 3.1(P1, P2) or 3.2(L1, L3) holds. Then there exists an $\varepsilon > 0$ such that $|\mathrm{cov}(\underline{X}_0, \underline{X}_h)|_1 < C|h|^{-(2+\varepsilon)}$, $\sup_t |\mathrm{cov}(\underline{X}_{t,T}, \underline{X}_{t+h,T})|_1 < C|h|^{-(2+\varepsilon)}$ and*

$$\sup_{t_1,j_1,\ldots,j_4} \sum_{t_2,t_3,t_4} (1+|t_i|)\cdot|\mathrm{cum}(X_{t_1,j_1},X_{t_2,j_2},X_{t_3,j_3},X_{t_4,j_4})| < \infty, \quad i=1,2,3,$$

Furthermore, if Assumption 3.1(P1, P4) or 3.2(L1, L5) holds, then for $1 \leq n \leq 8$ we have

$$\sup_{t_1} \sum_{t_2,\ldots,t_n} |\mathrm{cum}(X_{t_1,j_1},X_{t_2,j_2},\ldots,X_{t_n,j_n})| < \infty.$$

PROOF. The proof immediately follows from Lemma A.9, thus we omit the details. □

Theorem A.1 *Suppose that Assumption 3.1 holds, then we have*

$$\sqrt{T}\widehat{c}(r,\ell) = \sqrt{T}\widetilde{c}(r,\ell) + O_p\left(\frac{1}{b\sqrt{T}} + b + b^2T^{1/2}\right). \quad \text{(A.21)}$$

Under Assumption 3.2, we have

$$\sqrt{T}\widehat{c}(r,\ell) = \sqrt{T}\widetilde{c}(r,\ell) + \sqrt{T}\mathcal{S}_T(r,\ell) + \sqrt{T}\mathcal{B}_T(r,\ell) + O_p\left(\frac{\log T}{b\sqrt{T}} + b\log T + b^2\sqrt{T}\right). \quad \text{(A.22)}$$

PROOF. To prove (A.21), we use the expansion (A.9) to give

$$\begin{aligned}
\sqrt{T}\big(\widehat{c}(r,\ell)-\widetilde{c}(r,\ell)\big) &= \underbrace{A_{1,1}}_{\text{Lemma } A.4(i)} + \underbrace{A_{1,2}}_{\text{Lemma } A.5(i)} + \underbrace{O_p(A_2)+O_p(B_2)}_{(A.16)} + \sqrt{T}\big(\underbrace{\mathcal{S}_T(r,\ell)+\mathcal{B}_T(r,\ell)}_{\text{Lemma } A.7(i)}\big) \\
&= O\left(\frac{1}{T^{1/2}} + \frac{1}{bT} + b + \frac{1}{b\sqrt{T}} + b^2\sqrt{T}\right)
\end{aligned}$$

To prove (A.22) we first note that by Lemma A.7(ii) we have $\|\mathcal{S}_T(r,\ell)\|_2 = O(T^{-1/2})$ and $|\mathcal{B}_T(r,\ell)| = O(b^{1/2})$, therefore we use expansion (A.9) to give

$$\begin{aligned}
\sqrt{T}\big(\widehat{c}(r,\ell)-\widetilde{c}(r,\ell)\big) - \sqrt{T}\Big(\big(\mathcal{S}_T(r,\ell)+\mathcal{B}_T(r,\ell)\Big) &= \underbrace{A_{1,1}}_{\text{Lemma } A.4(ii)} + \underbrace{A_{1,2}}_{\text{Lemma } A.5(ii)} + \underbrace{O_p(A_2)+O_p(B_2)}_{(A.16)} \\
&= O\left(\frac{\log T}{b\sqrt{T}} + b\log T + \frac{1}{b\sqrt{T}} + b^2\sqrt{T}\right).
\end{aligned}$$

This proves the result. □

Proof of Theorems 3.1 and 3.5 The proof of Theorems 3.1 and 3.5 follows immediately from Theorem A.1. □

A.3 Proof of Theorem 3.2 and Lemma 3.1

Throughout the proof, we will assume that T is sufficiently large, i.e. such that $0 < r < \frac{T}{2}$ and $0 \leq \ell < \frac{T}{2}$ hold. This avoids issues related to symmetry and periodicity of the DFTs. The proof relies on the following important lemma. We mention, that unlike the previous (and future) sections in the Appendix, we will prove the result for the multivariate case. This is because for the variance calculation there are subtle differences between the multivariate and univariate cases.

Lemma A.10 *Suppose that $\{\underline{X}_t\}$ is fourth order stationary such that $\sum_h |h| \cdot |\mathrm{cov}(\underline{X}_0, \underline{X}_h)|_1 < \infty$ and $\sum_{t_1,t_2,t_3}(1+|h_1|) \cdot |\mathrm{cum}(\underline{X}_0, \underline{X}_{t_1}, \underline{X}_{t_3}, \underline{X}_{t_3})|_1 < \infty$ (where $|\cdot|_1$ is taken pointwise over every cumulant combination). We mention that these conditions are satisfied under Assumption 3.1(P1,P2) (see Corollary A.1). Then, for all fixed $r_1, r_2 \in \mathbb{N}$ and $\ell_1, \ell_2 \in \mathbb{N}_0$ and all $j_1, j_2, j_3, j_4 \in \{1, \ldots, d\}$, we have*

$$
\begin{aligned}
T\mathrm{cov}\left(\widetilde{c}_{j_1,j_2}(r_1,\ell_1), \widetilde{c}_{j_3,j_4}(r_2,\ell_2)\right) &= \left\{\delta_{j_1j_3}\delta_{j_2j_4}\delta_{\ell_1\ell_2} + \delta_{j_1j_4}\delta_{j_2j_3}\delta_{\ell_1,-\ell_2}\right\}\delta_{r_1,r_2} \\
&\quad + \kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)\delta_{r_1,r_2} + O\left(\frac{1}{T}\right), \\
T\mathrm{cov}\left(\widetilde{c}_{j_1,j_2}(r_1,\ell_1), \overline{\widetilde{c}_{j_3,j_4}(r_2,\ell_2)}\right) &= O\left(\frac{1}{T}\right), \\
T\mathrm{cov}\left(\overline{\widetilde{c}_{j_1,j_2}(r_1,\ell_1)}, \widetilde{c}_{j_3,j_4}(r_2,\ell_2)\right) &= O\left(\frac{1}{T}\right), \\
T\mathrm{cov}\left(\overline{\widetilde{c}_{j_1,j_2}(r_1,\ell_1)}, \overline{\widetilde{c}_{j_3,j_4}(r_2,\ell_2)}\right) &= \left\{\delta_{j_1j_3}\delta_{j_2j_4}\delta_{\ell_1\ell_2} + \delta_{j_1j_4}\delta_{j_2j_3}\delta_{\ell_1,-\ell_2}\right\}\delta_{r_1,r_2} \\
&\quad + \kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)\delta_{r_1,r_2} + O\left(\frac{1}{T}\right),
\end{aligned}
$$

where $\delta_{jk} = 1$ if $j = k$ and $\delta_{jk} = 0$ otherwise.

PROOF. Straightforward calculations give

$$
\begin{aligned}
& T\mathrm{cov}\left(\widetilde{c}_{j_1,j_2}(r_1,\ell_1), \widetilde{c}_{j_3,j_4}(r_2,\ell_2)\right) \\
= \;& \frac{1}{T}\sum_{k_1,k_2=1}^{T}\sum_{s_1,s_2,s_3,s_4=1}^{d} L_{j_1s_1}(\omega_{k_1})\overline{L_{j_2s_2}(\omega_{k_1+r_1})L_{j_3s_3}(\omega_{k_2})}L_{j_4s_4}(\omega_{k_2+r_2}) \\
& \times \mathrm{cov}(J_{k_1,s_1}\overline{J_{k_1+r_1,s_2}}, J_{k_2,s_3}\overline{J_{k_2+r_2,s_4}})\exp(i\ell_1\omega_{k_1} - i\ell_2\omega_{k_2}).
\end{aligned}
$$

and by using the identity $cum(Z_1, \overline{Z_2}, \overline{Z_3}, Z_4) = \text{cov}(Z_1\overline{Z_2}, Z_3\overline{Z_4}) - E(Z_1\overline{Z_3})E(\overline{Z_2}Z_4) - E(Z_1Z_4)E(\overline{Z_2Z_3})$ for complex-valued and zero mean random variables Z_1, Z_2, Z_3, Z_4, we get

$$\begin{aligned}
& T\text{cov}(\widetilde{c}_{j_1,j_2}(r_1,\ell_1), \widetilde{c}_{j_3,j_4}(r_2,\ell_2)) \\
= & \frac{1}{T}\sum_{k_1,k_2=1}^{T}\sum_{s_1,s_2,s_3,s_4=1}^{d} L_{j_1s_1}(\omega_{k_1})\overline{L_{j_2s_2}(\omega_{k_1+r_1})L_{j_3s_3}(\omega_{k_2})}L_{j_4s_4}(\omega_{k_2+r_2}) \\
& \big\{\text{E}(J_{k_1,s_1}\overline{J_{k_2,s_3}})\text{E}(\overline{J_{k_1+r_1,s_2}}J_{k_2+r_2,s_4}) + \text{E}(J_{k_1,s_1}J_{k_2+r_2,s_4})\text{E}(\overline{J_{k_1+r_1,s_2}J_{k_2,s_3}}) \\
& + cum(J_{k_1,s_1}, \overline{J_{k_1+r_1,s_2}}, \overline{J_{k_2,s_3}}, J_{k_2+r_2,s_4})\big\}\exp(i\ell_1\omega_{k_1} - i\ell_2\omega_{k_2}) \\
=: & I + II + III.
\end{aligned}$$

Substituting the identity $\text{E}(J_{k_1,s_1}\overline{J_{k_2,s_2}}) = \frac{1}{T}\sum_{h=-T+1}^{T-1}\kappa_{s_1s_2}(h)\sum_{t=1}^{T-|h|}e^{it(\omega_{k_1}-\omega_{k_2})}$ into I and replacing the inner sum with $\sum_{t=1}^{T}e^{it(\omega_{k_1}-\omega_{k_2})}$ gives

$$\begin{aligned}
I = & \frac{1}{T}\sum_{k_1,k_2=1}^{T}\sum_{s_1,s_2,s_3,s_4=1}^{d} L_{j_1s_1}(\omega_{k_1})\overline{L_{j_2s_2}(\omega_{k_1+r_1})L_{j_3s_3}(\omega_{k_2})}L_{j_4s_4}(\omega_{k_2+r_2}) \\
& \times\left(\frac{1}{2\pi T}\sum_{h=-(T-1)}^{T-1}\kappa_{s_1s_3}(h)e^{-ih\omega_{k_2}}\left(\sum_{t=1}^{T}e^{-it(\omega_{k_1}-\omega_{k_2})} + O(h)\right)\right) \\
& \times\left(\frac{1}{2\pi T}\sum_{h=-(T-1)}^{T-1}\kappa_{s_2s_4}(h)e^{ih\omega_{k_2+r_2}}\left(\sum_{t=1}^{T}e^{it(\omega_{k_1+r_1}-\omega_{k_2+r_2})} + O(h)\right)\right)\exp(i\ell_1\omega_{k_1} - i\ell_2\omega_{k_2}),
\end{aligned}$$

where it is clear that the $O(h) = -\sum_{t=T-h+1}^{T}e^{-it(\omega_{k_1}-\omega_{k_2})}$ term is uniformly bounded over all frequencies and h. By using $\sum_h|h\kappa_{s_1s_2}(h)| < \infty$, we have

$$\begin{aligned}
I = & \frac{1}{T}\sum_{k_1,k_2=1}^{T}\left(\sum_{s_1,s_3=1}^{d}L_{j_1s_1}(\omega_{k_1})f_{s_1s_3}(\omega_{k_2})\overline{L_{j_3s_3}(\omega_{k_2})}\exp(i\ell_1\omega_{k_1})\right) \\
& \times\left(\sum_{s_2,s_4=1}^{d}\overline{L_{j_2s_2}(\omega_{k_1+r_1})f_{s_2s_4}(\omega_{k_2+r_2})}L_{j_4s_4}(\omega_{k_2+r_2})\exp(-i\ell_2\omega_{k_2})\right)\delta_{k_1k_2}\delta_{r_1r_2} + O\left(\frac{1}{T}\right) \\
= & \delta_{j_1j_3}\delta_{j_2j_4}\delta_{r_1r_2}\delta_{\ell_1,\ell_2} + O\left(\frac{1}{T}\right),
\end{aligned}$$

where $\mathbf{L}(\omega_k)\mathbf{f}(\omega_k)\overline{\mathbf{L}(\omega_k)}' = \mathbf{I}_d$ and $\frac{1}{T}\sum_{k=1}^{T}\exp(-i(\ell_1-\ell_2)\omega_k) = 1$ if $\ell_1 = \ell_2$ and zero otherwise have been used. Using similar arguments, we obtain

$$\begin{aligned}
II = & \frac{1}{T}\sum_{k_1,k_2=1}^{T}\left(\sum_{s_1,s_4=1}^{d}L_{j_1s_1}(\omega_{k_1})\overline{f_{s_1s_4}(\omega_{k_2+r_2})}L_{j_4s_4}(\omega_{k_2+r_2})\exp(i\ell_1\omega_{k_1})\right) \\
& \left(\sum_{s_2,s_3=1}^{d}\overline{L_{j_2s_2}(\omega_{k_1+r_1})}f_{s_2s_3}(\omega_{k_2})\overline{L_{j_3s_3}(\omega_{k_2})}\exp(-i\ell_2\omega_{k_2})\right)\delta_{k_1,-k_2-r_2}\delta_{k_1+r_1,-k_2} + O\left(\frac{1}{T}\right) \\
= & \delta_{j_1j_4}\delta_{j_3j_2}\delta_{\ell_1,-\ell_2}\delta_{r_1r_2} + O\left(\frac{1}{T}\right),
\end{aligned}$$

where $\exp(-i\ell_2\omega_{r_1}) \to 1$ as $T \to \infty$ and $\frac{1}{T}\sum_{k=1}^{T}\exp(-i(\ell_1+\ell_2)\omega_k) = 1$ if $\ell_1 = -\ell_2$ and zero otherwise have been used. Finally, by using Theorem 4.3.2, (Brillinger, 1981), we have

$$\begin{aligned}
III &= \frac{1}{T}\sum_{k_1,k_2=1}^{T}\sum_{s_1,s_2,s_3,s_4=1}^{d} L_{j_1s_1}(\omega_{k_1})\overline{L_{j_2s_2}(\omega_{k_1+r_1})L_{j_3s_3}(\omega_{k_2})}L_{j_4s_4}(\omega_{k_2+r_2})\exp(i\ell_1\omega_{k_1}-i\ell_2\omega_{k_2}) \\
&\quad \times\left(\frac{2\pi}{T^2}f_{4;s_1,s_2,s_3,s_4}(\omega_{k_1},-\omega_{k_1+r_1},-\omega_{k_2})\sum_{t=1}^{T}e^{it(-\omega_{r_1}+\omega_{r_2})}+O\left(\frac{1}{T}\right)\right) \\
&= \frac{2\pi}{T^2}\sum_{k_1,k_2=1}^{T}\sum_{s_1,s_2,s_3,s_4=1}^{d} L_{j_1s_1}(\omega_{k_1})\overline{L_{j_2s_2}(\omega_{k_1+r_1})L_{j_3s_3}(\omega_{k_2})}L_{j_4s_4}(\omega_{k_2+r_2})\exp(i\ell_1\omega_{k_1}-i\ell_2\omega_{k_2}) \\
&\quad \times f_{4;s_1,s_2,s_3,s_4}(\omega_{k_1},-\omega_{k_1+r_1},-\omega_{k_2})\delta_{r_1r_2}+O\left(\frac{1}{T}\right) \\
&= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{2\pi}\sum_{s_1,s_2,s_3,s_4=1}^{d} L_{j_1s_1}(\lambda_1)\overline{L_{j_2s_2}(\lambda_1)L_{j_3s_3}(\lambda_2)}L_{j_4s_4}(\lambda_2)\exp(i\ell_1\lambda_1-i\ell_2\lambda_2) \\
&\quad \times f_{4;s_1,s_2,s_3,s_4}(\lambda_1,-\lambda_1,-\lambda_2)d\lambda_1 d\lambda_2\delta_{r_1r_2}+O\left(\frac{1}{T}\right) \\
&= \kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)\delta_{r_1r_2}+O\left(\frac{1}{T}\right),
\end{aligned}$$

which concludes the proof of the first part. This result immediately implies the fourth part by taking into account that $\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)$ is real-valued by (A.23) below. In the computations for the second and the third part, a $\delta_{r_1,-r_2}$ crops up, which is always zero due to $r_1, r_2 \in \mathbb{N}$. □

PROOF of Theorem 3.2

To prove part (i), we consider the entries of $\widetilde{\mathbf{C}}_T(r,\ell)$

$$\mathrm{E}(\widetilde{c}_{j_1,j_2}(r,\ell)) = \frac{1}{T}\sum_{k=1}^{T}\sum_{s_1,s_2=1}^{d} L_{j_1,s_1}(\omega_k)\mathrm{E}\left(J_{k,s_1}\overline{J_{k+r,s_2}}\right)\overline{L_{j_2,s_2}(\omega_{k+r})}\exp(i\omega_k\ell)$$

and using Lemma A.8(i) yields $E\left(J_{k,s_1}\overline{J_{k+r,s_2}}\right) = O(\frac{1}{T})$ for $r \neq Tk$, $k \in \mathbb{Z}$, which gives the assertion. Part (ii) follows from $\Re Z = \frac{1}{2}(Z+\overline{Z})$, $\Im Z = \frac{1}{2i}(Z-\overline{Z})$ and Lemma A.10. □

Lemma A.11 *Under suitable assumptions,* $\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)$ *satisfies*

$$\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4) = \overline{\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)} = \kappa^{(\ell_2,\ell_1)}(j_3,j_4,j_1,j_2) = \kappa^{(-\ell_1,-\ell_2)}(j_2,j_1,j_4,j_3). \quad \text{(A.23)}$$

In particular, $\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)$ *is always real-valued. Furthermore,* (A.23) *causes the limits of* $\mathrm{var}\left(\sqrt{T}\mathrm{vec}\left(\Re\widetilde{\mathbf{C}}_T(r,0)\right)\right)$ *and* $\mathrm{var}\left(\sqrt{T}\mathrm{vec}\left(\Im\widetilde{\mathbf{C}}_T(r,0)\right)\right)$ *to be singular.*

PROOF. The first identity in (A.23) follows by substituting $\lambda_1 \to -\lambda_1$ and $\lambda_2 \to -\lambda_2$ in $\overline{\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)}$, $\overline{L_{js}(-\lambda)} = L_{js}(\lambda)$ and $\overline{f_{4;s_1,s_2,s_3,s_4}(-\lambda_1,\lambda_1,\lambda_2)} = f_{4;s_1,s_2,s_3,s_4}(\lambda_1,-\lambda_1,-\lambda_2)$.

The second follows from exchanging λ_1 and λ_2 in $\overline{\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)}$ and $\overline{f_{4;s_1,s_2,s_3,s_4}(\lambda_2,-\lambda_2,-\lambda_1)} = f_{4;s_3,s_4,s_1,s_2}(\lambda_1,-\lambda_1,-\lambda_2)$. The third identity follows by substituting $\lambda_1 \to -\lambda_1$ and $\lambda_2 \to -\lambda_2$ in $\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)$ and $f_{4;s_1,s_2,s_3,s_4}(-\lambda_1,\lambda_1,\lambda_2) = f_{4;s_2,s_1,s_4,s_3}(\lambda_1,-\lambda_1,-\lambda_2)$. The first identity immediately implies that $\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)$ is real-valued. To prove the second part of this lemma, we consider only the real part of $\widetilde{\mathbf{C}}_T(r,0)$ and we can assume wlog that $d=2$. From Lemma A.10, we get immediately

$$\begin{aligned}
&\operatorname{var}\left(\sqrt{T}vec\left(\Re\widetilde{\mathbf{C}}_T(r,0)\right)\right) \\
\to\ &\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \kappa^{(0,0)}(1,1,1,1) & \kappa^{(0,0)}(1,1,2,1) & \kappa^{(0,0)}(1,1,1,2) & \kappa^{(0,0)}(1,1,2,2) \\ \kappa^{(0,0)}(2,1,1,1) & \kappa^{(0,0)}(2,1,2,1) & \kappa^{(0,0)}(2,1,1,2) & \kappa^{(0,0)}(2,1,2,2) \\ \kappa^{(0,0)}(1,2,1,1) & \kappa^{(0,0)}(1,2,2,1) & \kappa^{(0,0)}(1,2,1,2) & \kappa^{(0,0)}(1,2,2,2) \\ \kappa^{(0,0)}(2,2,1,1) & \kappa^{(0,0)}(2,2,2,1) & \kappa^{(0,0)}(2,2,1,2) & \kappa^{(0,0)}(2,2,2,2) \end{pmatrix},
\end{aligned}$$

and due to (A.23), the second and third rows are equal leading to the singularity. $\square$

PROOF of Lemma 3.1 By using Lemma A.8(ii) (generalized to the multivariate setting) we have

$$\begin{aligned}
\mathrm{E}(\widetilde{\mathbf{C}}_T(r,\ell)) &= \frac{1}{T}\sum_{k=1}^{T}\mathbf{L}(\omega_k)\mathrm{E}(\underline{J}_T(\omega_k)\overline{\underline{J}_T(\omega_{k+r})}')\overline{\mathbf{L}(\omega_{k+r})}'\exp(i\ell\omega_k)) \\
&= \frac{1}{T}\sum_{k=1}^{T}\mathbf{L}(\omega_k)\left(\int_0^1 \mathbf{f}(u;\omega_k)\exp(-2\pi iru)du\right)\overline{\mathbf{L}(\omega_{k+r})}'\exp(i\ell\omega_k)) + O(\frac{1}{T})\ (\text{by } (A.18)) \\
&= \frac{1}{2\pi}\int_0^{2\pi}\mathbf{L}(\omega)\left(\int_0^1 \mathbf{f}(u;\omega)\exp(-2\pi iru)du\right)\overline{\mathbf{L}(\omega+\omega_r)}'\exp(i\ell\omega_k)) + O(\frac{1}{T}) \\
&= \frac{1}{2\pi}\int_0^{2\pi}\int_0^1 \mathbf{L}(\omega)\mathbf{f}(u;\omega)\overline{\mathbf{L}(\omega)}'\exp(i2\pi ru)\exp(i\ell\omega)dud\omega + O(\frac{1}{T}) \\
&= \boldsymbol{A}(r,\ell) + O(\frac{1}{T}).
\end{aligned}$$

Thus giving the required result. $\square$

PROOF of Lemma 3.2 The proof of (i) follows immediately from $\mathbf{L}(\omega)\mathbf{f}(u;\omega)\overline{\mathbf{L}(\omega)}' \in L_2(\mathbb{R}^{d^2})$.

To prove (ii), we note that if $\{\underline{X}_t\}$ is second order stationary, then $\mathbf{f}(u;\omega) = \mathbf{f}(\omega)$. Therefore, $\mathbf{L}(\omega)\mathbf{f}(u;\omega)\overline{\mathbf{L}(\omega)}' = I_d$ and $\mathbf{A}(r,\ell) = 0$ for all r and ℓ, except $\mathbf{A}(0,0) = I_d$. To prove the only if part, suppose $\mathbf{A}(r,\ell) = 0$ for all $r \neq 0$ and all $\ell \in \mathbb{Z}$ then $\sum_{r,\ell}\mathbf{A}(r,\ell)\exp(-2\pi iru)\exp(i\ell\omega)$ is only a function of ω, thus $\mathbf{f}(u;\omega)$ is only a function of ω which immediately implies that the underlying process is second order stationary.

To prove (iii) we use integration by parts. Under Assumption 3.2(L2, L4) the first derivative of $\mathbf{f}(u;\omega)$ exists with respect to u and the second derivative exists with respect to ω (moreover with respect to ω $\mathbf{L}(\omega)\mathbf{f}(u;\omega)\overline{\mathbf{L}(\omega)}'$) is a periodic continuous function. Therefore by integration by parts, twice with respect to ω and once with respect to u, we have

$$\begin{aligned}
\boldsymbol{A}(r,\ell) &= \frac{1}{2\pi}\int_0^{2\pi}\int_0^1 \boldsymbol{G}(u;\omega)\exp(i2\pi ru)\exp(i\ell\omega)dud\omega \\
&= \frac{1}{(i\ell)^2(2\pi r)}\int_0^{2\pi}\frac{\partial^2\boldsymbol{G}(u;\omega)}{\partial\omega^2}\exp(i2\pi ru)d\omega\Big]_{u=0}^{u=1} + \frac{1}{(i\ell)^2(2\pi r)}\int_0^1\int_0^{2\pi}\frac{\partial^3\boldsymbol{G}(u;\omega)}{\partial u\partial\omega^2}\exp(i2\pi ru)d\omega du,
\end{aligned}$$

where $\boldsymbol{G}(u;\omega) = \mathbf{L}(\omega)\mathbf{f}(u;\omega)\overline{\mathbf{L}(\omega)}'$. Taking absolutes of the above, we have $|\boldsymbol{A}(r,\ell)|_1 \leq K|\ell|^{-2}|r|^{-1}$ for some finite constant K.

To prove (iv), we note that

$$\begin{aligned}
\mathbf{A}(-r,-\ell) &= \frac{1}{2\pi}\int_0^{2\pi}\int_0^1 \mathbf{L}(\omega)\mathbf{f}(u;\omega)\overline{\mathbf{L}(\omega)}'\exp(-i2\pi ru)\exp(-i\ell\omega)dud\omega \\
&= \frac{1}{2\pi}\int_0^{2\pi}\int_0^1 \mathbf{L}(\omega)\mathbf{f}(u;-\omega)\overline{\mathbf{L}(\omega)}'\exp(-i2\pi ru)\exp(i\ell\omega)dud\omega \text{ (by a change of variables)} \\
&= \frac{1}{2\pi}\int_0^{2\pi}\int_0^1 \mathbf{L}(\omega)\mathbf{f}(u;\omega)'\overline{\mathbf{L}(\omega)}'\exp(-i2\pi ru)\exp(i\ell\omega)dud\omega \text{ (since } \mathbf{f}(u;\omega)' = \mathbf{f}(u;-\omega)) \\
&= \mathbf{A}(-r,\ell)'.
\end{aligned}$$

Thus we have proven the lemma. □

A.4 Proof of Theorems 3.3 and 3.6

The objective in this section is to prove asymptotic normality of $\widehat{\mathbf{C}}_T(r,\ell)$. We start by studying its approximation $\widetilde{c}_{j_1,j_2}(r,\ell)$, which we use to show normality. Expanding $\widetilde{c}_{j_1,j_2}(r,\ell)$ gives the quadratic form

$$\begin{aligned}
&\widetilde{c}_{j_1,j_2}(r,\ell) \\
&= \frac{1}{T}\sum_{k=1}^{T}\overline{\underline{J}_T(\omega_{k+r})}'\overline{\mathbf{L}_{j_2,\bullet}(\omega_{k+r})}'\mathbf{L}_{j_1,\bullet}(\omega_k)\underline{J}_T(\omega_k)\exp(i\ell\omega_k) \\
&= \frac{1}{2\pi T}\sum_{t,\tau=1}^{T}\underline{X}'_{t,T}\left(\frac{1}{T}\sum_{k=1}^{T}\overline{\mathbf{L}_{j_2,\bullet}(\omega_{k+r})}'\mathbf{L}_{j_1,\bullet}(\omega_k)\exp(i\omega_k(t-\tau+\ell))\right)\underline{X}_{\tau,T}\exp(-i\tau\omega_r) \qquad \text{(A.24)}
\end{aligned}$$

In Lemmas A.12 and A.13 we will show that the inner sum decays sufficiently fast over $(t-\tau+\ell)$ to allow us to use show asymptotic normality of $\widetilde{c}_{j_1,j_2}(r,\ell)$.

Lemma A.12 *Suppose $\mathbf{f}(\omega)$ is a non-singular matrix and the second derivatives of the elements of $\mathbf{f}(\omega)$ with respect to ω are bounded. Then we have*

$$\sup_{z\in[0,2\pi]}\left|\frac{\partial^2 L_\bullet(\underline{f}(\omega))L_\bullet(\underline{f}(\omega+z))}{\partial\omega^2}\right| < \infty \qquad \text{(A.25)}$$

and

$$\sup_{z,\underline{j}} |a_{\underline{j}}(s;z)| \leq \frac{C}{|s|^2} \quad \textit{and} \quad \sup_{\underline{j}} |d_{\underline{j}}(s;z)|_1 \leq \frac{C}{|s|^2} \quad \textit{for } s \neq 0,$$

where

$$\begin{aligned} a_{\underline{j}}(s;z) &= \frac{1}{2\pi}\int_0^{2\pi} L_{j_1,j_2}((\underline{f}(\omega))L_{j_3,j_4}(\underline{f}(\omega+z))\exp(is\omega)d\omega, \\ d_{\underline{j}}(s;z) &= \frac{1}{2\pi}\int_0^{2\pi} h_{j_1j_3}(r;\omega_k)\nabla_{\underline{f}(\omega),\underline{f}(\omega+z)}L_{j_1,j_2}(\underline{f}(\omega))L_{j_3,j_4}(\underline{f}(\omega+z))\exp(is\omega)d\omega \end{aligned} \quad \text{(A.26)}$$

and $h_{j_2j_4}(\omega;r) = \int_0^1 f_{j_2j_4}(u;\omega)\exp(2\pi iur)du$ *with a finite constant* C*.*

PROOF. Implicit differentiation gives

$$\begin{aligned} \frac{\partial L_{js}(\underline{f}(\omega))}{\partial\omega} &= \frac{\partial\underline{f}(\omega)'}{\partial\omega}\nabla_{\underline{f}}L_{js}(\underline{f}(\omega)) \quad \text{and} \\ \frac{\partial^2 L_{js}(\underline{f}(\omega))}{\partial\omega^2} &= \frac{\partial^2\underline{f}(\omega)'}{\partial\omega^2}\nabla_{\underline{f}}L_{j,s}(\underline{f}(\omega)) + \frac{\partial\underline{f}(\omega)'}{\partial\omega}\nabla^2_{\underline{f}}L_{j,s}(\underline{f}(\omega))\frac{\partial\underline{f}(\omega)}{\partial\omega}. \end{aligned} \quad \text{(A.27)}$$

By using Lemma A.1, we have that $\sup_\omega |\nabla_{\underline{f}}L_{j,s}(\underline{f}(\omega))| < \infty$ and $\sup_\omega \left|\nabla^2_{\underline{f}}L_{j,s}(\underline{f}(\omega))\right| < \infty$. Since $\sum_h h^2|\boldsymbol{\kappa}(h)|_1 < \infty$ (or equivalently in the nonstationary case, the integrated covariance satisfies this assumption), then we have $|\frac{\partial\mathbf{f}(\omega)}{\partial\omega}|_1 < \infty$ and $|\frac{\partial^2\mathbf{f}(\omega)}{\partial\omega^2}|_1 < \infty$. Substituting these bounds into (A.27) gives (A.25).

To prove $\sup_z |a_{\underline{j}}(s;z)| \leq C|s|^{-2}$ (for $s \neq 0$), we use (A.25) and apply integration by parts twice to $a_{\underline{j}}(s;z)$ to obtain the bound (similar to the proof of Lemma 3.1, in Appendix A.3). We use the same method to obtain $\sup_z |d_{\underline{j}}(s;z)| \leq C|s|^{-2}$ (for $s \neq 0$). □

Lemma A.13 *Suppose that either Assumption 3.1(P1, P3, P4) or Assumption 3.2 (L1, L2, L3) holds (in the stationary case we let* $X_t = X_{t,T}$*). Then we have*

$$\widetilde{c}_{j_1,j_2}(r,\ell) = \frac{1}{2\pi T}\sum_{t,\tau=1}^{T} \underline{X}'_{t,T}G_{\omega_r}(t-\tau+\ell)\underline{X}_{\tau,T}\exp(-i\tau\omega_r) + O_p\left(\frac{1}{T}\right),$$

where $G_{\omega_r}(k) = \int_0^1 \overline{\mathbf{L}_{j_2,\bullet}(\omega+\omega_r)}'\mathbf{L}_{j_1,\bullet}(\omega)\exp(i\omega k)d\omega = \sum_{s_1,s_2=1}^{d} a_{j_1,s_1,j_2,s_2}(k;\omega_k)$*,* $|G_{\omega_r}(k)| \leq C/|k|^2$*.*

PROOF. We replace $\left(\frac{1}{T}\sum_{k=1}^{T}\overline{\mathbf{L}_{j_2,\bullet}(\omega_{k+r})}'\mathbf{L}_{j_1,\bullet}(\omega_k)\exp(i\omega_k(t-\tau+\ell))\right)$ in (A.24) with its integral limit $G_{\omega_r}(t-\tau+\ell)$, and by using (A.26), we obtain bounds on $G_{\omega_r}(s)$. This gives the required result. □

Theorem A.2 *Suppose that $\{\underline{X}\}_t$ satisfies Assumption 3.1(P1-P3). Then for all fixed $r \in \mathbb{N}$ and $\ell \in \mathbb{Z}$, we have*

$$\sqrt{T} vech\left(\Re \widetilde{\mathbf{C}}_T(r,\ell)\right) \xrightarrow{D} \mathcal{N}\left(\underline{0}_{d(d+1)/2}, \mathbf{W}_{\ell,\ell}\right) \quad and \quad \sqrt{T} vech\left(\Im \widetilde{\mathbf{C}}_T(r,\ell)\right) \xrightarrow{D} \mathcal{N}\left(\underline{0}_{d(d+1)/2}, \mathbf{W}_{\ell,\ell}\right), \tag{A.28}$$

where $\underline{0}_{d(d+1)/2}$ is the $d(d+1)/2$ zero vector and

$$\mathbf{W}_{\ell,\ell} = \mathbf{W}^{(1)}_{\ell,\ell} + \mathbf{W}^{(2)}_{\ell,\ell} \quad \textit{(defined in (2.11) and (2.15))}.$$

PROOF. Since each element of $\widetilde{\mathbf{C}}_T(r,\ell)$ can be approximated by the quadratic form given in Lemma A.13, to show asymptotic normality of $\widetilde{\mathbf{C}}_T(r,\ell)$, we use a central limit theorem for quadratic forms. One such central limit theorems is given in Lee and Subba Rao (2011), Corollary 2.2 (which holds for both stationary and nonstationary time series). Assumption 3.1(P1-P3) implies the conditions in Lee and Subba Rao (2011), Corollary 2.2, therefore by using Cramer-Wold device we have asymptotic normality of $\widetilde{\mathbf{C}}_T(r,\ell)$. □

PROOF of Theorem 3.3 Since $\sqrt{T}\widehat{\mathbf{C}}_T(r,\ell) = \sqrt{T}\widetilde{\mathbf{C}}_T(r,\ell) + o_p(1)$ to show asymptotic normality of $\sqrt{T}\widehat{\mathbf{C}}_T(r,\ell)$, we simply need to show asymptotic normality of $\sqrt{T}\widetilde{\mathbf{C}}_T(r,\ell)$. Now by applying identical methods to the proof of Theorem A.2, we obtain the result. □

PROOF of Theorem 3.4 Follows immediately from Theorem 3.3.

We now derive the distribution of $\widehat{\mathbf{C}}(r,\ell)$ under the assumption of local stationarity. We recall from Theorem 3.5 that the distribution of $\widehat{\mathbf{C}}_T(r,\ell)$ is determined by $\widetilde{\mathbf{C}}_T(r,\ell)$ and $\mathcal{S}_T(r,\ell)$. We have shown in Lemma A.13 that $\widetilde{\mathbf{C}}_T(r,\ell)$ can be approximated by a quadratic form. We now show that $\mathcal{S}_T(r,\ell)$ is also a quadratic form. Substituting the quadratic form expansion

$$\underline{f}_{k,r} - \mathrm{E}(\underline{f}_{k,r}) = \frac{1}{2\pi T}\sum_{t,\tau=1}^{T} \lambda_b(t-\tau)\underline{g}(\underline{X}_t \underline{X}_\tau')\exp(i(t-\tau)\omega_k)$$

into $\mathcal{S}_{T,j_1,j_2}(r,\ell)$ (defined in (A.7)) gives

$$\begin{aligned}
&\mathcal{S}_{T,j_1,j_2}(r,\ell)\\
=\ & \frac{1}{T}\sum_{k=1}^{T}\sum_{s_1,s_2=1}^{d} \underbrace{\mathrm{E}(J_{k,s_1}\overline{J_{k+r,s_2}})}_{h_{s_1 s_2}(r;\omega_k)} \left(\widehat{\underline{f}}_{k,r} - \mathrm{E}(\widehat{\underline{f}}_{k,r})\right)' \nabla A_{j_1,s_1,j_2,s_2}(\underline{f}_{k,r}) e^{i\ell\omega_k}\\
=\ & \sum_{s_1,s_2=1}^{d} \frac{1}{2\pi T}\sum_{t,\tau=1}^{T} \lambda_b(t-\tau)\underline{g}(\underline{X}_t\underline{X}_\tau')' \underbrace{\frac{1}{T}\sum_{k=1}^{T}\exp(i(t-\tau+\ell)\omega_k) h_{s_1 s_2}(\omega_k;r)\nabla A_{j_1,s_1,j_2,s_2}(\underline{f}_{k,r})}_{\approx \underline{d}_{j_1,s_1,j_2,s_2,\omega_r}(t-\tau+\ell) \text{ by } (A.26)}\\
=\ & \sum_{s_1,s_2=1}^{d} \frac{1}{2\pi T}\sum_{t,\tau=1}^{T} \lambda_b(t-\tau)\underline{g}(\underline{X}_t\underline{X}_\tau')'\underline{d}_{j_1,s_1,j_2,s_2,\omega_r}(t-\tau+\ell) + O\left(\frac{1}{T}\right)
\end{aligned} \tag{A.29}$$

where the random vector $\underline{g}(\underline{X}_t\underline{X}_\tau')$ is defined as

$$\underline{g}(\underline{X}_t\underline{X}_\tau') = \begin{pmatrix} \mathrm{vech}(\underline{X}_t\underline{X}_\tau') \\ \mathrm{vech}(\underline{X}_t\underline{X}_\tau')\exp(i(t-\tau)\omega_r) \end{pmatrix} - \mathrm{E}\begin{pmatrix} \mathrm{vech}(\underline{X}_t\underline{X}_\tau') \\ \mathrm{vech}(\underline{X}_t\underline{X}_\tau')\exp(i(t-\tau)\omega_r) \end{pmatrix}$$

and $\underline{d}_{j_1,s_1,j_2,s_2,\omega_r}(t-\tau-\ell)$ is defined in (A.26).

PROOF of Theorem 3.6 Theorem 3.5 implies that

$$\widehat{c}_{j_1,j_2}(r,\ell) - \mathcal{B}_{T,j_1j_2}(r,\ell) = \widetilde{c}_{j_1,j_2}(r,\ell) + \mathcal{S}_{T,j_1,j_2}(r,\ell) + o_p\left(\frac{1}{\sqrt{T}}\right).$$

By using Lemma A.13 and (A.29), we have that $\widetilde{c}_{j_1,j_2}(r,\ell) + \mathcal{S}_{T,j_1,j_2}(r,\ell)$ is a quadratic form. Therefore, by applying Lee and Subba Rao (2011), Corollary 2.2 to $\widetilde{c}_{j_1,j_2}(r,\ell) + \mathcal{S}_{T,j_1,j_2}(r,\ell)$, we can prove (3.12). □

A.5 Proof of results in Section 4

PROOF of Lemma 4.1. We first prove (i). Politis and Romano (1994) have shown that the stationary bootstrap leads to a bootstrap sample which is stationary with respect to the observations $\{X_t\}_{t=1}^T$. Therefore, by using the same argument, as those used to prove Lemma 1, Politis and Romano (1994), and conditioning on the block length for $0 < t_1 < t_2 \ldots < t_n$, we have

$$\begin{aligned} \mathrm{cum}^*(X_{t_1}^*, X_{t_2}\ldots, X_{t_n}^*) &= \mathrm{cum}^*(X_{t_1}^*, X_{t_2}^*, \ldots, X_{t_n}^*|L<|t_n|)P(L<|t_n|) \\ &\quad + \mathrm{cum}^*(X_{t_1}^*, X_{t_2}^*, \ldots, X_{t_n}^*|L\geq|t_n|)P(L\geq|t_n|). \end{aligned}$$

We observe that $\mathrm{cum}^*(X_{t_1}^*, \ldots, X_{t_n}^*|L<|t_n|) = 0$ (since the random variables in separate blocks are conditionally independent), $\mathrm{cum}^*(X_{t_1}^*, \ldots, X_{t_n}^*|L\geq|t_n|) = \widehat{\kappa}^C(t_2-t_1, \ldots, t_n-t_1)$ and $P(L\geq |t_n|) = (1-p)^{|t_n|}$. Thus altogether, we have $\mathrm{cum}^*(X_{t_1}^*, \ldots, X_{t_n}^*) = (1-p)^{|t_n|}\widehat{\kappa}^C(t_2-t_1, \ldots, t_n-t_1)$.

We now prove (ii). We first bound the difference $\widehat{\mu}_n^C(h_1, \ldots, h_{n-1}) - \widehat{\mu}_n(h_1, \ldots, h_{n-1})$. Without loss of generality, we consider the case $1 \leq h_1 \leq h_2 \cdots \leq h_{n-1} < T$. Comparing $\widehat{\mu}_n^C$ with $\widehat{\mu}_n$, we observe that the only difference is that $\widehat{\mu}_n^C$ contains a few additional terms due to Y_t for $t > T$, therefore

$$\widehat{\mu}_n^C(h_1, \ldots, h_{n-1}) - \widehat{\mu}_n(h_1, \ldots, h_{n-1}) = \frac{1}{T}\sum_{t=T-h_{n-1}+1}^{T} Y_t \prod_{i=1}^{n-1} Y_{t+h_i}.$$

Since $Y_t = X_{t\bmod T}$, we have

$$\left\|\widehat{\mu}_n^C(h_1, \ldots, h_{n-1}) - \widehat{\mu}_n(h_1, \ldots, h_{n-1})\right\|_{q/n} \leq \frac{|h_{n-1}|}{T}\sup_{t,T}\|X_{t,T}\|_q^n$$

and substituting this bound into (4.2) gives (ii).

We partition the proof of (iii) in two stages. First, we derive the sampling properties of the sample moments and using these results we derive the sampling properties of the sample cumulants. Assume $0 \leq h_1 \leq \ldots \leq h_{n-1}$ and define the product $Z_t = X_t \prod_{i=1}^{n-1} X_{t+h_i}$, then by using Ibragimov's inequality, we have $\|\mathrm{E}(Z_t|\mathcal{F}_{t-i}) - \mathrm{E}(Z_t|\mathcal{F}_{t-i-1})\|_m \leq C\|Z_t\|_r |i|^{-\alpha(\frac{1}{m}-\frac{1}{r})}$. Let $M_i(t) = \mathrm{E}(Z_t|\mathcal{F}_{t-i}) - \mathrm{E}(Z_t|\mathcal{F}_{t-i-1})$, then $Z_t - \mathrm{E}(Z_t) = \sum_i M_i(t)$. Using the above and Burkhölder's inequality (in the case that $m \geq 2$), we obtain the bound

$$
\begin{aligned}
& \|\widehat{\mu}_n(h_1,\ldots,h_{n-1}) - \mathrm{E}(\widehat{\mu}_n(h_1,\ldots,h_{n-1}))\|_m \\
\leq & \Big\| \frac{1}{T}\sum_{t=1}^{T}\Big(X_t \prod_{j=1}^{n-1} X_{t+h_j} - \mathrm{E}(X_t \prod_{j=1}^{n-1} X_{t+h_j})\Big)\Big\|_m \\
= & \Big\|\frac{1}{T}\sum_{t=1}^{T}(Z_t - \mathrm{E}(Z_t))\Big\|_m \leq \Big\|\frac{1}{T}\sum_{t=1}^{T}\sum_i M_i(t)\Big\|_m \\
\leq & \frac{1}{T}\sum_i \Big\|\sum_{t=1}^{T} M_i(t)\Big\|_m \leq \frac{1}{T}\sum_i \Big(\sum_{t=1}^{T}\|M_i(t)\|_m^2\Big)^{1/2} \leq \frac{C}{\sqrt{T}} \underbrace{\sum_i |i|^{-\alpha(\frac{1}{m}-\frac{1}{r})}}_{<\infty \text{ if } \alpha(m^{-1}-r^{-1})>1},
\end{aligned}
$$

which is finite if for some $r > m\alpha/(\alpha - m)$ we have $\sup_t \|Y_t\|_r < \infty$. Now we write this in terms of moments of X_t. Since $\sup_t \|Y_t\|_r \leq (\sup_t \|X_t\|_{rn})^n$, if $\alpha > m$ and $\|X_t\|_r < \infty$ where $r > nm\alpha/(\alpha - m)$, then $\|\widehat{\mu}_n(h_1,\ldots,h_{n-1}) - \mathrm{E}(\widehat{\mu}_n(h_1,\ldots,h_{n-1}))\|_m = O(T^{-1/2})$. As the sample cumulant is the product of sample moments, we use the above to bound the difference in the product of sample moments:

$$
\begin{aligned}
& \Big\|\prod_{k=1}^{m}\widehat{\mu}_{d_k}(h_{k,1},\ldots,h_{k,d_k}) - \prod_{k=1}^{m}\mathrm{E}(\widehat{\mu}_{d_k}(h_{k,1},\ldots,h_{k,d_k})\Big\|_{q/n} \\
\leq & \sum_{j=1}^{m}\Big\|\widehat{\mu}_{d_j}(h_{j,1},\ldots,h_{j,d_j}) - \mathrm{E}(\widehat{\mu}_{d_j}(h_{j,1},\ldots,h_{j,d_j}))\Big\|_{qD_j/(nd_j)} \\
& \times\Big(\prod_{k=1}^{j-1}\|\hat{\mu}_{d_k}(h_{k,1},\ldots,h_{k,d_k})\|_{qD_j/(nd_k)}\Big)\prod_{k=j+1}^{m}\mu_{d_k}(h_{k,1},\ldots,h_{k,d_k}),
\end{aligned}
$$

where $D_j = \sum_{k=1}^{j} d_k$ (sum of all the sample moment orders). Applying the previous discussion to this situation, we see that if the mixing rate is such that $\alpha > \frac{qD_j}{nd_j}$ and the moment bound satisfies

$$
r > \frac{d_j\alpha\frac{qD_j}{nd_j}}{\alpha - \frac{qD_j}{nd_j}}, \tag{A.30}
$$

for all j, then

$$
\Big\|\prod_{k=1}^{m}\widehat{\mu}_{d_k}(h_{k,1},\ldots,h_{k,d_k}) - \prod_{k=1}^{m}\mathrm{E}(\widehat{\mu}_{d_k}(h_{k,1},\ldots,h_{k,d_k})\Big\|_{q/n} = O(T^{-1/2}).
$$

We use this to bound

$$\|\widehat{\kappa}_n(h_1,\ldots,h_{n-1}) - \widetilde{\kappa}_n(h_1,\ldots,h_{n-1})\|_{q/n} \leq \sum_{\pi}(|\pi|-1)!\Big\|\prod_{B\in\pi}\widehat{\mu}_n(\pi_{i\in B}) - \prod_{B\in\pi}\mathrm{E}(\widehat{\mu}_n(\pi_{i\in B}))\Big\|_{q/n}.$$

In order to show that the above difference $O(T^{-1/2})$ we will use (A.30) to obtain sufficient mixing and moment conditions. To get the minimum mixing rate α, we consider the case $D=n$ and $d_k=1$, this correspond to $\alpha > q$. To get the minimum moment rate, we consider the case $D=n$ and $d_k=n$, this gives $r > q\alpha/(\alpha-\frac{q}{n})$. Therefore, under these conditions we have $\|\widehat{\kappa}_n(h_1,\ldots,h_{n-1})) - \widetilde{\kappa}_n(h_1,\ldots,h_{n-1}))\|_q = O(\frac{1}{T^{1/2}})$. This proves (4.6)

We now prove (4.7). It is straightforward to show that if $h_{k,1}=0$ and $0\leq h_{k,2}\leq\ldots\leq h_{k,d_k}\leq T$, then we have

$$\Big|\frac{1}{T}\sum_{t=1}^{T-h_{k,d_k}}\mathrm{E}(X_tX_{t+h_{k,2}}\ldots X_{t+h_{k,d_k}}) - \frac{1}{T}\sum_{t=1}^{T}\mathrm{E}(X_tX_{t+h_{k,2}}\ldots X_{t+h_{k,d_k}})\Big| \leq C\frac{h_{k,d_k}}{T}.$$

Using this and the same methods as above we have (4.7) thus we have shown (iii).

To prove (iv), we note that it is immediately clear that $\bar{k}_n$ is the nth order cumulant of a stationary time series. However, in the case that the time series is nonstationary, the story is different. To prove (iva), we note that under the assumption that $\mathrm{E}(X_t)$ is constant for all t, we have

$$\begin{aligned}
\bar{\kappa}_2(h) &= \frac{1}{T}\sum_{t=1}^{T}\mathrm{E}(X_tX_{t+h}) - \big(\frac{1}{T}\sum_{t=1}^{T}\mathrm{E}(X_t)\big)^2 \\
&= \frac{1}{T}\sum_{t=1}^{T}\mathrm{E}\big((X_t-\mu)(X_{t+h}-\mu)\big) \quad (\text{using } \mathrm{E}(X_t)=\mu) \\
&= \frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_t, X_{t+h}).
\end{aligned}$$

To prove (ivb), we note that by using the same argument as above, we have

$$\begin{aligned}
\bar{\kappa}_3(h_1,h_2) &= \frac{1}{T}\sum_{t=1}^{T}\mathrm{E}(X_tX_{t+h_1}X_{t+h_2}) - \big(\frac{1}{T}\sum_{t=1}^{T}\big[\mathrm{E}(X_tX_{t+h_1}) + \mathrm{E}(X_{t+h_1}X_{t+h_2}) + \mathrm{E}(X_tX_{t+h_2})\big]\big)\mu + 2\mu^3 \\
&= \frac{1}{T}\sum_{t=1}^{T}\mathrm{E}\big((X_t-\mu)(X_{t+h_1}-\mu)(X_{t+h_2}-\mu)\big) = \frac{1}{T}\sum_{t=1}^{T}\mathrm{cum}(X_t, X_{t+h_1}, X_{t+h_2}),
\end{aligned}$$

which proves (ivb).

So far, the above results are the average cumulants. However, this pattern does not continue for $n \geq 4$. We observe that

$$\kappa_4(h_1,h_2,h_3) = \frac{1}{T}\sum_{t=1}^{T}\mathrm{E}\big[(X_t-\mu)(X_{t+h_1}-\mu)(X_{t+h_2}-\mu)(X_{t+h_2}-\mu)\big] -$$
$$\left(\frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_t,X_{t+h_1})\right)\left(\frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_{t+h_2},X_{t+h_3})\right) - \left(\frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_t,X_{t+h_2})\right)\times$$
$$\left(\frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_{t+h_1},X_{t+h_3})\right) - \left(\frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_t,X_{t+h_3})\right)\left(\frac{1}{T}\sum_{t=1}^{T}\mathrm{cov}(X_{t+h_1},X_{t+h_2})\right),$$

which cannot be written as the average of the fourth order cumulant. However, it is straightforward to show that the above can be written as the average of the fourth order cumulants plus the additional average covariances. This proves (ivc). The proof of (ivd) is similar and we omit the details. □

PROOF of Lemma 4.2. To prove (i), we use the triangle inequality to obtain

$$\left|\widehat{h}_n(\omega_1,\ldots,\omega_{n-1}) - \bar{f}_{n,T}(\omega_1,\ldots,\omega_{n-1})\right| \leq I + II,$$

where

$$I = \frac{1}{(2\pi)^{n-1}}\sum_{r_1,\ldots,r_{n-1}=-T}^{T}(1-p)^{\max(r_i,0)-\min(r_i,0)}\left|\widehat{\kappa}_n(r_1,\ldots,r_{n-1}) - \widetilde{\kappa}_n(r_1,\ldots,r_{n-1})\right|,$$
$$II = \frac{1}{(2\pi)^{n-1}}\sum_{r_1,\ldots,r_{n-1}=-T}^{T}(1-p)^{\max(r_i,0)-\min(r_i,0)}\left|\widetilde{\kappa}_n(r_1,\ldots,r_{n-1}) - \overline{\kappa}_n(r_1,\ldots,r_{n-1})\right|.$$

Term I can be split into two main cases:

(i) If $0 \leq r_1 \leq \ldots \leq r_{n-1}$ we have

$$\sum_{0\leq r_1\leq r_2\leq\ldots\leq r_{n-1}\leq T}(1-p)^{\max(r_{n-1},0)-\min(r_1,0)} \leq \sum_{r=1}^{T} r^{n-2}(1-p)^r \leq \frac{1}{p^{n-1}}$$

(ii) If $r_1 < 0$ but $r_{n-1} > 0$ we have

$$\sum_{\substack{-T\leq r_1\leq r_2\leq\cdots\leq r_{n-1}\leq T\\ r_1<0,r_{n-1}>0}}(1-p)^{\max(r_{n-1},0)-\min(r_1,0)} \leq \frac{1}{p^{n-1}}.$$

Altogether this gives

$$\sum_{r_1,\ldots,r_{n-1}=-T}^{T}(1-p)^{\max(r_i,0)-\min(r_i,0)} \leq Cp^{-(n-1)}. \tag{A.31}$$

Therefore, by using (4.6) and (A.31), we have

$$
\begin{aligned}
\|I\|_2 &\leq \frac{1}{(2\pi)^{n-1}} \sum_{r_1,\ldots,r_{n-1}=-T}^{T} (1-p)^{\max(r_i,0)-\min(r_i,0)} \underbrace{\left\|\widehat{\kappa}_n(r_1,\ldots,r_{n-1}) - \widetilde{\kappa}_n(r_1,\ldots,r_{n-1})\right\|_2}_{O(T^{-1/2}) \text{ (uniform in } r_i\text{) by eq. (4.6)}} \\
&= O\left(\frac{1}{T^{1/2}p^{(n-1)}}\right) \qquad \text{(by eq. (A.31))}.
\end{aligned}
$$

To bound $|II|$, we use (4.7) to give

$$
|II| \leq \frac{C}{T} \sum_{r_1,\ldots,r_{n-1}=-T}^{T} (1-p)^{\max(r_i,0)-\min(r_i,0)} (\max(r_i,0) - \min(r_i,0)) = O(\frac{1}{Tp^n}),
$$

where the final bound on the right hand side of the above is deduced using the same arguments used to bound (A.31). This proves (i).

To prove (ii), we note that

$$
\begin{aligned}
& \left|\widehat{h}_n(\omega_1,\ldots,\omega_{n-1}) - f_n(\omega_1,\ldots,\omega_{n-1})\right| \\
\leq\ & \left|\widehat{h}_n(\omega_1,\ldots,\omega_{n-1}) - \bar{f}_{n,T}(\omega_1,\ldots,\omega_{n-1})\right| + |\bar{f}_{n,T}(\omega_1,\ldots,\omega_{n-1}) - f_n(\omega_1,\ldots,\omega_{n-1})|.
\end{aligned}
$$

A bound for the first term on the right hand side of the above is given in part (i). To bound the second term, we note that $\bar{\kappa}_n(\cdot) = \kappa_n(\cdot)$ (where $\kappa_n(\cdot)$ are the cumulants of a nth order stationary time series). Furthermore, we have the inequality

$$
|\bar{f}_{n,T}(\omega_1,\ldots,\omega_{n-1}) - f_n(\omega_1,\ldots,\omega_{n-1})| \leq III + IV
$$

where

$$
\begin{aligned}
III &= \frac{1}{(2\pi)^{n-1}} \sum_{r_1,\ldots,r_{n-1}=-T}^{T} \left|(1-p)^{\max(r_i,0)-\min(r_i,0)} - 1\right| \cdot \left|\kappa_n(r_1,\ldots,r_{n-1})\right|, \\
IV &= \frac{1}{(2\pi)^{n-1}} \sum_{|r_1|,\ldots \text{ or } ,\ldots,|r_{n-1}|>T} |\kappa_n(r_1,\ldots,r_{n-1})|.
\end{aligned}
$$

Substituting the bound $|1-(1-p)^l| \leq Klp$, into III gives $|III| = O(p)$. Finally, by using similar arguments to those used in Brillinger (1981), Theorem 4.3.2 (based on the absolute summability of the kth order cumulant), we have $IV = O(\frac{1}{T})$. Altogether this gives (ii).

We now prove (iii). In the case that $n \in \{2,3\}$, the proof is identical to the stationary case since $\hat{h}_2$ an $\hat{h}_3$ are estimators of $f_{2,T}$ and $f_{3,T}$, which are the Fourier transforms of average covariances and average cumulants. Since the second and third order covariances decay at a sufficiently fast rate, $f_{2,T}$ and $f_{3,T}$ are finite. This proves (iiia)

On the other hand, we will prove that for $n \geq 4$, $\bar{f}_{n,T}$ depends on p. We prove the result for $n = 4$ (the result for the higher order cases follow similarly). Lemma A.9 implies that

$\sum_h |\frac{1}{T}\sum_{t=1}\mathrm{cov}(X_t, X_{t+h})| < \infty$ and $\sum_{h_1,h_2,h_3} |\frac{1}{T}\sum_{t=1}^{T}\mathrm{cum}(X_t, X_{t+h_1}, X_{t+h_2}, X_{t+h_3})| < \infty$. Therefore

$$\begin{aligned}
\sup_{\omega_1,\omega_2,\omega_3} |\bar{f}_{4,T}(\omega_1,\omega_2,\omega_3)| &\leq \frac{1}{(2\pi)^3}\sum_{h_1,h_2,h_3=-T}^{T}(1-p)^{\max(r_i,0)-\min(r_i,0)}|\overline{\kappa}_4(h_1,h_2,h_3)| \\
&\leq C\sum_{h=1}^{T}(1-p)^h = O(p^{-1}), \text{ (using the definition of } \overline{\kappa}_4(\cdot) \text{ in (4.8))},
\end{aligned}$$

where C is a finite constant (this proves (iiic)). The proof for the bound of the higher order $\bar{f}_{n,T}$ is similar. Thus we have shown (iii). □

PROOF of Theorem 4.1. Substituting Lemma 4.1(i) into $\mathrm{cum}^*(J_T^*(\omega_{k_1}),\ldots,J_{T,}^*(\omega_{k_n}))$ gives

$$\begin{aligned}
&\mathrm{cum}^*(J_T^*(\omega_{k_1}),\ldots,J_T^*(\omega_{k_n})) \\
&= \frac{1}{(2\pi T)^{n/2}}\sum_{t_1,\ldots,t_n=1}^{T}(1-p)^{\max_i((t_i-t_1),0)-\min_i((t_i-t_1),0)}\widehat{\kappa}_n^C(t_2-t_1,\ldots,t_n-t_1)e^{-it_1\omega_{k_1}-\ldots-it_n\omega_{k_n}}.
\end{aligned}$$

Let $r_i = t_i - t_1$

$$\begin{aligned}
&\mathrm{cum}^*(J_T^*(\omega_{k_1}),\ldots,J_T^*(\omega_{k_n})) \\
&= \frac{1}{(2\pi T)^{n/2}}\sum_{r_2,\ldots,r_n=-T+1}^{T-1}(1-p)^{g(\underline{r})}\widehat{\kappa}_n^C(r_2,\ldots,r_n)e^{-ir_2\omega_{k_2}-\ldots-ir_n\omega_{k_n}}\sum_{t=|\min_i(r_i,0)|+1}^{T-|\max_i(r_i,0)|}e^{-it(\omega_{k_1}+\omega_{k_2}+\ldots+\omega_{k_n})},
\end{aligned}$$

where $g(\underline{r}) = \max_i(r_i,0) - \min_i(r_i,0)$. Using that $\|\widehat{\kappa}_n^C(r_2,\ldots,r_n)\|_1 < \infty$, it is clear from the above that $\|\mathrm{cum}^*(J_T^*(\omega_{k_1}),\ldots,J_T^*(\omega_{k_n}))\|_1 = O(\frac{1}{T^{n/2-1}p^{n-1}})$, which proves (4.13). However, this is a crude bound and below we obtain more precise bounds (under stronger conditions).

Let

$$e(|\min_i(r_i,0)|,|\max_i(r_i,0)|) = \sum_{t=|\min_i(r_i,0)|+1}^{T-|\max_i(r_i,0)|}e^{-it(\omega_{k_1}+\omega_{k_2}+\ldots+\omega_{k_n})}.$$

Replacing $\widehat{\kappa}_n^C(r_2,\ldots,r_n)$ in the above with $\widehat{\kappa}_n(r_2,\ldots,r_n)$ gives

$$\begin{aligned}
&\mathrm{cum}^*(J_T^*(\omega_{k_1}),\ldots,J_T^*(\omega_{k_n})) \\
&= \frac{1}{(2\pi T)^{n/2}}\sum_{r_2,\ldots,r_n=-T+1}^{T-1}(1-p)^{g(\underline{r})}\widehat{\kappa}_n(r_2,\ldots,r_n)e^{-ir_2\omega_{k_2}-\ldots-ir_n\omega_{k_n}}e(|\min_i(r_i,0)|,|\max_i(r_i,0)|) + R_1,
\end{aligned}$$

where

$$R_1 = \frac{1}{(2\pi T)^{n/2}}\sum_{r_2,\ldots,r_n=-T+1}^{T-1}(1-p)^{g(\underline{r})}\Big(\widehat{\kappa}_n^C(r_2,\ldots,r_n)-\widehat{\kappa}_n(r_2,\ldots,r_n)\Big)e(|\min_i(r_i,0)|,|\max_i(r_i,0)|). \tag{A.32}$$

Therefore, by substituting (4.5) into R_1 and using a similar inequality to (A.31), that is,

$$\sum_{-T+1\leq r_2<\ldots<r_n\leq T-1}(1-p)^{r_n}r_n \leq \sum_r (1-p)^r r^{n-1} \leq \frac{1}{p^n}, \tag{A.33}$$

we have $\|R_1\|_{q/n} = O(\frac{n!}{p^nT^{n/2}})$. Finally, we replace $e(|\min_i(r_i,0)|,|\max_i(r_i,0)|)$ with $e(0,0)$ to give

$$\begin{aligned}
& \mathrm{cum}^*(J_T^*(\omega_{k_1}),\ldots,J_T^*(\omega_{k_n})) \\
= & \frac{1}{(2\pi T)^{n/2}}\sum_{r_2,\ldots,r_n=-T+1}^{T-1}(1-p)^{g(\underline{r})}\widehat{\kappa}_n(r_2,\ldots,r_n)e^{-ir_2\omega_{k_2}-\ldots-ir_n\omega_{k_n}}\sum_{t=1}^{T}e^{-it(\omega_{k_1}+\omega_{k_2}+\ldots+\omega_{k_n})} \\
& +R_1+R_2, \\
= & \frac{1}{(2\pi T)^{n/2}}h_n(\omega_{k_2},\ldots,\omega_{k_n})\sum_{t=1}^{T}e^{-it(\omega_{k_1}+\omega_{k_2}+\ldots+\omega_{k_n})}+R_1+R_2,
\end{aligned}$$

where R_1 is defined in (A.32) and

$$\begin{aligned}
R_2 = & \frac{1}{(2\pi T)^{n/2}}\sum_{r_2,\ldots,r_n=-T+1}^{T-1}(1-p)^{g(\underline{r})}\widehat{\kappa}_n(r_2,\ldots,r_n) \\
& \times e^{-ir_2\omega_{k_2}-\ldots-ir_n\omega_{k_n}}\Big(e(|\min_i(r_i,0)|,|\max_i(r_i,0)|)-e(0,0)\Big).
\end{aligned}$$

This leads to the bound

$$|R_2| \leq \frac{1}{(2\pi T)^{n/2}}\sum_{r_2,\ldots,r_n=-T+1}^{T-1}(1-p)^{g(\underline{r})}|\widehat{\kappa}_n(r_2,\ldots,r_n)|\big(|\max_i(r_i,0)|+|\min_i(r_i,0)|\big).$$

By using Hölder's inequality, it is straightforward to show that $\|\widehat{\kappa}_n(r_2,\ldots,r_n)\|_{q/n} \leq C < \infty$. This implies

$$\begin{aligned}
\|R_2\|_{q/n} \leq & \frac{C}{(2\pi T)^{n/2}}\sum_{r_2,\ldots,r_n=-T+1}^{T-1}(1-p)^{g(\underline{r})}\big(|\max_i(r_i,0)|+|\min_i(r_i,0)|\big) \\
\leq & \frac{2C}{(2\pi T)^{n/2}}\sum_{r_2,\ldots,r_n=-T+1}^{T-1}(1-p)^{g(\underline{r})}\max_i(|r_i|) = O(\frac{n!}{T^{n/2}p^n}),
\end{aligned}$$

where the above follows from (A.33). Therefore, by letting $R_{T,n} = R_1+R_2$ we have $\|R_{T,n}\|_{q/n} = O(\frac{n!}{T^{n/2}p^n})$. This proves (4.14).

To prove (a), we note that if $\sum_{l=1}^n \omega_{k_l} \notin \mathbb{Z}$, then the first term in (4.14) is zero and we have (4.15) (since $R_{T,n} = O_p(\frac{n!}{T^{n/2}p^n})$). On the other hand, if $\sum_l k_l \in \mathbb{Z}$, then the first term in (4.14) dominates and we use Lemma 4.2(ii) to obtain the other part of (4.15). The proof of (b) is similar, but uses Lemma 4.2(iii) rather than Lemma 4.2(ii), we omit the details. □

A.6 Proofs for Section 5

PROOF of Lemma 5.1. We first note that by Assumption 5.1, we have summability of the 2nd to 8th order cumulants (see Lemma A.9 for details). Therefore, to prove (ia) we can use Theorem 4.1(ii) to obtain

$$\begin{aligned}
\mathrm{cum}^*(J^*_{T,j_1}(\omega_{k_1}), J^*_{T,j_2}(\omega_{k_2})) &= f_{2;j_1,j_2}(\omega_{k_1})\frac{1}{T}\sum_{t=1}^{T}\exp(it(\omega_{k_1}+\omega_{k_2})) + O_p(\frac{1}{Tp^2}) \\
&= f_{2;j_1,j_2}(\omega_{k_1})I(k_1=-k_2) + O_p(\frac{1}{Tp^2})
\end{aligned}$$

The proof of (ib) and (ii) is identical, hence we omit the details. □

PROOF of Theorem 5.1 Since the only random component in $\widetilde{c}^*_{j_1,j_2}(r,\ell_1)$ are the DFTs, evaluating the covariance with respect to the bootstrap measure and using Lemma 5.2 to obtain an expression for the covariance between the DFTs gives

$$T\mathrm{cov}^*\big(\widetilde{c}^*_{j_1,j_2}(r,\ell_1), \widetilde{c}^*_{j_3,j_4}(r,\ell_2)\big) = \delta_{j_1j_3}\delta_{j_2j_4}\delta_{\ell_1\ell_2} + \delta_{j_1j_4}\delta_{j_2j_3}\delta_{\ell_1,-\ell_2} + \kappa_T^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4) + O_p\left(\frac{1}{Tp^4}\right)$$

$$T\mathrm{cov}^*\big(\widetilde{c}^*_{j_1,j_2}(r,\ell_1), \overline{\widetilde{c}^*_{j_3,j_4}(r,\ell_2)}\big) = \delta_{j_1j_3}\delta_{j_2j_4}\delta_{\ell_1\ell_2} + \delta_{j_1j_4}\delta_{j_2j_3}\delta_{\ell_1,-\ell_2} + \kappa_T^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4) + O_p\left(\frac{1}{Tp^4}\right),$$

which gives both part (i) and (ii). □

The proof of the above lemma is based on $\widetilde{c}^*_{j_1,j_2}(r,\ell)$ and we need to show that this is equivalent to $\widehat{c}^*_{j_1,j_2}(r,\ell)$ and $\acute{c}^*_{j_1,j_2}(r,\ell)$, which requires the following lemma.

Lemma A.14

Suppose $\{\underline{X}_t\}_t$ is a time series with a constant mean which satisfies Assumption 5.2(B2). Let $\widehat{\mathbf{f}}^*_T$ be defined in (2.21) and define $\widetilde{\mathbf{f}}^*_T(\omega_k) = \widehat{\mathbf{f}}^*_T(\omega_k) - \mathrm{E}(\widehat{\mathbf{f}}^*_T(\omega_k))$. Then we have

$$\big\|\mathrm{E}^*|\widetilde{\mathbf{f}}^*_T(\omega_k)|^2\big\|_4 = O\left(\frac{1}{bT} + \frac{1}{T^{3/2}p^3} + \frac{1}{T^{3/2}pb}\right), \tag{A.34}$$

$$\big\|\mathrm{cum}^*_4\big(\widetilde{\mathbf{f}}^*_T(\omega_k)\big)\big\|_2 = O\left(\frac{1}{(bT)^2} + \frac{1}{(Tp^2)^2}\right), \tag{A.35}$$

$$\big\|\mathrm{E}^*|\widetilde{\mathbf{f}}^*_T(\omega_k)|^4\big\|_2 = O\left(\frac{1}{(bT)^2} + \frac{1}{(Tp^2)^2}\right), \tag{A.36}$$

$$\big\|\mathrm{E}^*|J^*_{T,j_1}(\omega_k)\overline{J^*_{T,j_2}(\omega_{k+r})}|\big\|_8 = O\Big((1+\frac{1}{(T^{1/2}p)})^{1/2}\Big), \tag{A.37}$$

$$\|\mathrm{E}^*(J^*_{k,j_1}\overline{J}_{k+r,j_2})\|_8 = O(\frac{1}{(T^{1/2}p)^2}), \tag{A.38}$$

$$\left\|\mathrm{cov}^*(J^*_{T,j_1}(\omega_k)\overline{J^*_{T,j_2}(\omega_{k+r})}, J^*_{T,j_3}(\omega_s)\overline{J^*_{T,j_4}(\omega_s)})\right\|_4 = \begin{cases} O(\frac{1}{(pT^{1/2})^2}), & k=s \ \ or\ k=s+r \\ O(\frac{1}{(pT^{1/2})^3}), & otherwise \end{cases} \tag{A.39}$$

$$\left\|\mathrm{cum}^*_3(J^*_{T,j_1}(\omega_{k_1})\overline{J^*_{T,j_2}(\omega_{k_1+r})}, J^*_{T,j_3}(\omega_{k_2})\overline{J^*_{T,j_4}(\omega_{k_2+r})}, \widetilde{f}^*_{j_5j_6}(\omega_{k_1}))\right\|_{8/3}$$
$$= \begin{cases} O\big(\frac{1}{(pT^{1/2})^2}\big), & k_1=k_2 \ \ or\ k_1+r=k_2 \\ O\left(\frac{1}{bT(T^{1/2}p)^4}+\frac{1}{(pT^{1/2})^3}\right), & otherwise \end{cases}, \tag{A.40}$$

$$\left\|\mathrm{cum}^*(J^*_{T,j_1}(\omega_{k_1})\overline{J^*_{T,j_2}(\omega_{k_1+r})}), \overline{J_{T,j_1}}^*(\omega_{k_2})J_{T,j_2^*(\omega_{k_2+r})})\right\|_4 = \begin{cases} O(1), & k_1=k_2 \ \ or\ k_1=k_2+r \\ O(\frac{1}{(T^{1/2}p)^4}), & otherwise \end{cases} \tag{A.41}$$

$$\left\|\mathrm{cum}^*_2(J^*_{T,j_1}(\omega_{k_1})\overline{J^*_{T,j_2}(\omega_{k_1+r})}, \widetilde{f}^*_{j_3j_4}(\omega_{k_2}))\right\|_4 = O\left(\frac{1}{(pT^{1/2})^3}+\frac{1}{bT(pT^{1/2})^2}\right), \tag{A.42}$$

$$\left\|\mathrm{E}^*(\tilde{I}^*_{k_1,r,j_1,j_2}\tilde{I}^*_{k_2,r,j_3,j_4}\widetilde{f}^*_{k_1}\widetilde{f}^*_{k_2})\right\|_2 = \begin{cases} O(\frac{1}{(T^{1/2}p)^3}+\frac{1}{(Tb)(T^{1/2}p)^2}), & k_1=k_2 \ \ or\ k_1=k_2+r \\ O(\frac{1}{(T^{1/2}p)^4}+\frac{1}{(Tb)(T^{1/2}p)^3}), & otherwise \end{cases} \tag{A.43}$$

all these bounds are uniform over frequency and $\tilde{I}^*_{k,r} = I^*_{k,r} - \mathrm{E}^*(I^*_{k,r})$.

PROOF. Without loss of generality, we will prove the result in the univariate case (and under the assumption of nonstationarity). We will make wide use of (4.15) and (4.16) which we summarize below. For $n \in \{2,3\}$, we have

$$\left\|\mathrm{cum}^*(J^*_T(\omega_{k_1}),\ldots,J^*_T(\omega_{k_n}))\right\|_{q/n} = \begin{cases} O\left(\frac{1}{T^{n/2-1}}+\frac{1}{(T^{1/2}p)^{n-1}}\right), & \sum_{l=1}^n \omega_{k_l} \in \mathbb{Z} \\ O\left(\frac{1}{(T^{1/2}p)^n}\right), & \sum_{l=1}^n \omega_{k_l} \notin \mathbb{Z} \end{cases} \tag{A.44}$$

and, for $n \geq 4$,

$$\left\|\mathrm{cum}^*(J^*_T(\omega_{k_1}),\ldots,J^*_T(\omega_{k_n}))\right\|_{q/n} = \begin{cases} O\left(\frac{1}{T^{n/2-1}p^{n-3}}+\frac{1}{(T^{1/2}p)^{n-1}}\right), & \sum_{l=1}^n \omega_{k_l} \in \mathbb{Z} \\ O\left(\frac{1}{(T^{1/2}p)^n}\right), & \sum_{l=1}^n \omega_{k_l} \notin \mathbb{Z} \end{cases}$$

To simplify notation, let $J_T^*(\omega_k) = J_k^*$. To prove (A.34), we expand $\text{var}^*(\widehat{f}_T^*(\omega))$ to give

$$
\begin{aligned}
& \|\text{var}^*(\widehat{f}^*(\omega_k))\|_4 \\
\leq & \left\| \frac{1}{T^2} \sum_{l_1,l_2} K_b(\omega_k - \omega_{l_1}) K_b(\omega_k - \omega_{l_2}) \Big[\text{cov}(J_{l_1}^*, J_{l_2}^*)\text{cov}(\overline{J}_{l_1}^*, \overline{J}_{l_2}^*) + \text{cov}(J_{l_1}^*, \overline{J}_{l_2}^*)\text{cov}(\overline{J}_{l_1}^*, J_{l_2}^*) + \right. \\
& \left. \text{cum}(J_{l_1}^*, \overline{J}_{l_2}^*, J_{l_2}^*, \overline{J}_{l_2}^*) \right\|_4 \\
= & \left\| \frac{1}{T^2} \sum_{l_1 \neq l_2} K_b(\omega_k - \omega_{l_1}) K_b(\omega_k - \omega_{l_2}) \Big[\text{cov}(J_{l_1}^*, J_{l_2}^*)\text{cov}(\overline{J}_{l_1}^*, \overline{J}_{l_2}^*) + \text{cov}(J_{l_1}^*, \overline{J}_{l_2}^*)\text{cov}(\overline{J}_{l_1}^*, J_{l_2}^*) \Big] \right\|_4 + \\
& \left\| \frac{1}{T^2} \sum_{l_1,l_2} K_b(\omega_k - \omega_{l_1}) K_b(\omega_k - \omega_{l_2}) \text{cum}(J_{l_1}^*, \overline{J}_{l_2}^*, J_{l_2}^*, \overline{J}_{l_2}^*) \Big] \right\|_4 + \\
& \left\| \frac{1}{T^2} \sum_{l} K_b(\omega_k - \omega_l)^2 |\text{cov}(J_l^*, J_l^*)|^2 \right\|_4 \\
\leq & \frac{C}{T^2} \sum_{l_1,l_2} K_b(\omega_k - \omega_{l_1}) K_b(\omega_k - \omega_{l_2}) \big(\frac{1}{(T^{1/2}p)^4} + \frac{1}{(T^{1/2}p)^3} + \frac{1}{T} \big) + \\
& \frac{C}{T^2} \sum_{l} K_b(\omega_k - \omega_l)^2 \big(1 + \frac{1}{T^{1/2}p} \big) \text{ (by using } (A.44)) \\
= & O\left(\frac{1}{bT} + \frac{1}{(T^{1/2}p)^3} + \frac{1}{(bT)(Tp^{1/2})} \right) \text{ (these are just the leading terms).}
\end{aligned}
$$

We now prove (A.36), expanding $\text{cum}_4^*(\widehat{f}^*(\omega_k))$ gives

$$
\|\text{cum}_4^*(\widehat{f}^*(\omega_k))\|_2 = \Big\| \frac{1}{T^4} \sum_{l_1,l_2,l_3,l_4} \left(\prod_{i=1}^{4} K_b(\omega_k - \omega_{l_i}) \right) \text{cum}^*\big(|J_{l_1}^*|^2, |J_{l_2}^*|^2, |J_{l_3}^*|^2, |J_{l_4}^*|^2 \big) \Big\|_2.
$$

By using indecomposable partitions to decompose the cumulant $\text{cum}^*\big(|J_{l_1}^*|^2, |J_{l_2}^*|^2, |J_{l_3}^*|^2, |J_{l_4}^*|^2\big)$ in terms of cumulants of J_k and using (A.44), we can show that the leading term of $\text{cum}^*\big(|J_{l_1}^*|^2, |J_{l_2}^*|^2, |J_{l_3}^*|^2, |J_{l_4}^*|^2\big)$ is the product of four covariances of the type $\text{cum}(J_{l_1}, J_{l_2})\text{cum}(\bar{J}_{l_2}, J_{l_3})\text{cum}(\bar{J}_{l_3}, J_{l_4})\text{cov}(\bar{J}_{l_4}, \bar{J}_{l_1})$

$$
\|\text{cum}_4^*(\widehat{f}^*(\omega_k))\|_2 = O\left(\frac{1}{(Tb)^3} (1 + \frac{1}{T^{1/2}p})^4 + \frac{1}{(T^{1/2}p)^4} \frac{1}{(bT)^2} + \frac{1}{(T^{1/2}p)^3} + \frac{1}{(bT)^2 (Tp^{1/2})^2} \right),
$$

which proves (A.36). Since $\text{E}^*(\widetilde{f}^*(\omega)) = 0$, we have

$$
\text{E}^* |\widetilde{f}^*(\omega_k)|^4 = 3\text{var}^*(\widetilde{f}^*(\omega_k))^2 + \text{cum}_4(\widetilde{f}^*(\omega_k)),
$$

therefore by using (A.34) and (A.35), we obtain (A.36).

To prove (A.37), we note that $\text{E}^* |J_T^*(\omega_k) \overline{J_T^*(\omega_{k+r})}| \leq (\text{E}^* |J_T^*(\omega_k) \overline{J_T^*(\omega_{k+r})}|^2)^{1/2}$. Therefore, by using

$$
\begin{aligned}
\text{E}^* |J_T^*(\omega_k) \overline{J_T^*(\omega_{k+r})}|^2 = & \text{E}^* |J_T^*(\omega_k)|^2 \text{E}^* |\overline{J_T^*(\omega_{k+r})}|^2 + |\text{E}^*(J_T^*(\omega_k) J_T^*(\omega_{k+r}))|^2 + \\
& |\text{E}^*(J_T^*(\omega_k) \overline{J_T^*(\omega_{k+r})})|^2 + \text{cum}^*(J_T^*(\omega_k), \overline{J_T^*(\omega_{k+r})}, \overline{J_T^*(\omega_k)}, J_T^*(\omega_{k+r})),
\end{aligned}
$$

we have

$$\left\|\mathrm{E}^*|J_T^*(\omega_k)\overline{J_T^*(\omega_{k+r})}|\right\|_8 \leq \left\|\left[\mathrm{E}^*|J_T^*(\omega_k)\overline{J_T^*(\omega_{k+r})}|^2\right]^{1/2}\right\|_8 = \mathrm{E}\left|\left[\mathrm{E}^*(J_T^*(\omega_k)\overline{J_T^*(\omega_{k+r})})^2\right]^4\right|^{1/8}$$

$$= \left\|\mathrm{E}^*(J_T^*(\omega_k)\overline{J_T^*(\omega_{k+r})})^2\right\|_4^{1/2} = O\left(1+\frac{1}{(T^{1/2}p)}\right)^{1/2},$$

which proves (A.37). The proof of (A.38) immediately follows from (A.44).

To prove (A.39), we expand it in terms of covariances and cumulants

$$\begin{aligned}&\mathrm{cov}^*(J_k^*\overline{J_{k+r}}^*, J_s^*\overline{J_s}^*)\\ =\ &\mathrm{cum}^*(J_{k_2}^*, \bar{J}_j^*)\mathrm{cum}^*(\bar{J}_{k_2+r}^*, J_j^*) + \mathrm{cum}^*(J_{k_2}^*, J_j^*)\mathrm{cum}^*(\bar{J}_{k_2+r}^*, \bar{J}_j^*) + \mathrm{cum}^*(J_{k_2}^*, J_j^*, \bar{J}_{k_2+r}^*, \bar{J}_j^*),\end{aligned}$$

thus by using (A.44) we obtain (A.39).

To prove (A.40), we expand the sample bootstrap spectral density in terms of DFTs to obtain

$$\mathrm{cum}_3^*(J_{k_1}^*\overline{J}_{k_1+r}^*, J_{k_2}^*\overline{J}_{k_2+r}^*, \widetilde{f}^*(\omega_{k_1})) = \frac{1}{T}\sum_l K_b(\omega_{k_1}-\omega_l)\mathrm{cum}_3^*(J_{k_1}^*\overline{J}_{k_1+r}^*, J_{k_2}^*\overline{J}_{k_2+r}^*, |J_l^*|^2).$$

By using indecomposable partitions to partition $\mathrm{cum}_3^*(J_{k_1}^*\overline{J}_{k_1+r}^*, J_{k_2}^*\overline{J}_{k_2+r}^*, |J_l^*|^2)$ in terms of cumulants of the DFTs, we observe that the leading term is the product of covariances. This gives

$$\|\mathrm{cum}_3^*(J_{k_1}^*\overline{J}_{k_1+r}^*, J_{k_2}^*\overline{J}_{k_2+r}^*, |J_l^*|^2)\|_{8/3} = \begin{cases} O\left(\left(1+\frac{1}{T^{1/2}p}\right)^2\left(\frac{1}{(T^{1/2}p)^2}\right)\right), & k_i = l \text{ and } k_i + r = k_j \\ O\left(\left(1+\frac{1}{T^{1/2}p}\right)\left(\frac{1}{(T^{1/2}p)^4}\right)\right), & k_i = k_j \text{ or } k_i + r = l \\ O\left(\frac{1}{(T^{1/2}p)^6}\right), & \text{otherwise} \end{cases}.$$

for $i, j \in \{1, 2\}$. By substituting the above into (A.45), we get

$$\left\|\mathrm{cum}_3^*(J_{k_1}^*\overline{J}_{k_1+r}^*, J_{k_2}^*\overline{J}_{k_2+r}^*, \widetilde{f}^*(\omega_{k_1}))\right\|_{8/3} = \begin{cases} O\left(\frac{1}{(pT^{1/2})^2}\right), & k_1 = k_2 \text{ or } k_1 + r = k_2 \\ O\left(\frac{1}{bT(T^{1/2}p)^4} + \frac{1}{(pT^{1/2})^3}\right), & \text{otherwise} \end{cases},$$

which proves (A.40). The proofs of (A.41) and (A.42) are identical to the proof of (A.40), hence we omit the details.

To prove (A.43), we expand the expectation in terms of cumulants and use identical methods to those used above to obtain the result.

Finally to prove (A.47), we use the Minkowski inequality to give

$$\begin{aligned}&\left\|\mathrm{E}^*[\widehat{f}^*(\omega_k)] - f(\omega_k))\right\|_8 \leq \left\|\frac{1}{T}\sum_j K_b(\omega_k-\omega_j)\left[\mathrm{E}^*(J_j^*\overline{J_j}^*) - \widehat{h}_2(\omega_j)\right]\right\|_8 + \\ &\left\|\frac{1}{T}\sum_j K_b(\omega_k-\omega_j)[\widehat{h}_2(\omega_j) - f(\omega_j)]\right\|_8 + \left|\frac{1}{T}\sum_j K_b(\omega_k-\omega_j)f(\omega_j) - f(\omega_k)\right|, \quad \text{(A.45)}\end{aligned}$$

where $\widehat{h}_2$ is defined in (4.9). We now bound the above terms. By using Theorem 4.1(ii) (for $n = 2$), we have

$$\left\| \frac{1}{T} \sum_j K_b(\omega_k - \omega_j) \big[\mathrm{E}^*(J_j^* \overline{J_j}^*) - \widehat{h}_2(\omega_j) \big] \right\|_8 \leq \frac{1}{T} \sum_j K_b(\omega_k - \omega_j) \big\| \mathrm{E}^*(J_j^* \overline{J_j}^*) - \widehat{h}_2(\omega_j) \big\|_8 \leq O(\frac{1}{T^{1/2}p}).$$

By using Lemma 4.2(ii), we obtain

$$\left\| \frac{1}{T} \sum_j K_b(\omega_k - \omega_j) [\widehat{h}_2(\omega_j) - f(\omega_j)] \right\|_8 \leq \frac{1}{T} \sum_j K_b(\omega_k - \omega_j) \big\| \widehat{h}_2(\omega_j) - f(\omega_j) \big\|_8 = O(\frac{1}{T^{1/2}p} + p).$$

By using using similar methods to those used to prove Lemma A.2(i), we have $\frac{1}{T} \sum_j K_b(\omega_k - \omega_j) f(\omega_j) - f(\omega_k) = O(b)$. Substituting the above into (A.45) gives (A.42). □

Analogous to $\widehat{\mathbf{C}}_T(r, \ell)$, direct analysis of the variance of $\widehat{\mathbf{C}}_T^*(r, \ell)$ and $\acute{\mathbf{C}}_T(r, \ell)$ with respect to the bootstrap measure is extremely difficult because of the $\widehat{\mathbf{L}}^*(\omega_k)$ and $\widehat{\mathbf{L}}(\omega_k)$ in the definition of $\widehat{\mathbf{C}}_T^*(r, \ell)$ and $\acute{\mathbf{C}}_T(r, \ell)$. However, analysis of $\widetilde{\mathbf{C}}_T^*(r, \ell)$ is much easier, therefore to show that the bootstrap variance converges to the true variance we will show that $\mathrm{var}^*(\widehat{\mathbf{C}}_T^*(r, \ell))$ and $\mathrm{var}^*(\acute{\mathbf{C}}_T^*(r, \ell))$ can be replaced with $\mathrm{var}^*(\widetilde{\mathbf{C}}_T^*(r, \ell))$. To prove this result, we require the following definitions

$$\begin{aligned}
\widehat{c}_{j_1,j_2}^*(r, \ell) &= \frac{1}{T} \sum_{k=1}^T \sum_{s_1,s_2=1}^d A_{j_1,s_1,j_2,s_2}(\widehat{\underline{f}}_{k,r}^*) J_{k,s_1}^* \overline{J_{k+r,s_2}^*} \exp(i\ell\omega_k), \\
\breve{c}_{j_1,j_2}^*(r, \ell) &= \frac{1}{T} \sum_{k=1}^T \sum_{s_1,s_2=1}^d A_{j_1,s_1,j_2,s_2}(\mathrm{E}^*(\widehat{\underline{f}}_{k,r}^*)) J_{k,s_1}^* \overline{J_{k+r,s_2}^*} \exp(i\ell\omega_k), \\
\widetilde{c}_{j_1,j_2}^*(r, \ell) &= \frac{1}{T} \sum_{k=1}^T \sum_{s_1,s_2=1}^d A_{j_1,s_1,j_2,s_2}(\underline{f}_{k,r}) J_{k,s_1}^* \overline{J_{k+r,s_2}^*} \exp(i\ell\omega_k). \qquad (A.46)
\end{aligned}$$

We also require the following lemma which is analogous to Lemma A.2, but applied to the bootstrap spectral density estimator $\mathrm{E}^*(\underline{J}_T(\omega) \overline{\underline{J}}_T(\omega)')$.

Lemma A.15 *Suppose that $\{\underline{X}_t\}$ is an α-mixing second order stationary or locally stationary time series (which satisfies Assumption 3.2(L2)) with $\alpha > 4$ and the moment $\sup_t \|\underline{X}_t\|_s < \infty$ where $s > 4\alpha/(\alpha - 2)$. For $h \neq 0$ either the covariance of local covariance satisfies $|\boldsymbol{\kappa}(h)|_1 \leq C|h|^{-(2+\varepsilon)}$ or $|\boldsymbol{\kappa}(u; h)|_1 \leq C|h|^{-(2+\varepsilon)}$. Let $\underline{J}_T^*(\omega_k)$ be defined as in Step 5 of the bootstrap scheme.*

(a) If $Tp^2 \to \infty$ and $p \to 0$ as $T \to \infty$, then we have

- *(i)* $\sup_{1 \leq k \leq T} |\mathrm{E}\big[\mathrm{E}^*\big(\widehat{\mathbf{f}}_T^*(\omega_k)\big)\big] - \mathbf{f}(\omega_k)| = O(p + b + (bT)^{-1})$ *and* $\mathrm{var}(\widehat{\mathbf{f}}_T(\omega_k)) = O(\frac{1}{pT} + \frac{1}{T^{3/2}p^{5/2}} + \frac{1}{T^2p^4})$.
- *(ii)* $\sup_{1 \leq k \leq T} |\mathrm{E}^*\big(\widehat{\mathbf{f}}_T^*(\omega_k)\big) - \mathbf{f}(\omega_k)|_1 \xrightarrow{P} 0,$

(iii) *Further, if* $\mathbf{f}(\omega)$ *is nonsingular on* $[0, 2\pi]$*, then for all* $1 \leq s_1, s_2 \leq d$*, we have* $\sup_{1 \leq k \leq T} |L_{s_1,s_2}(\mathrm{E}^*(\underline{\widehat{f}}_T^*(\omega_k))) - L_{s_1,s_2}(\underline{f}(\omega_k))| \xrightarrow{P} 0$.

(b) *In addition, suppose for the mixing rate* $\alpha > 16$ *there exists a* $s > 16\alpha/(\alpha - 2)$ *such that* $\sup_t \|\underline{X}_t\|_s < \infty$*. Then, we have*

$$\|\mathrm{E}^*(\widehat{\mathbf{f}}^*(\omega_k)) - \mathbf{f}(\omega_k)\|_8 = O\left(\frac{1}{T^{1/2}p} + \frac{1}{p^2 T} + p + b + \frac{1}{bT}\right). \tag{A.47}$$

PROOF. To reduce notation, we prove the result in the univariate case. By using Theorem 4.1 and equation (4.14), we have

$$\mathrm{E}^*|J_T^*(\omega_j)|^2 = \widehat{h}_2(\omega_j) + R_1(\omega_j),$$

where $\| \sup_\omega |R_1(\omega_j)|\|_2 = O(\frac{1}{Tp^2})$. Therefore, substituting the above result into $\mathrm{E}^*(\widehat{f}_T^*(\omega_k)) = \mathrm{E}^*\big(\frac{1}{T}\sum_{j=-T/2}^{T/2} K_b(\omega_k - \omega_j)|J_T(\omega_j)|^2\big)$, we have

$$\mathrm{E}^*\big(\widehat{f}_T^*(\omega_k)\big) = \frac{1}{T}\sum_{|r|<T-1} \lambda_b(r)(1-p)^{|r|}\exp(ir\omega_k)\hat{\kappa}_n(r) + \tilde{R}_2(\omega_k) = \widetilde{f}_T(\omega_k) + \tilde{R}_2(\omega_k), \tag{A.48}$$

where $\widetilde{f}_T(\omega) = \frac{1}{T}\sum_{|r|<T-1} \lambda_b(r)(1-p)^{|r|}\exp(ir\omega)\hat{\kappa}_n(r)$ and $\| \sup_{\omega_k} \tilde{R}_2(\omega_k)\|_2 = O(\frac{1}{Tp^2})$. Thus for the remainder of the proof, we only need to analyze the leading term $\widetilde{f}_T(\omega)$ (note that unlike $\mathrm{E}^*\big(\widehat{f}_T^*(\omega_s)\big)$, $\widehat{f}_T(\omega)$ is defined over $[0, 2\pi]$ not just the fundamental frequencies).

Using the same methods as those used in the proof of Lemma A.2(a), it is straightforward to show that

$$\mathrm{E}(\widetilde{f}_T(\omega)) = f(\omega) + R_3(\omega), \tag{A.49}$$

where $\sup_\omega |R_3(\omega)| = O(\frac{1}{bT} + b + p)$, this proves $\sup_{1\leq k\leq T} |\mathrm{E}\big[\mathrm{E}^*\big(\widehat{f}_T^*(\omega_k)\big)\big] - f(\omega_k)| = O(p + b + (bT)^{-1})$. Using (4.6), it is straightforward to show that $\mathrm{var}(\widetilde{f}_T(\omega)) = O((pT)^{-1})$, therefore by (A.48) and the above we have (ai).

By using identical methods to those used to prove Lemma A.2(ci) we can show $\sup_\omega |\widetilde{f}_T(\omega) - \mathrm{E}(\widetilde{f}_T(\omega))| \xrightarrow{P} 0$. Thus from uniform convergence of $\widetilde{f}_T(\omega)$ and (ai) we immediately obtain uniform convergence of $\mathrm{E}^*(\widehat{f}_T(\omega_k))$ ($\sup_{1\leq s\leq T} |\mathrm{E}^*(\widehat{f}_T(\omega_k)) - f(\omega_k))|$). Similarly to show (aiii) we apply identical method to those used in the proof of Lemma A.2(cii) to $\widetilde{f}_T(\omega)$.

Finally, to show (b), we use that

$$\big\|\mathrm{E}^*(\widehat{f}^*(\omega_k)) - f(\omega_k)\big\|_8 \leq \big\|\widetilde{f}_T(\omega_k) - \mathrm{E}(\widetilde{f}_T(\omega_k))\big\|_8 + |\mathrm{E}(\widetilde{f}_T(\omega_k)) - f(\omega_k)| + \|R_2(\omega_k)\|_8.$$

By using (4.6) and the Minkowski inequality, we can show $\big\|\widetilde{f}_T(\omega_k) - \mathrm{E}(\widetilde{f}_T(\omega_k))\big\|_8 = O((Tp^2)^{-1/2})$, where this and the bounds above give (b). □

Lemma A.16 *Suppose that Assumption 5.2 and the conditions in Lemma A.15 hold. Then we have*

$$T\left(\mathrm{E}^*\left[\widehat{c}^*_{j_1,j_2}(r_1,\ell_1)\right]\overline{\mathrm{E}^*\left[\widehat{c}^*_{j_3,j_4}(r_2,\ell_2)\right]}-\mathrm{E}^*\left[\breve{c}^*_{j_1,j_2}(r_1,\ell_1)\right]\overline{\mathrm{E}^*\left[\breve{c}^*_{j_3,j_4}(r_2,\ell_2)\right]}\right)=O_p\left(a(T,b,p)\right) \quad \text{(A.50)}$$

$$T\left(\mathrm{E}^*\left[\widehat{c}^*_{j_1,j_2}(r_1,\ell_1)\overline{\widehat{c}^*_{j_3,j_4}(r_2,\ell_2)}\right]-\mathrm{E}^*\left[\breve{c}^*_{j_1,j_2}(r_1,\ell_1)\overline{\breve{c}^*_{j_3,j_4}(r_2,\ell_2)}\right]\right)=O_p\left(a(T,b,p)\right) \quad \text{(A.51)}$$

where $a(T,b,p)=\frac{1}{bTp^2}+\frac{1}{Tp^4}+\frac{1}{b^2T}+b+\frac{1}{Tp^2}+\frac{1}{T^{1/2}p}$.

PROOF. To simplify notation, we consider the case $d=1$, $\ell_1=\ell_2=0$ and $r_1=r_2=r$. We first prove (A.50). Recalling that the only difference between $\widehat{c}^*(r,0)$ and $\breve{c}^*(r,0)$, is that $A(\underline{\widehat{f}}^*_{k,r})$ is replaced with $A(\mathrm{E}^*(\underline{\widehat{f}}^*_{k,r}))$, the difference between their expectations squared (with respect to the stationary bootstrap measure) is

$$T\left(\left|\mathrm{E}^*(\widehat{c}^*(r,0))\right|^2-\left|\mathrm{E}^*(\breve{c}^*(r,0))\right|^2\right)=\frac{1}{T}\sum_{k_1,k_2}\Bigg(\mathrm{E}^*\left[A(\underline{\widehat{f}}^*_{k_1,r})J^*_{k_1}\overline{J}^*_{k_1+r}\right]\mathrm{E}^*\left[A(\underline{\widehat{f}}^*_{k_2,r})J^*_{k_2}\overline{J}^*_{k_2+r}\right]$$
$$-A(\mathrm{E}^*(\underline{\widehat{f}}^*_{k_1,r}))A(\mathrm{E}^*(\underline{\widehat{f}}^*_{k_2,r}))\mathrm{E}^*(J^*_{k_1}\overline{J}^*_{k_1+r})\mathrm{E}^*(J^*_{k_2}\overline{J}^*_{k_2+r})\Bigg)$$
$$=\frac{1}{T}\sum_{k_1,k_2}\left(\mathrm{E}^*\left[a^*_{k_1}I^*_{k_1,r}\right]\mathrm{E}^*\left[a^*_{k_2}I^*_{k_2,r}\right]-\widehat{a}_{k_1}\widehat{a}_{k_1}\mathrm{E}^*[I^*_{k_1,r}]\mathrm{E}^*[I^*_{k_2,r}]\right), \quad \text{(A.52)}$$

where

$$a^*_k=A(\underline{\widehat{f}}^*_{k,r}),\quad \widehat{a}_k=A(\mathrm{E}^*(\underline{\widehat{f}}^*_{k,r})),\quad \widetilde{f}^*_k=(\underline{\widehat{f}}^*_{k,r}-(\mathrm{E}^*(\underline{\widehat{f}}^*_{k,r}))\text{ and }I^*_{k,r}=J^*_k\overline{J}^*_{k+r}. \quad \text{(A.53)}$$

To bound the above, we use the above Taylor expansion

$$A(\underline{\widehat{f}}^*_{k,r})=A(\mathrm{E}^*(\underline{\widehat{f}}^*_{k,r}))+(\underline{\widehat{f}}^*_{k,r}-\mathrm{E}^*(\underline{\widehat{f}}^*_{k,r}))'\nabla A(\mathrm{E}^*(\underline{\widehat{f}}^*_{k,r}))+\frac{1}{2}(\underline{\widehat{f}}^*_{k,r}-\mathrm{E}^*(\underline{\widehat{f}}^*_{k,r}))'\nabla^2A(\underline{\bar{f}}^*_{k,r})(\underline{\widehat{f}}^*_{k,r}-\mathrm{E}^*(\underline{\widehat{f}}^*_{k,r})),$$

where ∇ and ∇^2 denotes the first and second partial derivative with respect to $\underline{f}_{k,r}$ and $\underline{\breve{f}}^*_{k,r}$ lies between $\underline{\widehat{f}}^*_{k,r}$ and $\mathrm{E}^*(\underline{\widehat{f}}^*_{k,r})$. To reduce cumbersome notation (and with a slight loss of accuracy, since it will not effect the calculation) we shall ignore that $\underline{\widehat{f}}_{k,r}$ is a vector and that many parts of the proof below require us to the complex conjugate $\overline{\tilde{I}^*_{k,r}}$ (we use instead $\tilde{I}^*_{k,r}$) and use (A.53) to rewrite the Taylor expansion as

$$a^*_k = \widehat{a}_k+\widetilde{f}^*_k\frac{\partial\widehat{a}_k}{\partial f}+\widetilde{f}^{*2}_k\frac{1}{2}\frac{\partial^2\bar{a}^*_k}{\partial f^2}, \quad \text{(A.54)}$$

where $\bar{a}^*_k=\nabla^2A(\underline{\bar{f}}^*_{k,r})$. Substituting (A.54) into (A.52), we obtain the decomposition

$$T\left(\left|\mathrm{E}^*(\widehat{c}^*(r,0))\right|^2-\left|\mathrm{E}^*(\breve{c}^*(r,0))\right|^2\right)=\sum_{i=1}^{8}I_i,$$

where the terms $\{I_i\}_{i=1}^{8}$ are

$$\begin{aligned}
I_1 &= \frac{1}{T}\sum_{k_1,k_2}\widehat{a}_{k_1}\frac{\partial\widehat{a}_{k_2}}{\partial f}\mathrm{E}^*(I^*_{k_1,r})\mathrm{E}^*\big(I^*_{k_2,r}\widetilde{f}^*_{k_2}\big)\\
&= \frac{1}{T}\sum_{k_1,k_2}\widehat{a}_{k_1}\frac{\partial\widehat{a}_{k_2}}{\partial f}\mathrm{E}^*(I^*_{k_1,r})\mathrm{E}^*\big(\widetilde{I}^*_{k_2,r}\widetilde{f}^*_{k_2}\big) \qquad (A.55)\\
I_2 &= \frac{1}{T}\sum_{k_1,k_2}\frac{\partial\widehat{a}_{k_1}}{\partial f}\frac{\partial\widehat{a}_{k_2}}{\partial f}\mathrm{E}^*\big[I^*_{k_1,r}\widetilde{f}^*_{k_1}\big]\mathrm{E}^*\big[I^*_{k_2,r}\widetilde{f}^*_{k_2}\big]\\
&= \frac{1}{T}\sum_{k_1,k_2}\frac{\partial\widehat{a}_{k_1}}{\partial f}\frac{\partial\widehat{a}_{k_2}}{\partial f}\mathrm{E}^*\big[\widetilde{I}^*_{k_1,r}\widetilde{f}^*_{k_1}\big]\mathrm{E}^*\big[\widetilde{I}^*_{k_2,r}\widetilde{f}^*_{k_2}\big]\\
I_3 &= \frac{1}{2T}\sum_{k_1,k_2}\widehat{a}_{k_2}\mathrm{E}^*\left[I^*_{k_1,r}\widetilde{f}^{2*}_{k_1}\frac{\partial^2\bar{a}^*_{k_1}}{\partial f^2}\right]\mathrm{E}^*(I^*_{k_2,r})\\
I_7 &= \frac{1}{T}\sum_{k_1,k_2}\frac{\partial\widehat{a}_{k_2}}{\partial f}\mathrm{E}^*\left[I^*_{k_1,r}\widetilde{f}^{2*}_{k_1}\frac{\partial^2\bar{a}^*_{k_1}}{\partial f^2}\right]\mathrm{E}^*\big[\widetilde{I}^*_{k_2,r}\widetilde{f}^*_{k_2}\big]\\
I_8 &= \frac{1}{4T}\sum_{k_1,k_2}\mathrm{E}^*\left[I^*_{k_1,r}\widetilde{f}^{2*}_{k_1}\frac{\partial^2\bar{a}^*_{k_1}}{\partial f^2}\right]\mathrm{E}^*\left[I^*_{k_2,r}\widetilde{f}^{2*}_{k_2}\frac{\partial^2\bar{a}^*_{k_2}}{\partial f^2}\right]
\end{aligned}$$

(with $\widetilde{I}^*_{k,r} = I^*_{k,r} - \mathrm{E}^*(I^*_{k,r})$) and I_4, I_5, I_6 are defined similarly. We first bound I_1. Writing $\widetilde{f}^*_k = \frac{1}{T}\sum_{j=1}^{T}K_b(\omega_k-\omega_j)\widetilde{I}^*_{j,0}$ gives

$$I_1 = \frac{1}{T}\sum_{k_1,k_2,j}K_b(\omega_{k_2}-\omega_j)\widehat{a}_{k_1}\frac{\partial\widehat{a}_{k_2}}{\partial f}\mathrm{E}^*(I^*_{k_1,r})\mathrm{cov}^*(\widetilde{I}^*_{k_2,r},\widetilde{I}^*_{j,0}).$$

By using the uniform convergence result in Lemma A.15(a), we have $\sup_k|\hat{a}_k - a_k| \stackrel{P}{\to} 0$. Therefore, $|I_1| = O_p(1)\tilde{I}_1$, where

$$\tilde{I}_1 = \frac{1}{T}\sum_{k_1,k_2,j}|K_b(\omega_{k_2}-\omega_j)|\Big|a_{k_1}\frac{\partial a_{k_2}}{\partial f}\Big|\Big|\mathrm{E}^*(I^*_{k_1,r})\mathrm{cov}^*(\widetilde{I}^*_{k_2,r},\widetilde{I}^*_{j,0})\Big|$$

with $a_k = A(\underline{f}_{k,r})$ and $\frac{\partial a_{k_1}}{\partial f} = \nabla_f A(\underline{f}_{k,r})$. Taking expectations inside $\tilde{I}_1$ gives

$$\|\tilde{I}_1\|_1 \leq \frac{1}{T}\sum_{k_1,k_2,j}|K_b(\omega_{k_2}-\omega_j)|\Big|a_{k_1}\frac{\partial a_{k_2}}{\partial f}\Big|\underbrace{\|\mathrm{E}^*(I^*_{k_1,r})\|_2}_{(A.37)}\underbrace{\big\|\mathrm{cov}^*(\widetilde{I}^*_{k_2,r},\widetilde{I}^*_{j,0})\big\|_2}_{(A.39)},$$

thus we have $\tilde{I}_1 = O_p(\frac{1}{p^4T}+\frac{1}{p^6T^2}) = O(\frac{1}{Tp^4})$ and $I_1 = O_p(\frac{1}{p^4T}+\frac{1}{p^6T^2}) = O(\frac{1}{Tp^4})$. We now bound I_2. By using an identical method to that given above, we have $|I_2| = O_p(1)\tilde{I}_2$, where

$$\begin{aligned}
\tilde{I}_2 &= \frac{1}{T}\sum_{k_1,k_2,j_1,j_2}|K_b(\omega_{k_2}-\omega_{j_1})||K_b(\omega_{k_2}-\omega_{j_2})|\Big|\frac{\partial a_{k_1}}{\partial f}\frac{\partial a_{k_2}}{\partial f}\Big|\Big|\mathrm{cov}^*(\widetilde{I}^*_{k_1,r},\widetilde{I}^*_{j_1,0})\mathrm{cov}^*(\widetilde{I}^*_{k_2,r},\widetilde{I}^*_{j_2,0})\Big|\\
\Rightarrow \|\tilde{I}_2\| &\leq \frac{1}{T}\sum_{k_1,k_2,j_1,j_2}|K_b(\omega_{k_2}-\omega_{j_1})||K_b(\omega_{k_2}-\omega_{j_2})|\Big|\frac{\partial a_{k_1}}{\partial f}\frac{\partial a_{k_2}}{\partial f}\Big|\underbrace{\big\|\mathrm{cov}^*(\widetilde{I}^*_{k_1,r},\widetilde{I}^*_{j_1,0})\big\|_2}_{(A.39)}\cdot\underbrace{\big\|\mathrm{cov}^*(\widetilde{I}^*_{k_2,r},\widetilde{I}^*_{j_2,0})\big\|_2}_{(A.39)},
\end{aligned}$$

which gives $\tilde{I}_2 = O(\frac{1}{T^2p^4})$ and thus $I_2 = O_p(\frac{1}{T^2p^4})$.

To bound I_3, we use Hölder's inequality

$$|I_3| \leq \frac{1}{T}\sum_{k_1,k_2} |\widehat{a}_{k_2}| \cdot \mathrm{E}^*|\widetilde{f}_{k_1}^{4*}|^{1/2}\mathrm{E}^*\big|I_{k_1,r}^{6*}\big|^{1/6}\mathrm{E}^*\big|(\frac{\partial^2 \bar{a}_{k_1}^*}{\partial f^2})^3\big|^{1/3}|\mathrm{E}^*(I_{k_2,r}^*)|.$$

Under Assumption 5.2(B1), we have that $\mathrm{E}^*\big|(\frac{\partial^2 \bar{a}_{k_1}^*}{\partial f^2})^3\big|^{1/3}$ is uniformly bounded in probability. Therefore, using this and Lemma A.15(a), we have $|I_3| = O_p(1)\tilde{I}_3$, where

$$\tilde{I}_3 = \frac{1}{T}\sum_{k_1,k_2} |a_{k_2}| \cdot \mathrm{E}^*|\widetilde{f}_{k_1}^{4*}|^{1/2}\mathrm{E}^*\big|I_{k_1,r}^{6*}\big|^{1/6}|\mathrm{E}^*(I_{k_2,r}^*)|.$$

Taking expectations of the above and using Hölder's inequality gives

$$\begin{aligned}
\mathrm{E}(\tilde{I}_3) &\leq \frac{1}{T}\sum_{k_1,k_2} |a_{k_2}| \cdot \|\mathrm{E}^*|\widetilde{f}_{k_1}^{4*}|^{1/2}\|_2 \cdot \|\mathrm{E}^*\big|I_{k_1,r}^{6*}\big|^{1/6}\|_6 \cdot \|\mathrm{E}^*(I_{k_2,r}^*)\|_3 \\
&= \frac{1}{T}\sum_{k_1,k_2} |a_{k_2}| \cdot \underbrace{\|\mathrm{E}^*(\widetilde{f}_{k_1}^{4*})\|_1^{1/2}}_{(A.36)}\underbrace{\|\mathrm{E}^*\big|I_{k_1,r}^*\big|^6\|_1^{1/6}}_{(A.37)}\underbrace{\|\mathrm{E}^*(I_{k_2,r}^*)\|_3}_{(A.38)}.
\end{aligned}$$

Thus by using Lemma A.14, we obtain $|I_3| = O_p(\frac{1}{bTp^2})$ Using a similar method, we obtain $|I_7| = O_p(1)\tilde{I}_7$, where

$$\begin{aligned}
\tilde{I}_7 &= \frac{1}{T}\sum_{k_1,k_2} \big|\frac{\partial a_{k_2}}{\partial f}\big|\mathrm{E}^*\big|I_{k_1,r}^{6*}\big|^{1/6}\mathrm{E}^*|\widetilde{f}_{k_1}^{4*}|^{1/2}\mathrm{E}^*\big[\widetilde{I}_{k_2,r}^*\widetilde{f}_{k_2}^*\big] \quad \text{and} \\
\|\tilde{I}_7\|_1 &\leq \frac{1}{T}\sum_{k_1,k_2} \underbrace{\|\mathrm{E}^*\big|I_{k_1,r}^*\big|^6\|_1^{1/6}}_{(A.37)}\underbrace{\|\mathrm{E}^*(\widetilde{f}_{k_1}^{4*})\|_1^{1/2}}_{(A.36)}\underbrace{\big\|\mathrm{cov}^*[\widetilde{I}_{k_2,r}^*\widetilde{f}_{k_2}^*]\big\|_3}_{(A.41)}.
\end{aligned}$$

Finally we use identical arguments as above to show that $|I_8| = O_p(1)\tilde{I}_8$, where

$$\tilde{I}_8 \leq \frac{1}{T}\sum_{k_1,k_2} \mathrm{E}^*|\widetilde{f}_{k_1}^{4*}|^{1/2}\mathrm{E}^*\big|I_{k_1,r}^{6*}\big|^{1/6}\mathrm{E}^*|\widetilde{f}_{k_2}^{4*}|^{1/2}\mathrm{E}^*\big|I_{k_2,r}^{6*}\big|^{1/6}.$$

Thus, using similar arguments as those used to bound $\|\tilde{I}_3\|_1$, we have $|I_8| = O((b^2T)^{-1})$. Similar arguments can be used to obtain the same bounds for $I_4, \ldots, I_6$, which altogether gives (A.50).

To bound (A.51), we write $I_{k,r}$ as $I_{k,r} = \tilde{I}_{k,r} + \mathrm{E}(I_{k,r})$ and substitute this in the difference to give

$$\begin{aligned}
& T\big(\mathrm{E}^*|\widehat{c}^*(r,0)\big|^2 - \big|\mathrm{E}^*|\check{c}^*(r,0)|^2\big) = \frac{1}{T}\sum_{k_1,k_2}\Big(\mathrm{E}^*\big[a_{k_1}^*a_{k_2}^*I_{k_1,r}^*I_{k_2,r}^*\big] - \widehat{a}_{k_1}\widehat{a}_{k_1}\mathrm{E}^*[I_{k_1,r}^*I_{k_2,r}^*]\Big) \\
= & \frac{1}{T}\sum_{k_1,k_2}\Big(\mathrm{E}^*\big[2\tilde{I}_{k_1,r}^*\mathrm{E}(I_{k_2,r}^*) - \mathrm{E}^*(I_{k_1,r}^*)\mathrm{E}^*(I_{k_2,r}^*)a_{k_1}^*a_{k_2}^*\big] - \widehat{a}_{k_1}\widehat{a}_{k_1}\mathrm{E}^*[2\tilde{I}_{k_1,r}^*\mathrm{E}(I_{k_2,r}^*) - \mathrm{E}^*(I_{k_1,r}^*)\mathrm{E}^*(I_{k_2,r}^*)]\Big] \\
& + \frac{1}{T}\sum_{k_1,k_2}\Big(\mathrm{E}^*\big[a_{k_1}^*a_{k_2}^*\tilde{I}_{k_1,r}^*\tilde{I}_{k_2,r}^*\big] - \widehat{a}_{k_1}\widehat{a}_{k_1}\mathrm{E}^*[\tilde{I}_{k_1,r}^*\tilde{I}_{k_2,r}^*]\Big].
\end{aligned}$$

We now substitute the Taylor expansion into (A.54) to get

$$T\big(\mathrm{E}^*|\widehat{c}^*(r,0)\big|^2 - \big|\mathrm{E}^*|\breve{c}^*(r,0)|^2\big) = \sum_{i=0}^{8} II_i, \tag{A.56}$$

where

$$II_0 = \frac{2}{T}\sum_{k_1,k_2}\Big(2\widetilde{I}^*_{k_1,r}\mathrm{E}(I^*_{k_2,r}) - \mathrm{E}^*(I^*_{k_1,r})\mathrm{E}^*(I^*_{k_2,r})\Big)\Big(\widehat{a}_{k_1} + \widetilde{f}^*_{k_1}\frac{\partial\widehat{a}_{k_1}}{\partial f} + \widetilde{f}^{*2}_{k_1}\frac{1}{2}\frac{\partial^2\bar{a}^*_k}{\partial f^2}\Big)\Big(\widetilde{f}^*_{k_2}\frac{\partial\widehat{a}_{k_2}}{\partial f} + \widetilde{f}^{*2}_{k_2}\frac{1}{2}\frac{\partial^2\bar{a}^*_{k_2}}{\partial f^2}\Big),$$

$$\begin{aligned}
II_1 &= \frac{1}{T}\sum_{k_1,k_2}\widehat{a}_{k_1}\frac{\partial\widehat{a}_{k_2}}{\partial f}\mathrm{E}^*\big(\widetilde{I}^*_{k_1,r}\widetilde{I}^*_{k_2,r}\widetilde{f}^*_{k_2}\big),\\
II_2 &= \frac{1}{T}\sum_{k_1,k_2}\frac{\partial\widehat{a}_{k_1}}{\partial f}\frac{\partial\widehat{a}_{k_2}}{\partial f}\mathrm{E}^*\big[\widetilde{I}^*_{k_1,r}\widetilde{f}^*_{k_1}\widetilde{I}^*_{k_2,r}\widetilde{f}^*_{k_2}\big],\\
II_3 &= \frac{1}{2T}\sum_{k_1,k_2}\widehat{a}_{k_2}\mathrm{E}^*\Big[\widetilde{I}^*_{k_1,r}\widetilde{f}^{2*}_{k_1}\frac{\partial^2\bar{a}^*_{k_1}}{\partial f^2}\widetilde{I}^*_{k_2,r}\Big],\\
II_7 &= \frac{1}{T}\sum_{k_1,k_2}\frac{\partial\widehat{a}_{k_2}}{\partial f}\mathrm{E}^*\Big[\widetilde{I}^*_{k_1,r}\widetilde{f}^{2*}_{k_1}\frac{\partial^2\bar{a}^*_{k_1}}{\partial f^2}\widetilde{I}^*_{k_2,r}\widetilde{f}^*_{k_2}\Big],\\
II_8 &= \frac{1}{4T}\sum_{k_1,k_2}\mathrm{E}^*\Big[\widetilde{I}^*_{k_1,r}\widetilde{f}^{2*}_{k_1}\frac{\partial^2\bar{a}^*_{k_1}}{\partial f^2}\widetilde{I}^*_{k_2,r}\widetilde{f}^{2*}_{k_2}\frac{\partial^2\bar{a}^*_{k_2}}{\partial f^2}\Big]
\end{aligned}$$

and II_4, II_5, II_6 are defined similarly. By using similar methods to those used to bound (A.50), Assumption 5.2(B1), (A.36), (A.37) and (A.38), we can show that $|II_0| = O_p((Tp^2b)^{-1})$. To bound $|II_1|,\ldots,|III_8|$ we use the same methods as those used to bound (A.50) and the bound in (A.36), (A.37), (A.40), (A.41), (A.42) and (A.43) to show (A.51), we omit the details as they are identical to the proof of (A.50). □

Lemma A.17

Suppose that Assumption 5.2(B2) and the conditions in Lemma A.15 hold. Then, we have

$$T\Big(\mathrm{E}^*\big[\breve{c}^*_{j_1,j_2}(r_1,\ell_1)\big]\overline{\mathrm{E}^*\big[\breve{c}^*_{j_3,j_4}(r_2,\ell_2)\big]} - \big|\mathrm{E}^*\big[\tilde{c}^*_{j_1,j_2}(r_1,\ell_1)\big]\overline{\mathrm{E}^*\big[\tilde{c}^*_{j_3,j_4}(r_2,\ell_2)\big]}\Big) = O_p\left(a(T,b,p)\right), \tag{A.57}$$

$$T\Big(\mathrm{E}^*\big[\breve{c}^*_{j_1,j_2}(r_1,\ell_1)\overline{\breve{c}^*_{j_3,j_4}(r_2,\ell_2)}\big] - \mathrm{E}^*\big[\tilde{c}^*_{j_1,j_2}(r_1,\ell_1)\overline{\tilde{c}^*_{j_3,j_4}(r_2,\ell_2)}\big]\Big) = O_p\left(a(T,b,p)\right), \tag{A.58}$$

$$T\Big(\mathrm{E}^*\big[\acute{c}^*_{j_1,j_2}(r_1,\ell_1)\big]\overline{\mathrm{E}^*\big[\acute{c}^*_{j_3,j_4}(r_2,\ell_2)\big]} - \big|\mathrm{E}^*\big[\tilde{c}^*_{j_1,j_2}(r_1,\ell_1)\big]\overline{\mathrm{E}^*\big[\tilde{c}^*_{j_3,j_4}(r_2,\ell_2)\big]}\Big) = O_p\left(a(T,b,p)\right), \tag{A.59}$$

$$T\Big(\mathrm{E}^*\big[\acute{c}^*_{j_1,j_2}(r_1,\ell_1)\overline{\acute{c}^*_{j_3,j_4}(r_2,\ell_2)}\big] - \mathrm{E}^*\big[\tilde{c}^*_{j_1,j_2}(r_1,\ell_1)\overline{\tilde{c}^*_{j_3,j_4}(r_2,\ell_2)}\big]\Big) = O_p\left(a(T,b,p)\right), \tag{A.60}$$

where $a(T,b,p) = \frac{1}{bTp^2} + \frac{1}{Tp^4} + \frac{1}{b^2T} + b + \frac{1}{Tp^2} + \frac{1}{T^{1/2}p}$.

PROOF. Without loss of generality, we consider the case $d = 1$, $\ell_1 = \ell_2 = 0$ and $r_1 = r_2 = r$ and use the same notation introduced in the proof of Lemma A.16. To bound (A.57), we use the Taylor expansion

$$\begin{aligned}
& \widehat{a}_{k_1}\widehat{a}_{k_2} - \widehat{a}_{k_1}\widehat{a}_{k_2} \\
= \;& \tilde{f}_{k_2}\widehat{a}_{k_1}\frac{\partial \widehat{a}_{k_2}}{\partial f} + \tilde{f}_{k_1}\widehat{a}_{k_2}\frac{\partial \widehat{a}_{k_1}}{\partial f} + \frac{1}{2}\tilde{f}_{k_2}^2\widehat{a}_{k_1}\frac{\partial^2 \bar{a}_{k_2}}{\partial f^2} + \frac{1}{2}\tilde{f}_{k_2}\widehat{a}_{k_2}\frac{\partial^2 \bar{a}_{k_1}}{\partial f^2} + \tilde{f}_{k_1}\tilde{f}_{k_2}\frac{\partial \bar{a}_{k_2}}{\partial f}\frac{\partial \bar{a}_{k_1}}{\partial f},
\end{aligned} \tag{A.61}$$

which gives

$$T\big(\mathrm{E}^*\left[\breve{c}^*_{j_1,j_2}(r,0)\right]\overline{\mathrm{E}^*\left[\breve{c}^*_{j_3,j_4}(r,0)\right]} - \mathrm{E}^*\left[\tilde{c}^*_{j_1,j_2}(r,0)\right]\overline{\mathrm{E}^*\left[\tilde{c}^*_{j_3,j_4}(r,0)\right]}\big) = \sum_{i=1}^{3} III_i, \tag{A.62}$$

where

$$\begin{aligned}
III_1 &= \frac{2}{T}\sum_{k_1,k_2} a_{k_1}\frac{\partial a_{k_2}}{\partial f}\widetilde{f}_{k_2}\mathrm{E}^*(I^*_{k_1,r})\mathrm{E}^*(I^*_{k_2,r}), \\
III_2 &= \frac{1}{T}\sum_{k_1,k_2} a_{k_1}\frac{\partial \bar{a}^2_{k_2}}{\partial f^2}\widetilde{f}^2_{k_2}\mathrm{E}^*(I^*_{k_1,r})\mathrm{E}^*(I^*_{k_2,r}), \\
III_3 &= \frac{1}{T}\sum_{k_1,k_2} \widetilde{f}_{k_1}\widetilde{f}_{k_2}\frac{\partial \bar{a}_{k_1}}{\partial f}\frac{\partial \bar{a}_{k_2}}{\partial f}\mathrm{E}^*(I^*_{k_1,r})\mathrm{E}^*(I^*_{k_2,r}).
\end{aligned}$$

By using Lemma A.15, (A.37) and (A.47) and the same procedure used to bound (A.50), we obtain (A.57).

To bound (A.58), we use a similar decomposition to (A.62) to give

$$T\big(\left|\mathrm{E}^*|\breve{c}^*(r,0)\right|^2 - \left|\mathrm{E}^*|\widetilde{c}^*(r,0)\right|^2\big) = \sum_{i=0}^{3} IV_i,$$

where

$$\begin{aligned}
IV_0 = & -\frac{1}{T}\sum_{k_1,k_2}\left(2I^*_{k_1,r}\mathrm{E}^*(I^*_{k_2,r}) + \mathrm{E}(I^*_{k_1,r})\mathrm{E}^*(I^*_{k_2,r})\right)\left(\tilde{f}_{k_2}\widehat{a}_{k_1}\frac{\partial \widehat{a}_{k_2}}{\partial f} + \tilde{f}_{k_1}\widehat{a}_{k_2}\frac{\partial \widehat{a}_{k_1}}{\partial f} + \frac{1}{2}\tilde{f}^2_{k_2}\widehat{a}_{k_1}\frac{\partial^2 \bar{a}_{k_2}}{\partial f^2} + \right. \\
& \left. \frac{1}{2}\tilde{f}_{k_2}\widehat{a}_{k_2}\frac{\partial^2 \bar{a}_{k_1}}{\partial f^2} + \tilde{f}_{k_1}\tilde{f}_{k_2}\frac{\partial \bar{a}_{k_2}}{\partial f}\frac{\partial \bar{a}_{k_1}}{\partial f}\right),
\end{aligned}$$

$$\begin{aligned}
IV_1 &= \frac{2}{T}\sum_{k_1,k_2} a_{k_1}\frac{\partial a_{k_2}}{\partial f}\widetilde{f}_{k_2}\mathrm{E}^*(\tilde{I}^*_{k_1,r}\tilde{I}^*_{k_2,r}), \\
IV_2 &= \frac{1}{T}\sum_{k_1,k_2} a_{k_1}\frac{\partial \bar{a}^2_{k_2}}{\partial f^2}\widetilde{f}^2_{k_2}\mathrm{E}^*(\tilde{I}^*_{k_1,r}\tilde{I}^*_{k_2,r}), \\
IV_3 &= \frac{1}{T}\sum_{k_1,k_2} \widetilde{f}_{k_1}\widetilde{f}_{k_2}\frac{\partial \bar{a}_{k_1}}{\partial f}\frac{\partial \bar{a}_{k_2}}{\partial f}\mathrm{E}^*(\tilde{I}^*_{k_1,r}\tilde{I}^*_{k_2,r}).
\end{aligned}$$

Again using the same methods to bound (A.50), Lemma A.15 (A.38), (A.41) and (A.47) we obtain $IV_i = O_p(b + \frac{1}{Tp^2} + \frac{1}{T^{1/2}p} + \frac{1}{Tp^4})$, and thus (A.58).

To bound (A.59) and (A.60) we use identical methods to those given above, hence we omit the details. $\square$

PROOF of Lemma 5.2 We will prove (ii), the proof of (i) is similar. We observe that

$$\begin{aligned}
& T\Big|\mathrm{cov}^*(\widehat{c}^*_{j_1,j_2}(r,0_1), \widehat{c}^*_{j_3,j_4}(r,0_2)) - \mathrm{cov}^*(\widetilde{c}^*_{j_1,j_2}(r,0_1), \widetilde{c}^*_{j_3,j_4}(r,0_2))\Big| \\
\leq\ & T\Bigg(\mathrm{E}^*\big(\widehat{c}^*_{j_1,j_2}(r,0_1)\overline{\widehat{c}^*_{j_3,j_4}(r,0_2)}\big) - \mathrm{E}^*\big(\widetilde{c}^*_{j_1,j_2}(r,0)\overline{\widetilde{c}^*_{j_3,j_4}(r,0_2)}\big)\Bigg) \\
& +T\Bigg(\mathrm{E}^*(\widehat{c}^*_{j_1,j_2}(r,0))\overline{\mathrm{E}^*(\widehat{c}^*_{j_3,j_4}(r,0_2))} - \mathrm{E}^*(\widetilde{c}^*_{j_1,j_2}(r,0))\overline{\mathrm{E}^*(\widetilde{c}^*_{j_3,j_4}(r,0_2))}\Bigg).
\end{aligned}$$

Substituting (A.50)-(A.58) into the above gives the bound $O_p(a(T,b,p))$. By using a similar method, we can show

$$T\Big|\mathrm{cov}^*(\widehat{c}^*_{j_1,j_2}(r,0_1), \overline{\widehat{c}^*_{j_3,j_4}(r,0_2)}) - \mathrm{cov}^*(\widetilde{c}^*_{j_1,j_2}(r,0_1), \overline{\widetilde{c}^*_{j_3,j_4}(r,0_2)})\Big| = O_p(a(T,b,p)).$$

Together, these two results give the bounds in Lemma 5.2. $\square$

PROOF of Theorem 5.2 The proof in the fourth order stationary case follows by using that $\mathbf{W}^*_n$ is a consistent estimator of $\mathbf{W}_n$ (see Theorem 5.1 and Lemma 5.2) therefore

$$|\mathcal{T}^*_{m,n,d} - \mathcal{T}_{m,n,d}| \overset{P}{\to} 0.$$

Since $\mathcal{T}_{m,n,d}$ is asymptotically a chi-squared (see Theorem 3.4). Thus we have proven (i)

To prove (ii), we need to consider the case that $\{\underline{X}_t\}$ is locally stationary with $A_n(r,\ell) \neq 0$. From Theorem 3.6, we know that $\sqrt{T}(\Re\widehat{\mathbf{K}}_n(r) - \underbrace{\Re\mathbf{A}_n(r)}_{O(1)} - \underbrace{\Re\mathbf{B}_n(r)}_{O(b)})$ is asymptotically normal with mean zero. Therefore, since $\mathbf{W}^*_n = O(p^{-1})$, we have $(\mathbf{W}^*_n)^{-1/2} = O(p^{1/2})$. This altogether gives

$$|\sqrt{T}(\mathbf{W}^*_n)^{-1/2}\Re\widehat{\mathbf{K}}_n(r)|^2 + |\sqrt{T}(\mathbf{W}^*_n)^{-1/2}\Im\widehat{\mathbf{K}}_n(r)|^2 = O_p(Tp),$$

and thus the required result. $\square$

References

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

Billingsley, P. (1995). *Probability and measure* (Third ed.). New York: Wiley.

Bloomfield, P., Hurd, H., & Lund, R. (1994). Periodic correlation in stratospheric data. *Journal of Time Series Analysis*, *15*, 127-150.

Bousamma, F. (1998). *Ergodicité, mélange et estimation dans les modèles GARCH.* Unpublished doctoral dissertation, Paris 7.

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

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Annals of Statistics*, *16*, 1-37.

Dahlhaus, R. (2012). Handbook of statistics, volume 30. In S. Subba Rao, T. Subba Rao & C. Rao (Eds.), (p. 351-413). Elsevier.

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

Dahlhaus, R., & Polonik, W. (2009). Empirical spectral processes for locally stationary time series. *Bernoulli*, *15*, 1-39.

Dette, H., Preuss, P., & Vetter, M. (2011). A measure of stationarity in locally stationary processes with applications to testing. *Journal of the Americal Statistical Association*, *106*, 1113-1124.

Dwivedi, Y., & Subba Rao, S. (2011). A test for second order stationarity based on the Discrete Fourier Transform. *Journal of Time Series Analysis*, *32*, 68-91.

Eichler, M. (2012). Graphical modelling of multivariate time series. *Probability Theory and Related Fields*, *153*, 233-268.

Escanciano, J. C., & Lobato, I. N. (2009). An automatic Portmanteau test for serial correlation. *Journal of Econometrics*, *151*, 140-149.

Feigin, P., & Tweedie, R. L. (1985). Random coefficient Autoregressive processes: A Markov chain analysis of stationarity and finiteness of moments. *Journal of Time Series Analysis*, *6*, 1-14.

Franke, J., Stockis, J. P., & Tadjuidje-Kamgaing, J. (2010). On the geometric ergodicity of CHARME models. *Journal of Time Series Analysis*, *31*, 141-152.

Fryzlewicz, P., & Subba Rao, S. (2011). On mixing properties of ARCH and time-varying ARCH processes. *Bernoulli*, *17*, 320-346.

Goodman, N. R. (1965). *Statistical test for stationarity within the framework of harmonizable processes* (Tech. Rep. No. 65-28 (2nd August)). Office for Navel Research.

Hallin, M. (1984). Spectral factorization of nonstationary moving average processes. *Annals of Statistics*, *12*, 172-192.

Hosking, J. R. M. (1980). The multivariate Portmanteau statistic. *Journal of the American Statistical Association*, *75*, 602-608.

Hosking, J. R. M. (1981). Lagrange-tests of multivariate time series models. *Journal of the Royal Statistical Society (B)*, *43*, 219-230.

Hurd, H., & Gerr, N. (1991). Graphical methods for determining the presence of periodic correlation. *Journal of Time Series Analysis*, *12*, 337-350.

Jentsch, C. (2012). A new frequency domain approach of testing for covariance stationarity and for periodic stationarity in multivariate linear processes. *Journal of Time Series Analysis*, *33*, 177-192.

Lahiri, S. N. (2003). *Resampling methods for dependent data.* New York: Springer.

Laurent, S., Rombouts, J., & Violante, F. (2011). On the forecasting accuracy of multivariate GARCH models. *To appear in Journal of Applied Econometrics*.

Lee, J., & Subba Rao, S. (2011). A note on general quadratic forms of nonstationary processes. *Preprint*.

Lei, J., Wang, H., & Wang, S. (2012). A new nonparametric test for stationarity in the time domain. *Preprint*.

Loretan, M., & Phillips, P. C. B. (1994). Testing the covariance stationarity of heavy tailed time series. *Journal of Empirical Finance*, *1*, 211-248.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis.* Berlin: Springer.

Meyn, S. P., & Tweedie, R. (1993). *Markov chains and stochastic stability.* London: Springer.

Mokkadem, A. (1990). Propertiés de mélange des processus autoregressifs polnomiaux. *Annales de l'I H. P.*, *26*, 219-260.

Nason, G. (2013). A test for second-order stationarity and approximate confidence intervals for localized autocovariances for locally stationary time series. *Journal of the Royal Statistical Society, Series (B)*, *75*.

Neumann, M. (1996). Spectral density estimation via nonlinear wavelet methods for stationary nonGaussian time series. *J. Time Series Analysis*, *17*, 601-633.

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

Paparoditis, E. (2010). Validating stationarity assumptions in time series analysis by rolling local periodograms. *Journal of the American Statistical Association*, *105*, 839-851.

Paparoditis, E., & Politis, D. N. (2004). Residual-based block bootstrap for unit root testing. *Econometrica*, *71*, 813-855.

Parker, C., Paparoditis, E., & Politis, D. N. (2006). Unit root testing via the stationary bootstrap. *Journal of Econometrics*, *133*, 601-638.

Parzen, E. (1999). *Stochastic processes.* San Francisco: SIAM.

Patton, A., Politis, D., & White, H. (2009). Correction to "automatic block-length selection for the dependent bootstrap" by D. Politis and H. White',. *Econometric Reviews*, *28*, 372-375.

Pelagatti, M. M., & Sen, K. S. (2013). Rank tests for short memory stationarity. *Journal of Econometrics*, *172*, 90-105.

Pham, T. D., & Tran, L. T. (1985). Some mixing properties of time series models. *Stochastic processes and their applicatons*, *19*, 297-303.

Politis, D., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, *89*, 1303-1313.

Politis, D., & White, H. (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, *23*, 53-70.

Priestley, M. B. (1965). Evolutionary spectral and nonstationary processes. *Journal of the Royal Statistical Society (B)*, *27*, 385-392.

Priestley, M. B., & Subba Rao, T. (1969). A test for non-stationarity of a time series. *Journal of the Royal Statistical Society (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.

Robinson, P. (1991). Automatic frequency inference on semiparametric and nonparametric models. *Econometrica*, *59*, 1329-1363.

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

Subba Rao, T. (1970). The fitting of non-stationary time-series models with time-dependent parameters. *Journal of the Royal Statistical Society* B, *32*, 312-22.

Subba Rao, T., & Gabr, M. M. (1984). *An introduction to bispectral analysis and bilinear time series models* (Vol. 24). Berlin: Springer.

Terdik, G. (1999). *Bilinear stochastic models and related problems of nonlinear time series analysis* (Vol. 142). Berlin: Springer.

Vogt, M. (2012). Nonparametric regression with locally stationary time series. *Annals of Statistics*, *40*, 2601-2633.

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

Walker, A. M. (1963). Asymptotic properties of least square estimate of the parameters of the spectrum of a stationary non-deterministic time series. *J. Austral. Math Soc.*, *4*, 363-384.

Whittle, P. (1953). The analysis of multiple stationary time series. *Journal of the Royal Statistical Society (B)*.