A test for second order stationarity of a multivariate time series

Carsten Jentsch∗ and Suhasini Subba Rao†

December 27, 2012

Abstract

It is well known that the discrete Fourier transform (DFT) of a second order stationary time series between two distinct Fourier frequencies are asymptotically uncorrelated. In contrast we show that for a large class of second order nonstationary time series, including locally stationary time series, the covariance between the Fourier frequencies is non-zero. Indeed the correlations between the DFTs contain important information about the underlying nonstationary time series. In this paper we use these starkly differing properties to define a global test for stationarity based on the DFT of a vector time series. We show under the null of stationarity the test statistic has a chi-squared distribution and under the alternative of local stationarity a noncentral chi-squared distribution. In addition, under the assumption of Gaussianity of the time series, the test statistic is pivotal. However, in many econometric applications, such assumptions can be too strong, therefore the proposed test does not make any model assumptions (such as linearity) on the underlying time series. Under these general conditions the test statistic involves an unknown variance that is extremely difficult to directly estimate from the data. To overcome this issue, we propose a scheme based on the stationary bootstrap to estimate the unknown variance and show that the resulting estimator is a consistent estimator of the variance.

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

1 Introduction

In several disciplines, including the finance, geosciences 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, vector GARCH, 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), Paparaditis (2009, 2010), Dwivedi and Subba Rao (2011), Dette, Preuss, and Vetter (2011), Lei, Wang, and Wang (2012) and Jentsch (2012). 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 detect the alternative of a periodic multivariate stationary time series). One solution is to individually test for stationarity for each of the univariate processes. In addition to the problematic multiple testing issue, this strategy can lead to misleading conclusions. For example, consider the independent bivariate time series \( \{X_t\} \), where

\[
\text{var}(X_t) = \begin{pmatrix} 1 & \rho_t \\ \rho_t & 1 \end{pmatrix}.
\]

Each element of \( \{X_t\} \) is second-order stationary so a univariate test for stationarity on each element of the vector is unlikely to reject the null, however \( \{X_t\} \) is a nonstationary multivariate time series if \( \rho_t \) is not constant over time \( t \in \mathbb{Z} \). This simple example illustrates the 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 two possible disadvantages. The first is that locally the spectral density may 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. Secondly, for multivariate time series the spectral density estimator constructed over short periods of time will have a large mean squared error, leading to an unreliable test statistic. For these reasons, 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 linearity of the underlying time series, however, in many situations such an assumption is untenable, whereas our tests does not make any model assumptions, such as linearity, and allows for a wide class of time series.

In Section 2 we consider the covariance between the DFTs for stationary and nonstationary time series, focusing on the class of locally stationary time series defined in Dahlhaus (1997), Dahlhaus and Polonik (2006) (see Dahlhaus (2012) for a review). Based on these results we define the weighted DFT covariance and study its asymptotic sampling properties under second order stationarity and local stationarity. In particular, we show that under the null hypothesis the mean of the weighted DFT covariance is asymptotically zero. Whereas, under the alternative of local stationarity we show that the weighted DFT covariance estimates nonstationary characteristics in the time series. In Section 3 we propose the test statistic, which is based on the properties of the weighted DFT covariance, derived in Section 2. However, the resulting test statistic involves a variance which for non-Gaussian time series models is unknown and in
practice is extremely difficult to estimate. A test that does not correct for this variance, can lead to unreliable results, therefore we propose a bootstrap procedure to estimate the variance and show that the estimator is a consistent estimator of the variance.

Finally, in Section 5 we illustrate our test through an extensive simulation study. The performance of the test under the null and alternative hypotheses are considered, using both linear and nonlinear models and a wide range of nonstationary time series models including change point, slowly varying and unit root models. In our simulations, we compare the performance of the test statistic when standardised with the bootstrap variance and excluding the bootstrap variance correction. We show that for non-Gaussian models, omitting the bootstrap variance can lead to too many false positives. But by correcting for the variance, our test is able to detect a wide range of alternatives even when the sample size is relatively small.

The proofs can be found in the Appendix.

2 The weighted DFT covariance

2.1 Motivation

Let \{X_t = (X_{t,1}, \ldots, X_{t,d})', t \in \mathbb{Z}\} be a d-dimensional zero mean, multivariate time series and suppose we observe \{X_t\}_{t=1}^{\infty}. We define the vector discrete Fourier transform (DFT) as

\[ J_T(\omega_k) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} X_t e^{-it\omega_k}, \quad k = 1, \ldots, T, \]

where \( \omega_k = 2\pi k \) are the Fourier frequencies. Suppose that \{\overrightarrow{X}_t\} is a second order stationary multivariate time series, where all the entries of the autocovariance matrices of \{\overrightarrow{X}_t\} satisfy

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

Then, it is well known for \( k_1 - k_2 \neq Ts, s \in \mathbb{Z}, \) that \( \text{Cov}(J_{T,m}(\omega_{k_1}), J_{T,n}(\omega_{k_2})) = O\left(\frac{1}{T}\right)\), in other words the DFT has transformed a time series into a sequence which is approximately uncorrelated. The behaviour 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 devise introduced by Dahlhaus (1997) to study a large class of nonstationary time series, called locally stationary time series. More precisely, \{\overrightarrow{X}_{t,T}\} is a locally second order stationary time series, if its covariance structure changes slowly over time such that \( \text{cov}(\overrightarrow{X}_{t,T}, \overrightarrow{X}_{\tau,T}) = \kappa(\frac{1}{T}; t - \tau) + O\left(\frac{1}{|t - \tau|^{1+\varepsilon}}\right) \) (for some \( \varepsilon > 0 \)), and \( \kappa(u; r) \) is a smooth matrix function defined over \([0, 1]\). Using Lee and Subba Rao (2011), it can be shown that for such processes the covariance between the DFTs is

\[ \text{Cov}(J_{T,m}(\omega_{k_1}), J_{T,n}(\omega_{k_2})) = \int_0^1 f(u; \omega_{k_1}) \exp(i2\pi u(k_1 - k_2))du + O\left(\frac{1}{T}\right), \quad (2.1) \]

where \( f(u; \omega) = \frac{1}{2\pi \sum_{r=\infty}^{\infty} \kappa(u; r) \exp(-ir\omega)} \) is the local spectral density matrix. We observe if \( f(u; \omega_{k_1}) \) does not depend on \( u \) (thus \{\overrightarrow{X}_t\} is a second order stationary time series), then the above expression reduces to \( \text{Cov}(J_{T,m}(\omega_{k_1}), J_{T,n}(\omega_{k_2})) = O\left(\frac{1}{T}\right) \) for \( k_1 - k_2 \neq Ts, s \in \mathbb{Z}. \)

Moreover, the DFT mimics the behaviour of a time series, in the sense that the correlation decays the further apart the frequencies \( \omega_{k_1} \) and \( \omega_{k_2} \) (for \( 1 \leq k_1, k_2 \leq T/2 \)) are. Equation (2.1) 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 Defining the weighted DFT covariance

The discussion in the previous section suggests that to test for stationarity, we can equivalently consider the problem in the frequency domain and test for uncorrelatedness of the vector DFT, \( \{ L_T(\omega_k) \} \). Testing for uncorrelatedness of a multivariate time series is a well established technique in time series analysis (see, for example, Hoskings (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 matrices of the DFT, \( \{ L_T(\omega_k) \} \), and use this as the basis of the proposed test statistic.

Under suitable conditions on the second order stationary process \( \{ X_t \} \), we have \( \mathbb{E}(L_T(\omega_k)) = 0 \) and \( \text{var}(L_T(\omega_k)) \to f(\omega_k) \) as \( T \to \infty \), where \( f : [0, 2\pi] \to \mathbb{C}^{d \times d} \) with

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

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

\[
f(\omega) = B(\omega)\overline{B(\omega)}',
\]

where \( B(\omega) \) is a lower triangular matrix and \( \overline{B(\omega)}' \) denotes the transpose and complex conjugate of \( B(\omega) \). Let \( L(\omega_k) := B^{-1}(\omega_k) \), thus \( f^{-1}(\omega_k) = \overline{L(\omega_k)}L(\omega_k) \). Therefore, if \( \{ X_t \} \) is second order stationary the vector sequence, \( \{ L(\omega_1)L_T(\omega_1), \ldots, L(\omega_T)L_T(\omega_T) \} \), is approximately an uncorrelated sequence with the identity matrix as its variance.

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

\[
\hat{f}_T(\omega) = \frac{1}{T} \sum_{j=-\left\lfloor \frac{T}{2} \right\rfloor}^{\left\lfloor \frac{T}{2} \right\rfloor} K_b(\omega - \omega_j)I_T(\omega_j), \quad \omega \in [0, 2\pi],
\]

where \( I_T(\omega) = \{ L_T(\omega), \overline{L_T(\omega)}' \} \), \( |x| \) is the integer part of \( x \in \mathbb{R} \), \( K \) is a kernel function, \( b \) is the bandwidth and \( K_b(\cdot) = \frac{1}{b}K\left(\frac{\cdot}{b}\right) \). Let \( \hat{f}_T(\omega_k) \) be \( \hat{B}(\omega_k)\overline{B(\omega_k)}' \), where \( \hat{B}(\omega_k) \) is the (lower-triangular) Cholesky decomposition of \( \hat{f}(\omega_k) \) and \( \hat{L}(\omega_k) := \hat{B}^{-1}(\omega_k) \). Thus \( \hat{B}(\omega_k) \) and \( \hat{L}(\omega_k) \) are estimators of \( B(\omega_k) \) and \( L(\omega_k) \) respectively.

Using the above notation we now define the (standardized) weighted DFT covariance matrix at lags \( r \) and \( \ell \)

\[
\hat{C}_T(r, \ell) = \frac{1}{T} \sum_{k=1}^{T} \hat{L}(\omega_k)L_T(\omega_k)L_T(\omega_k+r)\overline{L(\omega_k+r)} \exp(i\ell\omega_k).
\]

The reason for including the exponential weight \( \exp(i\ell\omega_k) \) in the above definition is not immediately obvious, but it is analogous to the representation of the sample covariance of a stationary time series in terms of the integrated periodogram weighted with the exponential function. The role of \( \exp(i\ell\omega_k) \) will become clear in Section 2.5, when we show that \( \{ \hat{C}_T(r, \ell) \} \) can be used as a characterisation of the nonstationarity. We observe that due to the periodicity of the DFT, \( \hat{C}_T(r, \ell) \) is periodic in \( r \), where \( \hat{C}_T(r, \ell) = \hat{C}_T(r + T, \ell) \) for all \( r \in \mathbb{Z} \).
\section{2.3 Behaviour of $\widehat{C}_T(r, \ell)$ under stationarity}

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

$$\widehat{C}_T(r, \ell) = \frac{1}{T} \sum_{k=1}^{T} L(\omega_k) \overline{L}(\omega_k) \overline{L}(\omega_{k+r}) L(\omega_{k+r}) \exp(i\ell \omega_k). \quad (2.5)$$

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

\subsection*{2.3.1 Assumptions}

Let $| \cdot |_p$ denote the $\ell_p$-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 $\|X\|_p = (\mathbb{E}|X|^p)^{1/p}$.

**Assumption 2.1 (The process $\{X_t\}$)**

\begin{enumerate}
  \item [(P1)] Let us suppose that $\{X_t, t \in \mathbb{Z}\}$ is a $d$-variate zero mean, fourth order stationary, $\alpha$-mixing time series which satisfies
    $$\sup_{k \in \mathbb{Z}} \sup_{A, B \in \mathcal{P}(X_{t+k}, X_{t+k+1}, \ldots)} |P(A \cap B) - P(A)P(B)| \leq Ct^{-\alpha}, \quad t > 0$$
    where $C$ is a constant and $\alpha > 0$.
  \item [(P2)] For some $s > \frac{4\alpha}{\alpha - 0} > 0$, we have $\sup_{t \in \mathbb{Z}} \|X_t\|_s < \infty$.
  \item [(P3)] The $(d \times d)$ spectral density matrix $f(\omega)$ is non-singular on $[0, 2\pi]$.
  \item [(P4)] There exists an $s > \frac{8\alpha}{\alpha - 0} > 0$ such that $\sup_{t \in \mathbb{Z}} \|X_t\|_s < \infty$.
  \item [(P5)] The variance matrix $W_n$ introduced in Theorem 2.3, equation (2.14) is positive definite.
\end{enumerate}

**Assumption 2.2 (The kernel function $K$)**

\begin{enumerate}
  \item [(K1)] Let $\lambda_b(r) = \frac{1}{b} \sum_k K_b(\omega_k) \exp(-ir\omega_k)$, then $\sum_r |\lambda_b(r)| = O(b^{-1})$ and $\sum_r |r| \cdot |\lambda_b(r)| = O(b^{-2})$.
  \item [(K2)] $T^{-1/2} < b < T^{-1/4}$.
\end{enumerate}

Some comments on the assumptions are in order. The $\alpha$-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 & Tweedie, 1985), Mokkadem (1990), Meyn and Tweedie (1993), Bousamma (1998), Franke, Stockis, and Tadjuidje-Kamgaing (2010)). We show in Lemma A.3 that Assumption (P2) implies $\sum_{r=-\infty}^{\infty} |r| \cdot |\text{Cov}(X_{r,j_1}, X_{0,j_2})| < \infty$ for all $j_1, j_2 = 1, \ldots, d$ and absolute summability of the fourth order cumulants. Under Assumption (P1-P3) we will derive the limiting variance of $\widehat{C}_T(r, \ell)$ and show asymptotic normality of $\widehat{C}_T(r, \ell)$ (see Theorem 2.3). Assumption (P4) is slightly stronger than (P2) and it is used to show the asymptotic equivalence of $\sqrt{T}\overline{C}_T(r, \ell)$ and $\sqrt{T}\widehat{C}_T(r, \ell)$. In the case that the multivariate time series $\{X_t\}$ is geometric mixing, Assumption (P4) implies that for some $\delta > 0$, $(8 + \delta)$-moments of $\{X_t\}$ should exist. Assumption (P5) is immediately satisfied in the case that $\{X_t\}$ is a Gaussian time series, where $W_n$ becomes diagonal.
Remark 2.1 (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 \( \{X_t\} \) (see Theorem 2.3). The main advantage of this slightly stronger assumption is that it guarantees that the DFT covariances \( \tilde{C}_T(r_1, \ell) \) and \( \tilde{C}_T(r_2, \ell) \) are asymptotically uncorrelated at different lags \( r_1 \neq r_2 \).

2.4 The sampling properties of \( \tilde{C}_T(r, \ell) \)
Using the above assumptions we have the following result.

Theorem 2.1 (Asymptotical equivalence of \( \tilde{C}_T(r, \ell) \) and \( \tilde{C}_T(r, \ell) \))
Suppose Assumption 2.1 is satisfied and let \( \tilde{C}_T(r, \ell) \) and \( \tilde{C}_T(r, \ell) \) be defined as in (2.4) and (2.5), respectively. Then we have

\[
\sqrt{T} | \tilde{C}_T(r, \ell) - \tilde{C}_T(r, \ell) |_1 = O_p \left( \frac{1}{b \sqrt{T}} + b^2 \frac{1}{\sqrt{N}} \right).
\]

We now obtain the mean and variance of \( \tilde{C}_T(r, \ell) \) under the stated assumptions. Let \( \tilde{c}_{j_1,j_2}(r, \ell) = \tilde{C}_T(r, \ell)_{j_1,j_2} \) and \( \tilde{c}_{j_1,j_2}(r, \ell) = \tilde{C}_T(r, \ell)_{j_1,j_2} \) denote entry \((j_1,j_2)\) of the \((d \times d)\) DFT covariance matrices \( \tilde{C}_T(r, \ell) \) and \( \tilde{C}_T(r, \ell) \), respectively. In order to derive the properties we use that \( \tilde{c}_{j_1,j_2}(r, \ell) \) can be written as

\[
\tilde{c}_{j_1,j_2}(r, \ell) = \frac{1}{T} \sum_{k=1}^T L_{j_1, \bullet}(\omega_k)J_T(\omega_k)J_T(\omega_k+\ell)\cdot L_{j_2, \bullet}(\omega_k+\ell) \exp(i\ell\omega_k)
\]

\[
= \frac{1}{T} \sum_{k=1}^T \sum_{s_1,s_2=1}^d L_{j_1,s_1}(\omega_k)J_{T,s_1}(\omega_k)J_{T,s_2}(\omega_k+\ell)L_{j_2,s_2}(\omega_k+\ell) \exp(i\ell\omega_k),
\]

where \( L_{j,s}(\omega_k) \) is entry \((j,s)\) of \( L(\omega_k) \) and \( L_{j,\bullet}(\omega_k) \) denotes its \( j \)th row. Further, in order to obtain the variance of the above, we define

\[
\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4)
\]

\[
= \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} L_{j_1,s_1}(\lambda_1)L_{j_2,s_2}(\lambda_1)L_{j_3,s_3}(\lambda_2)L_{j_4,s_4}(\lambda_2) \exp(i\ell_1\lambda_1 - i\ell_2\lambda_2)
\]

\[
\times f_{4;\lambda_1,\lambda_2,\lambda_3}(\lambda_1, -\lambda_1, -\lambda_2)d\lambda_1d\lambda_2,
\]

where

\[
f_{4;\lambda_1,\lambda_2,\lambda_3}(\lambda_1, -\lambda_1, -\lambda_2) = \frac{1}{(2\pi)^3} \sum_{t_1,t_2,t_3=-\infty}^{\infty} \kappa^{(\lambda_1,\lambda_2,\lambda_3)}(t_1,t_2,t_3) \exp(i(-t_1\lambda_1 - t_2\lambda_2 - t_3\lambda_3))
\]

is the joint tri-spectra of \( \{X_t\} \) and \( \kappa^{(\lambda_1,\lambda_2,\lambda_3)}(t_1,t_2,t_3) = \text{cum}(X_t,X_{t_1},X_{t_2},X_{t_3}) \) for statistical properties of the tri-spectra see Brillinger (1981), Subba Rao and Gabr (1984) and Terdik (1999). The following symmetry properties will be useful in obtaining an expression for the variance of \( \tilde{C}_T(r, \ell) \)

\[
\kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4) = \kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4) = \kappa^{(\ell_2,\ell_1)}(j_3,j_4,j_1,j_2) = \kappa^{(-\ell_1,-\ell_2)}(j_2,j_1,j_4,j_3).
\]

The above implies that \( \kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4) \) is always real-valued.
Theorem 2.2 (First and second order structure of \( \{ \hat{C}_T(r, \ell) \} \))

Suppose Assumption 2.1(P1,P2) is satisfied. Then, the following assertions are true

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

(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

\[
TCov(\Re \tilde{c}_{j_1,j_2}(r_1, \ell_1), \Im \tilde{c}_{j_3,j_4}(r_2, \ell_2)) = O\left(\frac{1}{T}\right),
\]

(2.8)

\[
TCov(\Re \tilde{c}_{j_1,j_2}(r_1, \ell_1), \Re \tilde{c}_{j_3,j_4}(r_2, \ell_2)) = \frac{1}{2} \{ \delta_{j_1,j_2} \delta_{j_3,j_4} \delta_{\ell_1,\ell_2} + \delta_{j_1,j_4} \delta_{j_2,j_3} \delta_{\ell_1,-\ell_2} \} \delta_{r_1,r_2} + \frac{1}{2} \kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4) \delta_{r_1,r_2} + O\left(\frac{1}{T}\right),
\]

(2.9)

and

\[
TCov(\Im \tilde{c}_{j_1,j_2}(r_1, \ell_1), \Im \tilde{c}_{j_3,j_4}(r_2, \ell_2)) = \frac{1}{2} \{ \delta_{j_1,j_2} \delta_{j_3,j_4} \delta_{\ell_1,\ell_2} + \delta_{j_1,j_4} \delta_{j_2,j_3} \delta_{\ell_1,-\ell_2} \} \delta_{r_1,r_2} + \frac{1}{2} \kappa^{(\ell_1,\ell_2)}(j_1,j_2,j_3,j_4) \delta_{r_1,r_2} + O\left(\frac{1}{T}\right),
\]

(2.10)

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

By using the above results, we can show in the case of \( d = 2 \) and \( \ell = 0 \) that

\[
\text{var} \left( \sqrt{T} \text{vec} \left( \Re \hat{C}_T(r, 0) \right) \right) \rightarrow \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\
0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\
0 & 0 & 0 & 1 \\
\end{pmatrix} + \frac{1}{2} \begin{pmatrix}
\kappa^{(0,0)}(1,1,1,1) & \kappa^{(0,0)}(1,1,2,1) & \kappa^{(0,0)}(1,1,1,2) & \kappa^{(0,0)}(1,1,2,2) \\
\kappa^{(0,0)}(2,1,1,1) & \kappa^{(0,0)}(2,1,2,1) & \kappa^{(0,0)}(2,1,1,2) & \kappa^{(0,0)}(2,1,2,2) \\
\kappa^{(0,0)}(1,2,1,1) & \kappa^{(0,0)}(1,2,2,1) & \kappa^{(0,0)}(1,2,1,2) & \kappa^{(0,0)}(1,2,2,2) \\
\kappa^{(0,0)}(2,2,1,1) & \kappa^{(0,0)}(2,2,2,1) & \kappa^{(0,0)}(2,2,1,2) & \kappa^{(0,0)}(2,2,2,2) \\
\end{pmatrix},
\]

By (2.7), we can see that the second and third rows of the limiting variance of \( \sqrt{T} \hat{C}_T(r, 0) \) are identical, therefore \( \text{var} \left( \sqrt{T} \text{vec}(\Re \hat{C}_T(r, 0)) \right) \) becomes singular. This holds for any dimension \( d \).

Thus to remove this redundancy we transform it into a \( d(d + 1)/2 \)-dimensional vector, which is a vectorised version of its lower-triangular part. At this point, it is worth mentioning that our focus is on \( \hat{C}_T(r, \ell) \). As we have derived the asymptotic equivalence of \( \hat{C}_T(r, \ell) \) and \( \hat{C}_T(r, \ell) \) in Theorem 2.1 and the variance of \( \hat{C}_T(r, \ell) \), rather than transform the vector \( \hat{C}_T(r, \ell) \) we define a transformation of \( \hat{C}_T(r, \ell) \). That is

\[
\text{vech}(\hat{C}_T(r, \ell)) = (\widehat{c}_{1,1}(r, \ell), \widehat{c}_{2,1}(r, \ell), \ldots, \widehat{c}_{d,1}(r, \ell), \widehat{c}_{2,2}(r, \ell), \ldots, \widehat{c}_{d,2}(r, \ell), \ldots, \widehat{c}_{d,d}(r, \ell))'.
\]

In the theorem below we derive the limiting distribution of \( \text{vech}(\hat{C}_T(r, \ell)) \).

Theorem 2.3 (Asymptotic distribution of \( \text{vech}(\hat{C}_T(r, \ell)) \) under the null)

Suppose Assumptions 2.1 and 2.2 hold and \( \mathbf{W}_{\ell_1,\ell_2}^{(1)} \) and \( \mathbf{W}_{\ell_1,\ell_2}^{(2)} \) are \( (d(d + 1)/2 \times d(d + 1)/2) \)
matrices which are defined in (2.15) and (2.16), below. Then for all fixed \( r \in \mathbb{N} \) and \( \ell \in \mathbb{Z} \), we have
\[
\sqrt{T} \text{vech} \left( \mathbb{R} \mathbf{C}_T (r, \ell) \right) \overset{D}{\to} \mathcal{N} \left( \mathbf{0}_{d(d+1)/2}, \mathbf{W}_{\ell,\ell} \right) \quad \text{and} \quad \sqrt{T} \text{vech} \left( \mathbb{S} \mathbf{C}_T (r, \ell) \right) \overset{D}{\to} \mathcal{N} \left( \mathbf{0}_{d(d+1)/2}, \mathbf{W}_{\ell,\ell} \right)
\]
where \( \mathbf{0}_{d(d+1)/2} \) is the \( d(d+1)/2 \) zero vector and
\[
\mathbf{W}_{\ell,\ell} = \mathbf{W}_{\ell,\ell}^{(1)} + \mathbf{W}_{\ell,\ell}^{(2)}.
\]

Let the \( nd(d+1)/2 \)-dimensional vector \( \hat{\mathbf{K}}_n (r) \) be defined by
\[
\hat{\mathbf{K}}_n (r) = \left( \text{vech} (\mathbf{C}_T (r, 0))', \text{vech} (\mathbf{C}_T (r, 1))', \ldots, \text{vech} (\mathbf{C}_T (r, n-1))' \right)'.
\]
Then, for fixed \( m, n \in \mathbb{N} \), we have
\[
\sqrt{T} \begin{pmatrix} \mathbb{R} \hat{\mathbf{K}}_n (1) \\ \mathbb{S} \hat{\mathbf{K}}_n (1) \\ \vdots \\ \mathbb{R} \hat{\mathbf{K}}_n (m) \\ \mathbb{S} \hat{\mathbf{K}}_n (m) \end{pmatrix} \overset{D}{\to} \mathcal{N} \left( \mathbf{0}_{mn(d+1)}, \mathbf{W}_{m,n} \right),
\]
where \( \mathbf{W}_{m,n} \) is a \((mn(d+1) \times mn(d+1))\) block diagonal matrix
\[
\mathbf{W}_{m,n} = \text{diag} \left( \underbrace{\mathbf{W}_1, \ldots, \mathbf{W}_n}_{2m \text{ times}} \right),
\]
\( \mathbf{W}_n \) is an \((nd(d+1)/2 \times nd(d+1)/2)\) matrix consisting of \( n^2 \) blocks of dimension \( (d(d+1)/2 \times d(d+1)/2) \), with the \((\ell_1 + 1, \ell_2 + 1)\) block of \( \mathbf{W}_n \) is \((\mathbf{W}_n)_{\ell_1+1,\ell_2+1} = \mathbf{W}_{\ell_1,\ell_2}^{(1)} \delta_{\ell_1,\ell_2} + \mathbf{W}_{\ell_1,\ell_2}^{(2)} \).

The theorem above shows that the variance of the joint distribution of \( \{ \hat{\mathbf{K}}_n (r) \}_{r=1}^n \) is determined by \( (\mathbf{W}_n)_{\ell_1,\ell_2} \). Since, \( \mathbf{W}_n \) plays an important role in the test statistic, we obtain expressions for \( \mathbf{W}_n \) through a few examples and then derive a general expression.

Example 2.1 (The structure of \( \mathbf{W}_n = \{(\mathbf{W}_n)_{\ell_1,\ell_2}; 0 \leq \ell_1, \ell_2 \leq n-1\} \) for \( d = 2 \) and \( d = 3 \))

(i) For \( d = 2 \) and any \( n \in \mathbb{N} \), we have
\[
(\mathbf{W}_n)_{\ell_1,\ell_2} = \mathbf{W}_{\ell_1,\ell_2}^{(1)} \delta_{\ell_1,\ell_2} + \mathbf{W}_{\ell_1,\ell_2}^{(2)}
\]
\[
= \frac{1}{2} \begin{pmatrix} 1 + \delta_{\ell_1,0} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 + \delta_{\ell_1,0} \end{pmatrix} \delta_{\ell_1,\ell_2} + \frac{1}{2} \begin{pmatrix} \kappa_{\ell_1,\ell_2}(1,1,1,1) & \kappa_{\ell_1,\ell_2}(1,1,2,1) & \kappa_{\ell_1,\ell_2}(1,1,2,2) \\ \kappa_{\ell_1,\ell_2}(2,1,1,1) & \kappa_{\ell_1,\ell_2}(2,1,2,1) & \kappa_{\ell_1,\ell_2}(2,1,2,2) \\ \kappa_{\ell_1,\ell_2}(2,2,1,1) & \kappa_{\ell_1,\ell_2}(2,2,2,1) & \kappa_{\ell_1,\ell_2}(2,2,2,2) \end{pmatrix}.
\]

(ii) For \( d = 3 \) and any \( n \in \mathbb{Z}^+ \), we have
\[
(\mathbf{W}_n)_{\ell_1,\ell_2} = \mathbf{W}_{\ell_1,\ell_2}^{(1)} \delta_{\ell_1,\ell_2} + \mathbf{W}_{\ell_1,\ell_2}^{(2)}
\]
\[
= \frac{1}{2} \begin{pmatrix} 1 + \delta_{\ell_1,0} & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 + \delta_{\ell_1,0} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 + \delta_{\ell_1,0} \end{pmatrix} \delta_{\ell_1,\ell_2} + \mathbf{W}_{\ell_1,\ell_2}^{(2)}.
\]
In general, the matrix $W_{\ell_1, \ell_2}^{(1)}$ is a non-singular diagonal matrix with diagonal entries given by

$$
(W_{\ell_1, \ell_2}^{(1)})_{jj} = \begin{cases} 1, & j \in \left\{1 + \sum_{n=m}^d n \text{ for } m \in \{2, 3, \ldots, d\}\right\} \text{ and } \ell_1 = \ell_2 = 0 \\ \frac{1}{2}, & \text{otherwise} \end{cases}
$$

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

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

respectively, where $\lceil x \rceil$ is the smallest integer greater or equal than $x$. We observe that under the assumption of Gaussianity of $\{X_t\}$ all fourth order cumulants are zero, therefore $V_{\ell_1, \ell_2}^{(2)}$ vanishes and we obtain the diagonal matrix $(W_n)_{\ell_1, \ell_2} = W_{\ell_1, \ell_2}^{(1)} \delta_{\ell_1, \ell_2}$, which leads also to diagonal matrices $W_n$ and $W_{n,m}$ in (2.14).

### 2.5 Behaviour of $\hat{C}_T(r, \ell)$ for locally stationary time series

We now consider the behaviour of the DFT covariance $\hat{C}_T(r, \ell)$, under the assumption that the time series 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 rescaling device has been used for example in nonparametric time series by Robinson (1989) and by Dahlhaus (1997) in his definition of local stationarity. We use the rescaling device to define a locally stationary process as a time series which can ‘locally’ be approximated by a stationary time series (a similar definition was used, for example, in Dahlhaus and Subba Rao (2006), Dahlhaus and Polonik (2006), Subba Rao (2006), Vogt (2012), Dahlhaus (2012)).

#### 2.5.1 Assumptions

Typically, the vector time series $\{X_{u,T}\}$ is said to be locally stationary if there exists a (infinite dimensional) time series $\{X_t(u); t \in \mathbb{Z}, u \in [0, 1]\}$, which for fixed $u$, is strictly stationary, and can locally approximate $\{X_{u,T}\}$. In other words, if one can say that $|X_{u,T} - X_t(u)|_1 = O_p(T^{-1})$ and $|X_t(u) - X_t(v)|_1 = O_p(||u - v||)$, then $\{X_{u,T}\}$ can be called locally stationary. In order to prove the results in this paper for the case of local stationarity we require the following results. The assumptions appear quite unwielding, however, it is worth noting that the results in this paper have been derived without making any model assumptions on the process and allow for processes as diverse as linear time series, GARCH models and Hidden Markov models etc. These assumptions can be skipped on first reading.

**Assumption 2.3 (Locally stationary vector processes)** Let us suppose that the locally stationary process $\{X_{t,T}, t \in \mathbb{Z}\}$ is a $d$-variate zero mean time series and satisfies the following assumptions:
(L1) \( \{ X_{t,T}, t \in \mathbb{Z} \} \) is \( \alpha \)-mixing time series (\( \inf_{t,T} \lambda_{\min}(\text{var}(X_{t,T})) > 0 \), where \( \lambda_{\min}(A) \) denotes the smallest eigenvalue of a matrix \( A \)) with the rate
\[
\sup_{k,T \in \mathbb{Z}} \sup_{A \in \sigma(X_{t+k,T},X_{t+k+1,T},\ldots)} |P(A \cap B) - P(A)P(B)| \leq Ct^{-\alpha}, \quad t > 0
\]
where \( C \) is a constant and \( \alpha > 0 \).

(L2) There exists an associated time series \( \{ X_t(u) = (X_{t,1(u)}, \ldots, X_{t,d(u)}); t \in \mathbb{Z}, u \in [0,1] \} \) and constants \( \{ \kappa_2(r) \} \) and \( \{ \kappa_4(r_1, r_2, r_3) \} \), where for fixed \( u \), \( \{ X_t(u) \} \) is strictly stationary and satisfies
\[
(a) \ |\text{cov}(X_{t_1,T},X_{t_2,T}) - \text{cov}(X_{t_1}(t_1),X_{t_2}(t_2))|_{1} \leq \frac{1}{2} \kappa_2(t_1 - t_2)
\]
\[
(b) \ |\text{cov}(X_{t_1}(u_1),X_{t_2}(u_2)) - \text{cov}(X_{t_1}(v_1),X_{t_2}(v_2))|_{1} \leq (|u_1 - u_2| + |v_1 - v_2|) \kappa_2(t_1 - t_2)
\]
\[
(c) \ |\text{cum}(X_{t_1,T},X_{t_2,T},X_{t_3,T},X_{t_4,T}) - \text{cum}(X_{t_1}(t_1),X_{t_2}(t_2),X_{t_3}(t_3),X_{t_4}(t_4))|_{1} \leq \frac{1}{2} \kappa_4(t_2 - t_1, t_3 - t_1, t_4 - t_1)
\]

\[
(d) \ |\text{cum}(X_{t_1}(u_1),X_{t_2}(u_2),X_{t_3}(u_3),X_{t_4}(u_4)) - \text{cum}(X_{t_1}(v_1),X_{t_2}(v_2),X_{t_3}(v_3),X_{t_4}(v_4))|_{1} \leq \left( \sum_{j=1}^{4} |u_j - v_j| \right) \kappa_4(t_2 - t_1, t_3 - t_1, t_4 - t_1),
\]

where \( \sum_j |t| \cdot \kappa_2(t) \) < \( \infty \) and \( \sum_{t_1,t_2,t_3} (1 + |t_j|) \kappa_4(t_1, t_2, t_3) \) < \( \infty \).

(L3) \( \{ X_t(u), t \in \mathbb{Z} \} \) is \( \alpha \)-mixing time series with the rate
\[
\sup_{u \in \mathbb{Z}} \sup_{A \in \sigma(X_{t,u},X_{t+1,u},\ldots)} |P(A \cap B) - P(A)P(B)| \leq Ct^{-\alpha}, \quad t > 0
\]
where the mixing coefficients are defined in (L1).

(L4) For some \( s > \frac{4\alpha}{\alpha_0} > 0 \), we have \( \sup_{t,T} \| X_{t,T} \|_s \) < \( \infty \) and \( \sup_u \| X_t(u) \|_s \) < \( \infty \).

(L5) Define the localised spectral density and tri-spectra \( f(u; \omega) = \sum_{r=-\infty}^{\infty} \text{cov}(X_{t_0}(u),X_{t_r}(u)) \exp(-ir\omega) \) and \( f_4(u; \omega_1, \omega_2, \omega_3) = \sum_{r_1,r_2,r_3=-\infty}^{\infty} \text{cum}(X_{t_0}(u),X_{t_{r_1}}(u),X_{t_{r_2}}(u),X_{t_{r_3}}(u)) \exp(-ir_1\omega_1 - ir_2\omega_2 - ir_3\omega_3) \). The second partial derivative of \( f(u; \omega_1) \) and \( f_4(u; \omega_1, \omega_2, \omega_3) \) with respect to \( u \) is bounded and the first partial derivatives of these functions with respect to \( u, \omega_1, \omega_2, \omega_3 \) are bounded.

(L6) The \( (d \times d) \) integrated spectral density matrix \( f(\omega) = \int_0^1 f(u, \omega)du \) is non-singular on \( [0,2\pi] \).

(L7) There exists an \( s > \frac{8\alpha}{\alpha_0} > 0 \) such that \( \sup_{t,T} \| X_{t,T} \|_s \) < \( \infty \) and \( \sup_u \| X_t(u) \|_s \) < \( \infty \).

(L8) The variance matrix \( W_n \) defined in Theorem 2.3, where the tri-spectra is now the integrated tri-spectra \( \int_0^1 f_4(u; \omega_1, \omega_2, \omega_3)du \) is positive definite.

As in the stationary case, it can be shown that several nonlinear time series satisfy Assumption 2.3 (L1) (cf. Fryzlewicz and Subba Rao (2011) and Vogt (2012) who derive sufficient conditions for \( \alpha \)-mixing of a general class of nonstationary time series. Assumption 2.3(L2) is used to show that the covariance and the fourth order structure changes slowly over time (these assumptions are used in order to derive the limit of the DFT covariance under local stationarity). Assumption 2.3 (L1, L3, L4) are used to obtain that the fourth order cumulants of \( \{ X_t(u) \} \) are absolutely summable (Assumption 2.3 (L1, L4, L5) are used to show asymptotic normality). The stronger Assumption (L7) is required to replace \( \tilde{L}(\omega) \) with its deterministic limit (see below for the limit).
2.5.2 Sampling properties of $\hat{C}_T(r, \ell)$ under local stationarity

As in the stationary case, it is difficult to directly analyse $\hat{C}_T(r, \ell)$, therefore we show that it can be replaced by $\hat{C}_T(r, \ell)$ (defined in (2.5), where in the locally stationary case $L(\omega)$ are lower-triangular Cholesky matrices which satisfy $L(\omega)^T L(\omega) = f(\omega)^{-1}$ and $f(\omega) = \int_0^1 f(u; \omega) du$.

**Theorem 2.4 (Asymptotic equivalence of $\hat{C}_T(r, \ell)$ and $\hat{C}_T(r, \ell)$)**

Suppose Assumption 2.3 is satisfied and let $\hat{C}_T(r, \ell)$ and $\hat{C}_T(r, \ell)$ be defined as in (2.4) and (2.5), respectively. Then we have

$$\sqrt{T}|\hat{C}_T(r, \ell) - \hat{C}_T(r, \ell) - B_T(r, \ell) - S_T(r, \ell)|_1 = O_P \left( \frac{1}{b \sqrt{T}} + b^2 \sqrt{T} + b \right),$$

and

$$\hat{C}_T(r, \ell) = \mathbb{E}(\hat{C}_T(r, \ell)) + o_P(1)$$

where $B_T(r, \ell) = O(b)$ is a deterministic bias (which is defined in Appendix A.1, equation (A.5)) and $S_T(r, \ell)$ is the stochastic term (defined in (A.22)).

It is worth noting that there are some subtle differences between Theorems 2.4 and 2.1, in particular the inclusion of the additional terms $B_T(r, \ell)$ and $S_T(r, \ell)$.

We derive the limit of $C_T(r, \ell)$ under local stationarity.

**Lemma 2.1**

Suppose Assumption 2.3 is satisfied. Then for fixed $r$ we have

$$\mathbb{E}(\hat{C}_T(r, \ell)) \to A(r, \ell), \quad \hat{C}_T(r, \ell) \xrightarrow{P} A(r, \ell) \text{ as } T \to \infty, \text{ and } |A(r, \ell)|_1 \leq K|r|^{-1}|\ell|^{-2} \text{ (for some finite constant } K)$$

where $L(\omega)$ satisfies the Fourier representation

$$L(\omega) = \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 L(\omega) f(u; \omega) L(\omega)^T \exp(i2\pi ru) \exp(i\ell \omega) dud\omega.$$

We now discuss the implications of the above lemma. Since $A(r, \ell)$ are Fourier coefficients, then $L(\omega)f(u; \omega)L(\omega)^T$ satisfies the Fourier representation

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

Note that by construction $A(r, \ell) = \overline{A(-r, -\ell)}$ and since for fixed $u$, $f(u, \omega)$ is a Hermitian and periodic matrix function ($f(u, -\omega) = f(u, \omega)'$), then we have $\overline{A(r, \ell)} = A(-r, \ell)'$. Using the above representation, the local spectral density can be written in terms of these coefficients

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

An important observation about the above representations is that the coefficients $\{A(r, \ell)\}$ characterise the nonstationarity. This is because only for second order stationary time series will $A(r, \ell) = 0$ for all $r > 0$ and $\ell \geq 0$, or equivalently for any non-singular matrices $\{S_{r, \ell}\}$ and all $n, m \in \mathbb{Z}^+$

$$\sum_{r=1}^n \sum_{\ell=0}^{n-1} \|S_{r, \ell} \text{vech}(ReA(r, \ell))\|_2^2 + \|S_{r, \ell} \text{vech}(ImA(r, \ell))\|_2^2 = 0. \quad (2.17)$$
The proposed test for stationarity is based on this observation. Since by Lemma 2.1 the DFT covariances are estimators of \( \{ A(r, \ell) \} \), we use them to construct our test statistic. To do this we require the distribution of \( \hat{C}_T(r, \ell) \), which is derived in the following theorem.

**Theorem 2.5 (Limiting distributions of vech(\( \hat{K}_n(r) \)))**

Let us assume that Assumption 2.3 holds and let \( \hat{K}_n(r) \) be defined as in (2.12). Then for fixed \( m, n \in \mathbb{N} \) we have

\[
\sqrt{T} \begin{pmatrix}
\Re \hat{K}_n(1) - \Re A_n(1) - \Re B_n(1) \\
\Im \hat{K}_n(1) - \Im A_n(1) - \Im B_n(1) \\
\vdots \\
\Re \hat{K}_n(m) - \Re A_n(m) - \Re B_n(m) \\
\Im \hat{K}_n(m) - \Im A_n(m) - \Im B_n(m)
\end{pmatrix} \overset{D}{\to} N(0_{mn(d+1)}, \hat{W}_{m,n}),
\]

where \( \hat{W}_{m,n} \) is an \((mn(d+1) \times mn(d+1))\) variance matrix (which is not necessarily block diagonal), \( A_n(r) = (\text{vech}(A(r,0))', \ldots, \text{vech}(A(r,n-1))')' \) are the vectorised Fourier coefficients and \( B_n(r) = (\text{vech}(B(r,0))', \ldots, \text{vech}(B(r,n-1))')' = O(b) \).

### 3 The test statistic

#### 3.1 Motivation

We have shown that under the null of second order stationarity, the distribution of the DFT covariance matrices is asymptotically Gaussian with zero mean (see Theorem 2.3). In contrast, under the alternative of local stationarity the asymptotic distribution of the DFT covariance is not centered around zero (see Theorem 2.5).

Therefore, motivated by equation (2.17), to test for stationarity, we will test whether the limit of the squares of the DFT covariances are zero. Let \( \hat{K}_n(r) \) and \( W_n \) be defined as in (2.12) and (2.14), respectively. We define the test statistic \( T_{m,n,d} \)

\[
T_{m,n,d} = T \sum_{r=1}^{m} \left( |S_n \Re \hat{K}_n(r)|^2 + |S_n \Im \hat{K}_n(r)|^2 \right),
\]

where under the assumption that \( W_n \) is positive definite, \( S_n := U_n^{-1} \), and \( U_n \) is the lower-triangular matrix with \( W_n = U_n U'_n \).

If we can assume that \( \{ X_T \} \) is Gaussian, then \( T_{m,n,d} \) has the simple form

\[
T_{m,n,d,G} = T \sum_{r=1}^{m} \left( |\text{vech}(\Re \hat{C}_T(r,0))|^2 + |\text{vech}(\Im \hat{C}_T(r,0))|^2 \right)
+ 2T \sum_{r=1}^{m} \sum_{\ell=1}^{n-1} \left( |\text{vech}(\Re \hat{C}_T(r,\ell))|^2 + |\text{vech}(\Im \hat{C}_T(r,\ell))|^2 \right).
\]

#### 3.2 Distribution of the test statistic under the null

The following Theorem 3.1 gives the limiting distribution of \( T_{m,n,d} \) under the null of second order stationarity, and follows immediately from Theorem 2.3.
Theorem 3.1 (Limiting distribution of $T_{m,n,d}$ under the null)

Let us suppose that Assumptions 2.1 and 2.2 are satisfied. Then we have

$$T_{m,n,d} \overset{D}{\rightarrow} \chi^2_{mnd(d+1)},$$

(3.3)

where $\chi^2_{mnd(d+1)}$ is a $\chi^2$-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 $T_{m,n,d} > \chi^2_{mnd(d+1)}(1-\alpha)$, where $\chi^2_{mnd(d+1)}(1-\alpha)$ is the $(1-\alpha)$-quantile of the $\chi^2$-distribution with $mnd(d+1)$ degrees of freedom.

As we can see from the definition of the test statistic $T_{m,n,d}$ in (3.1), if $\{X_t\}$ is a non-Gaussian vector time series, the test statistic $S_n$ involves some parameters which are unknown. In particular, $S_n = U_n^{-1}$, where

$$W_n = U_n U_n'$$

and

$$(W_n)_{\ell_1,\ell_2} = (W_n)_{(1)}_{\ell_1,\ell_2} + (W_n)_{(2)}_{\ell_1,\ell_2} \quad \text{for } 0 \leq \ell_1, \ell_2 \leq n - 1.$$

A close inspection shows that (see Example 2.1) the matrix $(W_n)_{(1)}_{\ell_1,\ell_2}$ is parameter free, however $(W_n)_{(2)}_{\ell_1,\ell_2}$ is composed of $\{\kappa(\ell_1,\ell_2)(j_1,j_2,j_3,j_4)\}$, which one can see from (2.6) is an integral involving the spectral density, $f(\cdot)$ and the tri-spectra, $f_3 = \{f_{j_1,j_2,j_3,j_4}; 1 \leq j_1,j_2,j_3,j_4 \leq d\}$, which we need to estimate. In Section 4 we propose an estimator of $S_n$ based on the stationary bootstrap.

Remark 3.1 ($\kappa(\ell_1,\ell_2)(j_1,j_2,j_3,j_4)$ under linearity of $\{X_t\}$)

Suppose the additional assumption of linearity of the process $\{X_t\}$ is satisfied, that is, $\{X_t\}$ satisfies a representation

$$X_t = \sum_{\nu = -\infty}^{\infty} \Gamma_\nu \xi_{t-\nu}, \quad t \in \mathbb{Z},$$

where $\sum_{\nu = -\infty}^{\infty} |\Gamma_\nu| < \infty$, $\Gamma_0 = I_d$ and $\{\xi_t, t \in \mathbb{Z}\}$ are zero mean, i.i.d. random vectors with $\mathbb{E} (\xi_t \xi'_t) = \Sigma_e$ positive definite and whose fourth moments exist. Then the quantity $\kappa(\ell_1,\ell_2)(j_1,j_2,j_3,j_4)$ introduced in (2.6) becomes

$$\kappa(\ell_1,\ell_2)(j_1,j_2,j_3,j_4) = \sum_{s_{1,2,3,4} = 1}^{d} \kappa_{4,s_{1,2,3,4}} \left\{ \frac{1}{2\pi} \int_0^{2\pi} (L(\lambda_1)\Gamma(\lambda_1))_{j_2,j_4} (L(\lambda_1)\Gamma(\lambda_1))_{j_2,j_4} \exp(i\ell_1 \lambda_1) d\lambda_1 \right\} \times \left\{ \frac{1}{2\pi} \int_0^{2\pi} (L(\lambda_2)\Gamma(\lambda_2))_{j_3,j_4} (L(\lambda_2)\Gamma(\lambda_2))_{j_3,j_4} \exp(-i\ell_2 \lambda_2) d\lambda_2 \right\},$$

where $\Gamma(\omega) = \frac{1}{\sqrt{2\pi}} \sum_{\nu = -\infty}^{\infty} \Gamma_\nu e^{-i\nu \omega}$ is the transfer function of $\{X_t\}$ and $\kappa_{4,s_{1,2,3,4}} = \text{cum}(\xi_0_{s_1}, \xi_0_{s_2}, \xi_0_{s_3}, \xi_0_{s_4})$. 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 $\Gamma_\nu = 0$ for $\nu \neq 0$, we have $X_t = \xi_t$ and $\kappa(\ell_1,\ell_2)(j_1,j_2,j_3,j_4)$ simplifies to

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

where $\Sigma_e^{-1/2} \xi_t = (\tilde{\xi}_t, \ldots, \tilde{\xi}_{t,d})'$ and $\kappa_{4,j_1,j_2,j_3,j_4} = \text{cum}(\tilde{\xi}_0_{s_1}, \tilde{\xi}_0_{s_2}, \tilde{\xi}_0_{s_3}, \tilde{\xi}_0_{s_4})$.

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

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

where $\kappa_{4,j} = \text{cum}(\xi_0_j) / \sigma_j^4$ and $\Sigma_e = \text{diag}(\sigma_1^2, \ldots, \sigma_d^2)$. 

13
3.3 Distribution of the test statistic $T_{m,n,d}$ under the alternative of local stationarity

We now derive the limiting distribution of the test statistic under the alternative of local stationarity.

**Theorem 3.2 (Limiting distributions of $T_{m,n,d}$ under the alternative)**

Let us assume that Assumption 2.3 holds. The test statistic $T_{m,n,d}$ converges to a generalised non-central $\chi^2$-distribution, with noncentrality parameter determined by $\{\sqrt{T}A(r, \ell)\}$ and generalised parameter determined by $\tilde{W}_{m,n}$.

It immediately follows from the above that since under the alternative of local stationarity the distribution of $T_{m,n,d}$ converges to a non-central chi-squared distribution, with non-centrality parameter $\{\sqrt{T}A(r, \ell)\}$, then the power of the test converges to 100% as $T \to \infty$. Furthermore, because $|A(r, \ell)| \leq K|\ell|^{-2}|r|^{-1}$ (by Lemma 2.1), most of the nonstationarity in the time series is encoded in the first few parameters, therefore to detect the alternative we only need to use small $n$ and $m$ to construct the test statistic. Large $n$ and $m$, may result in a loss of power. In the simulation studies below we demonstrate that the test has good power (even for small sample sizes) for small values of $n$ and $m$. It is likely that penalisation methods, such as the AIC penalisation used in Lei et al. (2012) can be used to select the order $m$ and $n$.

4 A bootstrap estimator of the variance $S_n$

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 $W_n$ and thus of $S_n$, using a bootstrap procedure. As the test does not impose any structure on the observations $\{X_t\}$, we use a block-type bootstrap. 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), which is designed such that the bootstrap distribution is stationary. As we are testing for stationarity, we use the stationary bootstrap to estimate $S_n$. It is interesting to note that Parker, Paparoditis, and Politis (2005) proposed a unit root test based on the stationary bootstrap.

**The bootstrap testing scheme**

Step 1. Given the $d$-variate observations $X_1, \ldots, X_n$, evaluate $\text{vech}(\hat{\mathbf{H}}\hat{\mathbf{C}}_T(r, \ell))$ and $\text{vech}(\hat{\mathbf{S}}\hat{\mathbf{C}}_T(r, \ell))$ for $r = 1, \ldots, m$ and $\ell = 0, \ldots, n-1$.

Step 2. Define the blocks

$$ B_{i,L} = \{Y_{iL}, \ldots, Y_{iL+L-1}\}, $$

where $Y_j = X_{j T^{-1}}$ (hence there is wrapping on a torus if $j > T$) and $\bar{X} = \frac{1}{T} \sum_{t=1}^{T} X_t$. 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 $1 \leq s \leq T$ and $P(L_i = s) = p(1-p)^{s-1}$ for $s \geq 1$. 

14
Step 3. We draw blocks \( \{ B_{i_1, L_i} \}_i \) 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 \( X_1^*, \ldots, X_T^* \).

Step 4. Define the bootstrap spectral density estimator

\[
\hat{f}_T^*(\omega) = \frac{1}{T} \sum_{j=-\lfloor \frac{T-1}{2} \rfloor}^{\lfloor \frac{T-1}{2} \rfloor} K_b(\omega - \omega_j) \hat{I}_T^*(\omega_j) \hat{I}_T^*(\omega_j)', \tag{4.1}
\]

its lower-triangular Cholesky matrix \( \hat{B}_T^*(\omega) \), its inverse \( \hat{L}_T^*(\omega) = (\hat{B}_T^*(\omega))^{-1} \) and the bootstrap DFT covariances

\[
\hat{C}_T^*(r, \ell) = \frac{1}{T} \sum_{k=1}^T \hat{L}_T^*(\omega_k) \hat{I}_T^*(\omega_k) \hat{I}_T^*(\omega_{k+r}) \hat{I}_T^*(\omega_{k+r})' \exp(i\ell\omega_k), \tag{4.2}
\]

where \( \hat{I}_T^*(\omega_k) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^T X_t^* e^{-it\omega_k} \) is the bootstrap DFT.

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

\[
\{ \hat{W}_{\Re}^*(r) \}_{\ell_1+1, \ell_2+1} = T \left( \frac{1}{N} \sum_{j=1}^N \text{vech}(\Re \hat{C}_T^*(r, \ell_1))^{(j)} \text{vech}(\Re \hat{C}_T^*(r, \ell_2))^{(j)}' \right) - \left( \frac{1}{N} \sum_{j=1}^N \text{vech}(\Re \hat{C}_T^*(r, \ell_1))^{(j)} \right) \left( \frac{1}{N} \sum_{j=1}^N \text{vech}(\Re \hat{C}_T^*(r, \ell_2))^{(j)} \right)' \tag{4.3}
\]

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

Step 6. Define the bootstrap variance estimator \( \{ \hat{W}^*(r) \}_{\ell_1+1, \ell_2+1} \) as

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

let \( \hat{W}^*(r) = \hat{U}^*(r) \hat{U}^*(r)' \) be its Cholesky decomposition and define \( \hat{S}^*(r) = \hat{U}^*(r)^{-1}. \)

\( \hat{W}^*(r) \) is the bootstrap estimator of the rth block of \( \hat{W}_{m,n} \) defined in (2.14).

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

\[
\mathcal{T}_{m,n,d} = \frac{T}{m} \sum_{r=1}^m \left( |\hat{S}^*(r)\text{vech}(\Re \hat{K}_n(r))|^2 + |\hat{S}^*(r)\text{vech}(\Im \hat{K}_n(r))|^2 \right) \tag{4.4}
\]

and reject \( H_0 \) if \( \mathcal{T}_{m,n,d}^* > \chi^2_{\text{md}(d+1)}(1 - \alpha) \), where \( \chi^2_{\text{md}(d+1)}(1 - \alpha) \) is the \((1 - \alpha)\)-quantile of the \( \chi^2 \)-distribution with \( \text{md}(d+1) \) degrees of freedom to obtain a test of asymptotic level \( \alpha \in (0, 1) \).
We observe that because the blocks are random and their length is determined by a geometric distribution, then their lengths vary. However, the mean length of a block is approximately $1/p$ (since only block lengths less than length $T$ are used in the scheme). As it is assumed that $p \to 0$ and $Tp \to \infty$ as $T \to \infty$, then the mean block length increases as the sample size $T$ grows. However, we will show below 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.

**Remark 4.1**

An interesting variant on the above scheme is to use the bootstrap DFT covariances $\{\hat{C}_T^*(r, \ell)\}$ to construct bootstrap rejection regions for the test statistic. However, in this paper we focus on the asymptotic $\chi^2$-approximation.

### 4.1 Consistency of the bootstrap variance estimator

A close inspection of the bootstrap variance, shows that it involves cumulants of the bootstrap DFT. Therefore, we derive an expression for the cumulants of the bootstrap DFT conditioned on the observations. We start by defining the true, sample and sample circulant covariances as

$$
\kappa_{j_1, j_2}(r) = \text{cov}(X_{0, j_1}, X_{r, j_2}), \quad \hat{\kappa}_{j_1, j_2}(r) = \frac{1}{T} \sum_{t=1}^{T-\lfloor r \rfloor} (X_{t, j_1} - \overline{X}_{j_1})(X_{t+r, j_2} - \overline{X}_{j_2}),
$$

$$
\hat{\kappa}^C_{j_1, j_2}(r) = \frac{1}{T} \sum_{t=1}^{T} Y_{t, j_1} Y_{t+r, j_2},
$$

where $Y_{t, j} = X_{(t-1)\mod T+1, j} - \overline{X}_j$. In addition, we define the higher order circulant moments and cumulants as $\hat{\mu}^C_{n, j_1, \ldots, j_n}(r_2, \ldots, r_n) = \frac{1}{T} \sum_{t=1}^{T} \prod_{i=1}^{n} Y_{t+r_i, j_i}$, where $r_1 = 0$ and $\hat{\kappa}^C_{n, j_1, \ldots, j_n}$ denotes the sample nth order cumulant corresponding to the moments $\hat{\mu}^C_{n, j_1, \ldots, j_n}(r_2, \ldots, r_n)$, ie.

$$
\hat{\kappa}^C_{j_1, \ldots, j_n}(r_2, \ldots, r_n) = \sum_{\pi} (|\pi| - 1)!(-1)^{|\pi| - 1} \prod_{B \in \pi} \hat{\mu}^C_{j_1, \ldots, j_n}(\pi_i \in B) \quad (4.5)
$$

where $\pi$ runs through all partitions of $\{0, r_2, \ldots, r_n\}$ and $B$ are all blocks of the partition $\pi$. In order to obtain an expression for the cumulant of the DFT, we require the following lemma. We note that $E^*, \text{cov}^*, \text{cum}^*$ and $P^*$ denote the expectation, covariance, cumulant and probability measure with respect to the stationary bootstrap measure.

**Lemma 4.1**

Suppose Assumption 2.1(P1) holds, the moments $r$ and mixing rate $\alpha$ satisfy $\alpha(1 - 2n/r) > (2n - 1)$. Then we have

$$
\left\| \hat{h}_{n, j_1, \ldots, j_n}(\omega_{k_1}, \ldots, \omega_{k_{n-1}}) - f_{n, j_1, \ldots, j_n}(\omega_{k_1}, \ldots, \omega_{k_{n-1}}) \right\|_2 = O\left(\frac{1}{T} + \frac{1}{T^{1/2}p(n-1) + p}\right),
$$

where

$$
\hat{h}_{n, j_1, \ldots, j_n}(\omega_{k_1}, \ldots, \omega_{k_{n-1}}) = \frac{1}{(2\pi)^{n-1}} \sum_{r_1, \ldots, r_{n-1} = -T}^{T} (1 - p)^{\text{max}(r_i,0)-\min(r_i,0)} \hat{\kappa}_{n, j_1, \ldots, j_n}(r_1, \ldots, r_{n-1}) e^{-ir_1\omega_{k_1} - \cdots - ir_{n-1}\omega_{k_{n-1}}}. \quad (4.6)
$$
Theorem 4.1
Let us suppose \{X_t\} is a vector time series, and \{X_t^*\} is the stationary vector bootstrap sample. Let \(J_{T,j}^*(\omega)\) denote the \(j\)th component of the DFT of the stationary bootstrap.

(i) Suppose that \(0 \leq t_2 \leq t_3 \ldots \leq t_n\), then
\[
\kappa_{n,j_1,\ldots,j_n}^*(t_2, \ldots, t_n) := (1 - p)^{\max(t_n,0) - \min(t_1,0)} \kappa_{n,j_1,\ldots,j_n}^*(t_2, \ldots, t_n),
\]
where \(\kappa_{n,j_1,\ldots,j_n}^*(t_2, \ldots, t_n) := \text{cum}^*(X_{0,j_1}^*, \ldots, X_{t_n,j_n}^*)\).

(ii) Suppose \(\|X_t\|_r < \infty\), where \(r\) is such that the mixing size \(\alpha\) satisfies \(\alpha(1 - \frac{n}{r}) > (n - 1)\).
Let \(\omega_k = \frac{2\pi k}{T}\), then we have
\[
\text{cum}^*(J_{T,j_1}^*(\omega_{k_1}), \ldots, J_{T,j_n}^*(\omega_{k_n})) = \left(\frac{(2\pi)^{n/2 - 1}}{T^{n/2 - 1}} \hat{h}_{n,j_1,\ldots,j_n}(\omega_{k_1}, \ldots, \omega_{k_n})\right) \frac{1}{T} \sum_{t=1}^{T} \exp(-it(\omega_{k_1} + \ldots + \omega_{k_n})) + R_{T,n}
\]
\[
= \begin{cases} 
O_p\left(\frac{1}{T^{n/2-1}} + \frac{1}{(T^{n/2-1})^2} + \frac{1}{T^{n/2}p^n}\right), & \sum_{t=1}^{n} \omega_{k_t} \in \mathbb{Z}, \\
O_p\left(\frac{1}{T^{n/2}p^n}\right), & \sum_{t=1}^{n} \omega_{k_t} \notin \mathbb{Z}
\end{cases}
\]
where if \(\sup_{t,T} \|X_t\|_q < \infty\) we have \(\|R_{T,n}\|_{q/n} = O\left(\frac{1}{T^{n/2}p^n}\right)\).

Remark 4.2
(i) We observe that Theorem 4.1 does not require that \{\{X_t\}\} is a stationary vector time series.
In the case of local stationarity, \(\hat{h}_n(\cdot)\) can be considered as estimators of the integrated higher order spectral density similar to the integrated fourth order spectral density defined in Assumption 2.3(L8).

(ii) Combining Lemma 4.1 and Theorem 4.1, we observe that \(\text{cum}^*(J_{T,j_1}^*(\omega), J_{T,j_2}^*(\omega)) \xrightarrow{P} f_{j_1,j_2}(\omega)\). Thus \(\text{cum}^*(J_{T}^*(\omega), \hat{J}_T^*(\omega))\) is an estimator of the spectral density \(f(\omega)\).

In order to obtain the limit of the bootstrap variance estimator we define \(\tilde{C}_T^*(r, \ell)\),
\[
\tilde{C}_T^*(r, \ell) = \frac{1}{T} \sum_{k=1}^{T} L(\omega_k)J_{T}^*(\omega_k)J_{T}^*(\omega_{k+r}) L(\omega_{k+r}) \exp(i\ell\omega_k).
\]
We observe that this is almost identical to the bootstrap DFT \(\hat{C}_T^*(r, \ell)\), except that the inverse Cholesky decomposition \(L^*(\cdot)\), has been replaced with it’s limit \(L(\cdot)\). In order to show that the difference between \(\tilde{C}_T^*(r, \ell)\) and \(C_T^*(r, \ell)\) is negligible we require the following assumptions.

Assumption 4.1 (Bootstrap)
(B1) Let \(\hat{f}_T(\omega) = \text{vec}(\hat{F}_T(\omega))\) (where \(\hat{F}_T(\omega)\) is defined in (4.1)) and \(1 \leq j_1, j_2 \leq d, 0 \leq i \leq 2\).
We assume
\[
\sup_{\omega, \omega_2} \left| \left( \mathbb{E}^* \right| \nabla^i [L_{j_1,j_2}^*(\hat{f}_T(\omega_1)L_{j_2,j_4}(\hat{f}_T(\omega_2))] - \nabla^i [L_{j_1,j_2}(f(\omega_1)L_{j_3,j_4}(f(\omega_2))]\right|^4 \right)^{1/4} \xrightarrow{P} 0,
\]
where \(\nabla\) denotes the partial derivatives with respect to the elements in the vector \((f(\omega_1), f(\omega_2))\) and \(L(\hat{f}(\omega)) = B_T(\omega)\) (see Appendix A.1 for details).
(B2) The mixing rate $\alpha$ and moments $s$ (where $\|X_t\|_s < \infty$) are such that $s > \frac{8\alpha}{\alpha - 7} > 0$.

(B3) $T p^4 \to \infty$ as $bT p^2 \to \infty$, $b \to 0$ and $p \to 0$ as $T \to \infty$.

We note that due to the complexity of $\hat{W}_n$, the results in this section are technically challenging to prove, therefore we require the fairly strong assumptions given above.

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

Lemma 4.2
Suppose Assumptions 2.1 and 4.1 hold. Then we have

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

and

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

By using the above result, we can show that the stationary bootstrap estimator of the variance $\hat{W}^*(r)$ asymptotically converges to the desired limit.

Theorem 4.2 (Consistency of the bootstrap variance)
Suppose Assumptions 2.1 and 4.1 hold and let $W_{m,n}$ be defined as in (2.14). Then for all $m, n \in \mathbb{N}$, we have

$$\text{var}^* \left( \begin{array}{c} \Re \hat{K}_n^*(1) \\ \Im \hat{K}_n^*(1) \\ \vdots \\ \Re \hat{K}_n^*(m) \\ \Im \hat{K}_n^*(m) \end{array} \right) = W_{m,n} + o_p(1),$$

where $\hat{K}_n^*(r)$ is the bootstrap analogue of $\hat{K}_n^*(r)$ as defined in (2.12), that is,

$$\hat{K}_n^*(r) = \left( \text{vech}(\hat{C}_T^*(r,0))^t, \text{vech}(\hat{C}_T^*(r,1))^t, \ldots, \text{vech}(\hat{C}_T^*(r,n-1))^t \right)^t.$$

Finally, by using the above we can have the following result.

Theorem 4.3
Suppose Assumptions 2.1 and 4.1 hold. Then we have

$$|T_{m,n,d}^* - T_{m,n,d}| = o_p(1).$$

The theorem above means the test statistic with the bootstrap corrected variance asymptotically has the same distribution as the test statistic with known variance.
5 Simulations

In this section we illustrate the performance of the test for stationarity of multivariate time series through some simulations. As we do not assume any structure on the time series, we will use the bootstrap variance estimator proposed in Section 4 to estimate the unknown parameters in the test statistic. More precisely, we reject the null of stationarity at the $\alpha \times 100\%$-level if $T_{m,n,d}^* > \chi^2_{mnd(d+1)}(1 - \alpha)$, where $T_{m,n,d}^*$ is defined in (4.4). We call this the SB-test. To demonstrate that the bootstrap estimation step is necessary for obtaining the correct distribution for the test statistic under the null, we compare the SB-test with the test statistic constructed as if the process were Gaussian (even when the observations are in reality non-Gaussian). More precisely, we define the non-bootstrap test statistic $T_{m,n,d|G}$ as in (3.2), and we reject the null of stationarity at the $\alpha \times 100\%$-level if $T_{m,n,d|G} > \chi^2_{mnd(d+1)}(1 - \alpha)$. Noting that this test is asymptotically a test of level $\alpha$ only in the case that the fourth order cumulants are zero (which includes the Gaussian case).

5.1 The set-up

In the simulations below we will consider several stationary and nonstationary bivariate time series models, $(X_t = (X_{1t}, X_{2t})', t = 1, \ldots, T)$. For all models, we have generated $M = 300$ replications each of length $T = 50, 100, 200$. To estimate the spectral density matrices we have used the Bartlett–Priestley kernel, see Priestley (1981, p448) and have used $B = 200$ bootstrap samples for each time series to estimate the variance $S_1$. To see whether the tuning parameters have an impact on the results we have used several different values of $b$ and $p$, where $b$ is the bandwidth used for the smoothed periodogram in (2.3) and $p$ is define in Step 2 of the bootstrap testing scheme. To construct the test statistics $T_{m,n,d}^*$ and $T_{m,n,d|G}$ we have used $m \in \{1, 3, 5\}$, $n = 1$ and $d = 2$. In each case, the SB-based test has been executed for several combinations of bandwidths $b \in \{0.1, 0.2, 0.3\}$ and approximately expected block length $1/p \in \{1, 2, 3, 4, 5, 6\}$ and the non-bootstrap test also for different bandwidths $b \in \{0.1, 0.2, 0.3\}$. It is worth noting that $p = 1$ corresponds to the i.i.d. bootstrap and since there is dependence in the data we do not expect that this bootstrap to give good results. However we have included this case in our simulations to illustrate that a bootstrap for dependent data is necessary.

5.1.1 Models under the null hypothesis

To investigate the behaviour of the tests under the null of (second order) stationarity of the process $\{X_t\}$, we consider realizations from two vector autoregressive models (VAR), two vector moving-average models (VMA) to cover linear time series and also two non-linear time series models. More precisely, we consider data from the bivariate VAR(1) model

Model I & II $X_t = \begin{pmatrix} 0.9 & 0.3 \\ 0 & -0.5 \end{pmatrix} X_{t-1} + \varepsilon_t, \quad t = 1, \ldots, T,$

where $\{\varepsilon_t\}$ is i.i.d. zero mean random vectors, with $\varepsilon_t \sim \mathcal{N}(0, I_2)$ for Model I, and for Model II the first component of $\{\varepsilon_t, t \in \mathbb{Z}\}$ consists of i.i.d. uniformly distributed random variables, $\varepsilon_{t,1} \sim \mathcal{U}(-\sqrt{3}, \sqrt{3})$ and the second component are $t$-distributed random variables with 5 degrees of freedom multiplied with $(3/5)^{1/2}$, i.e. $(5/3)^{1/2}\varepsilon_{t,2} \sim t_5$ (noting that the variance and the excess kurtosis for these innovations are $-6/5$ and 6, respectively). Furthermore, we consider data from the bivariate VMA(1) model

Model III & IV $X_t = \begin{pmatrix} 0.9 & 0.3 \\ 0 & -0.5 \end{pmatrix} \varepsilon_{t-1} + \hat{\varepsilon}_t, \quad t = 1, \ldots, T,$
where for Model III \( \{ \varepsilon_t \} \) are i.i.d. standard normal random vectors, and for Model IV \( \{ \varepsilon_t \} \) are distributed as the innovations in Model II above. The non-linear models are based on two independent but identically distributed univariate GARCH(1,1) processes \( \{ Y_{t,i}, t \in \mathbb{Z} \}, i = 1, 2 \), with

Model V & VI \[ Y_{t,i} = \sigma_{t,i} \varepsilon_{t,i}, \quad \sigma_{t,i}^2 = 0.1 + 0.3 Y_{t-1,i}^2 + 0.6 \sigma_{t-1,i}^2, \quad t = 1, \ldots, T, \]

where \( \{ \varepsilon_{t,i}, t \in \mathbb{Z} \} \) are i.i.d. standard normal random vectors. Model V corresponds to the bivariate process \( \{ X_t = (Y_{t,1}, Y_{t,2})', t \in \mathbb{Z} \} \) and Model VI to \( \{ X_t = (|Y_{t,1}|, |Y_{t,2}|)' , t \in \mathbb{Z} \} \). Note that Model VI is not centered around zero and but no centering has to be done in advance to apply the test, since for constant mean the expectation of the DFT is zero for all frequencies except \( \omega = 0, \pi \).

5.1.2 Models under the alternative

To illustrate the behaviour of the tests under the alternative of (second order) non-stationarity of the process \( \{ X_t \} \), we consider realizations from three models with non-stationary models with different nonstationary behaviour. As we focus on locally stationary alternatives, where non-stationarity is caused by smoothly changing dynamics, we consider the time-varying VAR(1) model (tvVAR(1))

Model VII \[ X_t = AX_{t-1} + 2 \sin \left( 2\pi \frac{t}{T} \right) \varepsilon_t, \quad t = 1, \ldots, T. \]

where

\[ A = \begin{pmatrix} 0.9 & 0.3 \\ 0 & -0.5 \end{pmatrix}. \]

It can be shown that the time-varying spectral density for Model VII is \( f(u, \omega) = \frac{1}{2} (1 - \cos(4\pi u)) f(\omega) \), where \( f(\omega) \) is the spectral density matrix corresponding to the stationary time series \( Y_t = AX_{t-1} + 2\varepsilon_t \). To understand how the test performs for non-stationary time series with change points, we consider the VAR(1) model with one change point

Model VIII \[ X_t = \begin{cases} \begin{pmatrix} 0.9 & 0.3 \\ 0 & -0.5 \end{pmatrix} X_{t-1} + \varepsilon_t, & t = 1, \ldots, T/2 \\ \begin{pmatrix} -0.5 & 0.3 \\ 0 & 0.9 \end{pmatrix} X_{t-1} + \varepsilon_t, & t = T/2 + 1, \ldots, T. \end{cases} \]

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

Model IX \[ X_t = X_{t-1} + \varepsilon_t, \quad t = 1, \ldots, T. \]

In all Models VII–IX above \( \{ \varepsilon_t, t \in \mathbb{Z} \} \) are an i.i.d. standard normal random variables.

Realizations of stationary linear Models I–IV are shown in Figure 1 and non-linear Models V and VI in Figure 2, while Figure 3 is devoted to non-stationary Models VII–IX that cover three different types of non-stationary behaviour.

In Tables 1–9, the performance of both tests are reported for levels \( \alpha = 1\% \) and \( \alpha = 5\% \) and sample sizes \( T = 50, 100, 200 \), where Tables 1–6 show the results for the stationary models, while Tables 8–9 show those for non-stationary models.
5.2 Discussion of the simulations

To illustrate the limit of \( \hat{\mathbf{C}}_T(r,0) \) under the null and alternative, in Figures 4 and 5 we plot averages over 300 replications of the squared entries of DFT covariances \( T|\hat{C}_{11}(r,0)|^2 \), \( T|\hat{C}_{21}(r,0)|^2 \) and \( T|\hat{C}_{22}(r,0)|^2 \) for lag \( r = 1, \ldots, T/2 \). Figure 4 is for the stationary Models I (Gaussian VAR(1)) and III (Gaussian VMA(1)), here we make plots of the empirical quantiles of \( T|\hat{C}_{11}(r,0)|^2 \), \( T|\hat{C}_{21}(r,0)|^2 \) and \( T|\hat{C}_{22}(r,0)|^2 \) over \( r \) for both of the models. In contrast, in Figure 5, we plot the average of \( \hat{\mathbf{C}}_T(r,0) \) for the non-stationary Models VII-IX. In both figures, we show the corresponding empirical 95%-quantiles and the 95%-quantile of the \( \chi^2 \) distribution with two degrees of freedom.

In Figure 4, we observe that for small samples the empirical quantiles tend to be slightly lower than the \( \chi^2 \) approximation (suggesting the test can be quite conservative), but even for these realitively small sample sizes the empirical quantile is converging towards the asymptotic quantile as the sample size increases.

In Figure 5, we give plots of the averages of the weighted DFT for the nonstationary Models VII-IX. From Lemma 2.1 we see that for Models VII-VIII these are estimators of

\[
T^{-1} \int_0^{2\pi} \int_0^1 \mathbf{L}(\omega)\mathbf{f}(u,\omega)\overline{\mathbf{L}(\omega)} \exp(i2\pi ru)\,dud\omega \bigg| T^2 \, O(Tb^2) .
\]

The above explains why the plots are very different for all three types of non-stationarity. For all the nonstationary models, as the sample size increases, the empirical quantiles of \( T|\hat{C}_{11}(r,0)|^2 \), \( T|\hat{C}_{21}(r,0)|^2 \) and \( T|\hat{C}_{22}(r,0)|^2 \) become larger and cross the \( \chi^2 \)-quantile for certain subset of lags \( r \). Interestingly, the lags which are significant vary over the different models, which adds empirical evidence to our theoretical derivations in Section 2.5 that show that the weighted DFT covariance are estimating components of the time-varying spectral density. In particular, as the sample size grows the general shape for the locally stationary models stays the same, just the magnitude grows with the sample size. For example, for model VII the time-varying spectral density is \( f(u,\omega) = \frac{1}{2}(1 - \cos(4\pi u))f(\omega) \), thus \( \mathbf{L}(\omega)\mathbf{f}(u,\omega)\overline{\mathbf{L}(\omega)} = \frac{1}{2}(1 - \cos(4\pi u)) \) and \( A(r,\ell) = 0 \) for all \( r \) and \( \ell \) except \( r = 2 \) and \( \ell = 0 \) (which can be seen in Figure 5). In contrast, for the random walk (Model IX) the correlation is increasing over all the frequencies.

The performance of the tests under the null are shown in Tables 1–6 and under the alternative in Tables 8–9. In the case that the time series is stationary and Gaussian (Models I and III), the Gaussian test statistic \( T_{m,1,d,G} \) tends to be closer to the asymptotic quantile. The benefit of using the bootstrap variance becomes apparent for the non-Gaussian, stationary models (Models II and IV), in this case the Gaussian test statistic tends to over reject whereas using the bootstrap variance gives a test statistic which is closer to the \( \alpha \)% level. Neither of the tests appear to be particular sensitive to the tuning parameter \( b \), however, for the bootstrap test we need to choose a large \( p \) (near to one), which fits with out assumption that \( Tp^4 \to \infty \). Studying the performance of the test statistic in the nonstationary case (Tables 8–9) we see that the rejection rate depends on the choice of \( m \) (since the behaviour of the DFT covariance varies over \( r \)), however our theory has shown and the plots in Figure 5 demonstrate that the main correlations are for small values of \( r \).

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.
A Proofs

A.1 Proof of Theorems 2.1 and 2.4

Throughout this appendix we will use the notation \( f(\omega) = \text{vec}(f(\omega)) \), \( \tilde{f}(\omega) = \text{vec}(\tilde{f}(\omega)) \), \( J_{k,s} = J_{T,s}(\omega_k) \), \( f_k = f(\omega_k) \), \( \tilde{f}_k = \tilde{f}(\omega_k) \), \( f_{k,r} = (f_k') \). Let us suppose that \( G \) is a positive definite matrix, \( \mathbf{L}(G) \) and define the lower-triangular matrix \( \mathbf{L}(G) \) such that \( \mathbf{L}(G)\mathbf{G}(G)' = \mathbf{I} \) (hence \( \mathbf{L}(G) \) is the inverse of the Cholesky decomposition of \( G \)). Let \( L_{j,s}(G) \) denote the \((j,s)\)th element of the Cholesky matrix \( \mathbf{L}(G) \). Let \( \nabla L_{j,s}(G) = (\partial L_{j,s}(G)/\partial G_{11}, \ldots, \partial L_{j,s}(G)/\partial G_{dd})' \) and \( \nabla^n L_{j,s}(G) \) denote the vector of all partial \( n \)th order derivatives wrt \( G \). Furthermore, to reduce notation let \( \tilde{L}_{j,s}(\omega) = L_{j,s}(\tilde{f}(\omega)) \) and \( L_{j,s}(\omega) = L_{j,s}(f(\omega)) \).

Before proving Theorems 2.1 and 2.4 we first state some preliminary results.

Lemma A.1

Let us suppose that \( G \) is a positive definite \((d \times d)\) matrix and \( \mathbf{G} = \text{vec}(G) \). Then, for all \( 1 \leq j, s \leq d \) and \( r \in \mathbb{N} \), we have \( \|\nabla^r L_{j,s}(G)\|_2 < \infty \).

PROOF. Let \( G = \mathbf{B}\mathbf{B}' \) and be the lower-triangular Cholesky decomposition of \( G = (g_{kl}) \), set \( \mathbf{C} = \mathbf{B}^{-1} \) and let \( \Psi \) and \( \Phi \) be defined by \( \mathbf{B} = \Psi(G) \) and \( \mathbf{C} = \Phi(\mathbf{B}) \). This leads to \( \mathbf{L}(G) = \Phi(\Psi(G)) \). As there is no explicit formula for \( \mathbf{B} \) and \( \mathbf{C} \), their entries can be calculated recursively by

\[
b_{kl} = \begin{cases} 
0, & k < l \\
(g_{kk} - \sum_{j=1}^{k-1} b_{kj} b_{kj})^{1/2}, & k = l \\
\frac{1}{b_{kl}} (g_{kl} - \sum_{j=1}^{l-1} b_{kj} b_{lj}), & k > l
\end{cases}
\]

and

\[
c_{kl} = \begin{cases} 
0, & k < l \\
\frac{1}{b_{kl}}, & k = l \\
-\frac{1}{b_{kk}} \sum_{j=1}^{k-1} b_{kj} c_{jl}, & k > l
\end{cases}
\]

Due to

\[L_{kl}(G) = c_{kl} = \Phi_{kl}(\mathbf{B}) = \Phi_{kl}(\Psi(G))\]

and the chain rule, it suffices to consider \( \Psi \) and \( \Phi \) separately in the following and to prove that for all \( 1 \leq j, s \leq d \), we have \( (a) \|\nabla^r \Psi_{js}(G)\|_2 < \infty \) and \( (b) \|\nabla^r \Phi_{js}(\mathbf{B})\|_2 < \infty \). Ordering the non-zero entries of \( \mathbf{B} \) row-wise, we get for the first entry \( b_{11} = \sqrt{g_{11}} =: \Psi_{11}(G) \) such that \( \|\nabla^r \Psi_{11}(G)\|_2 < \infty \) obviously holds, because \( g_{11} > 0 \). Now suppose \( b_{kl} = \Psi_{kl}(G) \) with \( \|\nabla^r \Psi_{kl}(G)\|_2 < \infty \) for the first \( p \) non-zero elements of \( \mathbf{B} \). The \((p+1)\)th non-zero element is \( b_{mn} \), say. For \( m = n \), we get

\[b_{mn} = \left( g_{mn} - \sum_{j=1}^{m-1} b_{mj} \bar{b}_{mj} \right)^{1/2} = \left( g_{mn} - \sum_{j=1}^{m-1} \Psi_{mj}(G) \bar{\Psi}_{mj}(G) \right)^{1/2} =: \Psi_{mn}(G),\]

and for \( m > n \), we have

\[b_{mn} = \frac{1}{\Psi_{nn}(G)} \left( g_{mn} - \sum_{j=1}^{n-1} \Psi_{mj}(G) \bar{\Psi}_{nj}(G) \right) =: \Psi_{mn}(G),\]
such that $\|\nabla^r \Psi_{mn}(G)\|_2 < \infty$ is fulfilled. This is due to $g_{mm} - \sum_{j=1}^{k-1} b_{mj} b_{mj} > 0$ by $G > 0$ and the induction hypothesis for the first $p$ non-zero elements of $B$. This proves part (a). To prove part (b), we get immediately $c_{kk} = \Phi_{kk}(B)$ with $\|\nabla^r \Phi_{kk}(B)\|_2 < \infty$ for all $k$ due to $b_{kk} > 0$. We order the non-zero off-diagonal elements of $C$ row-wise and for the first such entry we get $c_{21} = -b_{21} c_{11} / b_{22} =: \Phi_{21}(B)$ with $\|\nabla^r \Phi_{21}(B)\|_2 < \infty$ as $b_{22} > 0$. Now suppose it holds $c_{kl} = \Phi_{kl}(B)$ with $\|\nabla^r \Phi_{kl}(B)\|_2 < \infty$ for the first $p$ non-zero off-diagonal elements of $C$. The $(p+1)$th non-zero element equals $c_{mn}$, say, and we have

$$c_{mn} = -\frac{1}{b_{mm}} \sum_{j=l}^{m-1} b_{mj} c_{jn} = -\frac{1}{b_{mm}} \sum_{j=l}^{m-1} b_{mj} \Phi_{jn}(B) =: \Phi_{mn}(B)$$

such that $\|\nabla^r \Phi_{mn}(B)\|_2 < \infty$, because $b_{mm} > 0$ and the induction hypothesis for the first $p$ non-zero off-diagonal elements of $C$. This proves part (b), which concludes this proof. □

We use the lemma above to prove the following lemma, which will be used frequently in this section and to prove Lemma A.14 in Section A.4.

**Lemma A.2 (Uniform convergence)**

Suppose that \(\{X_i\}\) is an $\alpha$-mixing vector time series and \(\hat{f}_T\) be defined in (2.3), where Assumption 2.2 and either Assumption 2.1 or 2.3 is satisfied. Then, if $b^2 T \to \infty$ we have

(i) \(\sup_\omega |\hat{f}_T(\omega) - f(\omega)|_1 \overset{P}{\to} 0\),

(ii) if $f(\omega)$ is nonsingular, then for all $1 \leq j, s \leq d$, we have \(\sup_\omega |L_{js}(\hat{f}_T(\omega)) - L_{js}(f(\omega))| \overset{P}{\to} 0\), as $T \to \infty$.

**PROOF.** We first prove (i), by considering the pointwise convergence on the matrix \(\hat{f}_T(\omega)\), and show

$$\sup_\omega |\hat{f}_{j_1,j_2}(\omega) - E[\hat{f}_{j_1,j_2}(\omega)]| \overset{P}{\to} 0 \text{ and } \sup_\omega |E[\hat{f}_{j_1,j_2}(\omega)] - f_{j_1,j_2}(\omega)| \to 0.$$  

To prove $\sup_\omega |\hat{f}_{j_1,j_2}(\omega) - E[\hat{f}_{j_1,j_2}(\omega)]| \overset{P}{\to} 0$, we use the Chebyshev's inequality, this requires showing that

$$E \left( \sup_\omega |\hat{f}_{j_1,j_2}(\omega) - E(\hat{f}_{j_1,j_2}(\omega))|^2 \right) \to 0 \quad \text{as } T \to \infty.$$  

To bound $E \sup_\omega |\hat{f}_{j_1,j_2}(\omega) - E(\hat{f}_{j_1,j_2}(\omega))|^2$, we will use Theorem 3B, page 85, Parzen (1999). More precisely, if \(\{X(\omega); \omega \in [0, \pi]\}\) is a zero mean stochastic process, then Parzen (1999) shows that

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

(A.2)

In our case we let $X(\omega) = \hat{f}_{j_1,j_2}(\omega) - E[\hat{f}_{j_1,j_2}(\omega)]$. The derivative in our case is

$$\frac{\partial \hat{f}_{j_1,j_2}(\omega_s)}{\partial \omega_s} = \frac{1}{T} \sum_{t, \tau = 1}^{T} i(t - \tau) X_t X_\tau \lambda_b(t - \tau) \exp(i(t - \tau) \omega_s).$$
Under Assumption 2.1 or 2.3, we have \( \sup_r \sum_i |r| \cdot |\text{cov}(X_{t,j_1}, X_{t+r,j_2})| < \infty \) and for \( 1 \leq i \leq 3 \), \( \sup_{r_1,r_2,r_3} |r| \cdot \text{cum}(X_{t,j_1}, X_{t+r_1,j_2}, X_{t+r_2,j_1}, X_{t+r_3,j_2}) \) \( < \infty \), this implies \( \text{var}(\tilde{f}_{j_1,j_2}(\omega)) = O\left(\frac{1}{T}\right) \) and \( \text{var}\left(\frac{\partial \tilde{f}_{j_1,j_2}(\omega)}{\partial \omega}\right) = O\left(\frac{1}{T}\right) \). Therefore, by using (A.2) we have

\[
\mathbb{E}\left( \sup_{0 \leq \omega \leq \pi} \left| \tilde{f}_{j_1,j_2}(\omega) - \mathbb{E}(\tilde{f}_{j_1,j_2}(\omega)) \right|^2 \right) \leq \frac{1}{2} \text{var}(\tilde{f}_{j_1,j_2}(0)) + \frac{1}{2} \text{var}(\tilde{f}_{j_1,j_2}(\pi)) + \int_0^\pi \left[ \text{var}(\tilde{f}_{j_1,j_2}(\omega)) \text{var}(\frac{\partial \tilde{f}_{j_1,j_2}(\omega)}{\partial \omega}) \right]^{1/2} d\omega = O\left(\frac{1}{b^{3/2}}\right).
\]

Thus by using the above and Chebyshev’s inequality, for any \( \varepsilon > 0 \) we have

\[
P\left( \sup \omega |\tilde{f}_{j_1,j_2}(\omega) - \mathbb{E}(\tilde{f}_{j_1,j_2}(\omega))| > \varepsilon \right) \leq \frac{\mathbb{E} \sup \omega |\tilde{f}_{j_1,j_2}(\omega) - \mathbb{E}(\tilde{f}_{j_1,j_2}(\omega))|^2}{\varepsilon^2} = O\left(\frac{1}{T \varepsilon}ight) \to 0
\]
as \( Tb^{3/2} \to \infty, b \to 0 \) and \( T \to \infty \). In addition, under the stated assumptions we have \( \sup_{\omega} |\mathbb{E}(\tilde{f}_{j_1,j_2}(\omega)) - \tilde{f}_{j_1,j_2}(\omega)| \to 0 \). Altogether, this gives (i).

To prove (ii), we use the mean value theorem on \( L_{js} \) to give

\[
\sup_{\omega} |L_{js}(\tilde{f}_T(\omega)) - L_{js}(\tilde{f}(\omega))| = \nabla L_{js}(\tilde{f}_T(\omega))(\tilde{f}_T(\omega) - \tilde{f}(\omega)).
\]

We note that under the stated assumptions \( \sup_{\omega} |\nabla L_{js}(\tilde{f}(\omega))| \) is bounded (see Lemma A.1), therefore \( |\nabla L_{js}(\tilde{f}(\omega))| \) is bounded in probability. Thus (ii) follows from (i).

In the results below we will make use of the lemma above and the Taylor expansion

\[
A_{j_1,j_2,s_1,s_2}(\tilde{f}_{k,r}) - A_{j_1,j_2,s_1,s_2}(f_{k,r}) = (\tilde{f}_{k,r} - f_{k,r})'\nabla A_{j_1,j_2,s_1,s_2}(\tilde{f}_{k,r}) + (\tilde{f}_{k,r} - f_{k,r})'\nabla^2 A_{j_1,j_2,s_1,s_2}(\tilde{f}_{k,r})(\tilde{f}_{k,r} - f_{k,r}), \tag{A.3}
\]

with \( \tilde{f}_{k,r} \) lying between \( f_{k,r} \) and \( \tilde{f}_{k,r} \) and \( A \) defined in (A.1). The main contribution in this section is to obtain a bound for the difference between \( \tilde{c}_{j_1,j_2}(r, \ell) \) and \( \tilde{c}_{j_1,j_2}(r, \ell) \). To show this, we decompose the difference as

\[
\sqrt{T}(\tilde{c}_{j_1,j_2}(r, \ell) - \tilde{c}_{j_1,j_2}(r, \ell)) = \frac{1}{\sqrt{T}} \sum_{k=1}^T \sum_{s_1,s_2=1}^d J_{k,s_1,s_2}T_{k+r,s_2} \left[ A_{j_1,j_2,s_1,s_2}(\tilde{f}_{k,r}) - A_{j_1,j_2,s_1,s_2}(f_{k,r}) \right] e^{i\ell \omega_k}
\]

\[
= \frac{1}{\sqrt{T}} \sum_{k=1}^T \sum_{s_1,s_2=1}^d \left[ J_{k,s_1,s_2}T_{k+r,s_2} - \mathbb{E}(J_{k,s_1,s_2}T_{k+r,s_2}) \right] \left[ A_{j_1,j_2,s_1,s_2}(\tilde{f}_{k,r}) - A_{j_1,j_2,s_1,s_2}(f_{k,r}) \right] e^{i\ell \omega_k} +
\]

\[
\frac{1}{\sqrt{T}} \sum_{k=1}^T \sum_{s_1,s_2=1}^d \mathbb{E}(J_{k,s_1,s_2}T_{k+r,s_2}) \left[ A_{j_1,j_2,s_1,s_2}(\tilde{f}_{k,r}) - A_{j_1,j_2,s_1,s_2}(f_{k,r}) \right] e^{i\ell \omega_k} \tag{I}
\]

\[
= I + II.
\]

We now substitute the Taylor expansion, (A.3), into the above to obtain \( I = A_1 + \tilde{A}_2 \) where

\[
A_1 = \frac{1}{\sqrt{T}} \sum_{k=1}^T \sum_{s_1,s_2=1}^d \left[ J_{k,s_1,s_2}T_{k+r,s_2} - \mathbb{E}(J_{k,s_1,s_2}T_{k+r,s_2}) \right] \left( \tilde{f}_{k,r} - f_{k,r} \right)' \nabla A_{j_1,j_2,s_1,s_2}(\tilde{f}_{k,r})(\tilde{f}_{k,r} - f_{k,r}) e^{i\ell \omega_k},
\]

\[
\tilde{A}_2 = \frac{1}{\sqrt{T}} \sum_{k=1}^T \sum_{s_1,s_2=1}^d \left[ J_{k,s_1,s_2}T_{k+r,s_2} - \mathbb{E}(J_{k,s_1,s_2}T_{k+r,s_2}) \right] \left( \tilde{f}_{k,r} - f_{k,r} \right)' \nabla^2 A_{j_1,j_2,s_1,s_2}(\tilde{f}_{k,r})(\tilde{f}_{k,r} - f_{k,r}) e^{i\ell \omega_k}.
\]
and $II = B_1 + \tilde{B}_2$, where

$$B_1 = \frac{1}{\sqrt{T}} \sum_{k=1}^{T} \sum_{s_1, s_2=1}^{d} \mathbb{E}(J_{k,s_1} J_{k+r,s_2}) (\hat{f}_{k,r} - f_{k,r})^\prime \nabla A_{j_1,s_1,j_2,s_2}(f_{k,r}) e^{i\ell \omega_k}$$

$$\tilde{B}_2 = \frac{1}{\sqrt{T}} \sum_{k=1}^{T} \sum_{s_1, s_2=1}^{d} \mathbb{E}(J_{k,s_1} J_{k+r,s_2}) (\hat{f}_{k,r} - f_{k,r})^\prime \nabla^2 A_{j_1,s_1,j_2,s_2}(\tilde{f}_{k,r})(\tilde{f}_{k,r} - f_{k,r}) e^{i\ell \omega_k}$$

with $\tilde{f}_{k,r}$ lying between $\hat{f}_{k,r}$ and $f_{k,r}$. We further decompose $A_1$ and $B_1$ by adding and subtracting $\mathbb{E}(\tilde{f}_{k,r})$ to obtain

$$A_{1,1} = \frac{1}{\sqrt{T}} \sum_{k=1}^{T} \sum_{s_1, s_2=1}^{d} \left[J_{k,s_1} J_{k+r,s_2} - \mathbb{E}(J_{k,s_1} J_{k+r,s_2})\right] (\hat{f}_{k,r} - \mathbb{E}(\tilde{f}_{k,r}))^\prime \nabla A_{j_1,s_1,j_2,s_2}(f_{k,r}) e^{i\ell \omega_k}$$

$$A_{1,2} = \frac{1}{\sqrt{T}} \sum_{k=1}^{T} \sum_{s_1, s_2=1}^{d} \left[J_{k,s_1} J_{k+r,s_2} - \mathbb{E}(J_{k,s_1} J_{k+r,s_2})\right] \left(\mathbb{E}(\tilde{f}_{k,r}) - f_{k,r}\right)^\prime \nabla A_{j_1,s_1,j_2,s_2}(f_{k,r}) e^{i\ell \omega_k}$$

and

$$S_{T,j_1,j_2}(r, \ell) = \frac{1}{T} \sum_{k=1}^{T} \sum_{s_1, s_2=1}^{d} \mathbb{E}(J_{k,s_1} J_{k+r,s_2})(\hat{f}_{k,r} - \mathbb{E}(\tilde{f}_{k,r}))^\prime \nabla A_{j_1,s_1,j_2,s_2}(f_{k,r}) e^{i\ell \omega_k} \quad (A.5)$$

$$B_{T,j_1,j_2}(r, \ell) = \frac{1}{T} \sum_{k=1}^{T} \sum_{s_1, s_2=1}^{d} \mathbb{E}(J_{k,s_1} J_{k+r,s_2}) \left(\mathbb{E}(\tilde{f}_{k,r}) - f_{k,r}\right)^\prime \nabla A_{j_1,s_1,j_2,s_2}(f_{k,r}) e^{i\ell \omega_k}.$$

Therefore we have $I = A_{1,1} + A_{1,2} + \tilde{A}_2$ and $II = \sqrt{T}(S_{T,j_1,j_2}(r, \ell) + B_{T,j_1,j_2}(r, \ell)) + \tilde{B}_2$. By using Lemma A.2 we have

$$\sup_{\omega_1, \omega_2} |\nabla^2 A_{j_1,s_1,j_2,s_2}(\tilde{f}_{k,r})(\omega_1)^\prime, \tilde{f}_{k,r}(\omega_2)^\prime) - \nabla^2 A_{j_1,s_1,j_2,s_2}(f(\omega_1)^\prime, f(\omega_2)^\prime)| \xrightarrow{P} 0.$$

Using this, we take the absolute values of $\tilde{A}_2$ and $\tilde{B}_2$, and replace $\nabla^2 A_{j_1,s_1,j_2,s_2}(\tilde{f}_{k,r})$ with its deterministic limit $\nabla^2 A_{j_1,s_1,j_2,s_2}(f_{k,r})$, to give

$$\sqrt{T} (\tilde{c}_{j_1,j_2}(r, \ell) - \tilde{c}_{j_1,j_2}(r, \ell)) = A_{1,1} + A_{1,2} + O_p(A_2) + \sqrt{T}(S_{T,j_1,j_2}(r, \ell) + B_{T,j_1,j_2}(r, \ell)) + O_p(B_2), \quad (A.6)$$

where we abuse notation and let $O_p(A_2)$ and $O_p(B_2)$ denote random variables which have the same orders as the positive random variables $A_2$ and $B_2$, where

$$A_2 = \frac{1}{\sqrt{T}} \sum_{k=1}^{T} \sum_{s_1, s_2=1}^{d} |J_{k,s_1} J_{k+r,s_2} - \mathbb{E}(J_{k,s_1} J_{k+r,s_2})| \left| (\hat{f}_{k,r} - f_{k,r})^\prime \nabla^2 A_{j_1,s_1,j_2,s_2}(f_{k,r})(\hat{f}_{k,r} - f_{k,r}) \right|,$$

$$B_2 = \frac{1}{\sqrt{T}} \sum_{k=1}^{T} \sum_{s_1, s_2=1}^{d} \mathbb{E}(J_{k,s_1} J_{k+r,s_2}) \left| (\hat{f}_{k,r} - f_{k,r})^\prime \nabla^2 A_{j_1,s_1,j_2,s_2}(f_{k,r})(\hat{f}_{k,r} - f_{k,r}) \right|. \quad (A.7)$$

Below we will bound $A_{1,1}, \ldots, B_2$, and show that for both the stationary and nonstationary cases $A_{1,1}, A_{1,2}, A_2$ and $B_2$ are asymptotically negligible. In contrast, in the case of second order stationarity, $\sqrt{T}S_{T,j_1,j_2}(r, \ell) = O\left(\frac{1}{\sqrt{T}}\right)$ and $\sqrt{T}B_{T,j_1,j_2}(r, \ell) = O\left(\frac{1}{\sqrt{T}}\right)$ whereas in the nonstationary
case these two terms are no longer asymptotically negligible, and play a role in the distribution of $\tilde{c}_{j_1,j_2}(r,\ell)$.

To prove the above, we require the following results. The following lemma derives conditions for summability of the cumulants based on mixing and moment assumptions. These results are used in this section and in Section A.4.

**Lemma A.3**

Let us suppose that $\{X_t\}$ is an $\alpha$-mixing time series with rate $\{K|t|^{-\alpha}\}$. If $t_1 \leq t_2 \leq \ldots \leq t_k$, then we have $|\text{cum}(X_{t_1,j_1}, \ldots, X_{t_k,j_k})| \leq C_k \sup_{t,T} \|X_{t,T}\|_{r}^{\frac{1}{r}} \prod_{i=2}^{k} |t_i - t_{i-1}|^{-\alpha \left( \frac{1-k/r}{1-\alpha} \right)}$,

$$\sup_{t_1} \sum_{t_2,\ldots,t_k=1}^{\infty} |\text{cum}(X_{t_1,j_1}, \ldots, X_{t_k,j_k})| \leq C_k \sup_{t,j} \|X_{t,j}\|_{r}^{\frac{1}{r}} \left( \sum_{t} |t|^{-\alpha \left( \frac{1-k/r}{1-\alpha} \right)} \right)^{k-1} < \infty, \quad (A.8)$$

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

$$\sup_{t_1} \sum_{t_2,\ldots,t_k=1}^{\infty} (1 + |t_j|)|\text{cum}(X_{t_1,j_1}, \ldots, X_{t_k,j_k})| \leq C_k \sup_{t,j} \|X_{t,j}\|_{r}^{\frac{1}{r}} \left( \sum_{t} |t|^{-\alpha \left( \frac{1-k/r}{1-\alpha} \right)} \right)^{k-1} < \infty, \quad (A.9)$$

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). \hfill \Box

**Corollary A.1**

Suppose Assumption 2.1(P1, P2) or 2.3(L1, L3, L4) holds, then we have

$$\sup_{t} \sum_{r} |r| \cdot |\text{cov}(X_{t,j_1}, X_{t+r,j_2})| < \infty, \quad \sup_{u} \sum_{r} |r| \cdot |\text{cov}(X_{t,j_1}(u), X_{t+r,j_2}(u))| < \infty$$

and for $1 \leq i \leq 3$ and $1 \leq j_1, j_2, j_3, j_4 \leq d$

$$\sup_{t_1} \sum_{t_2, t_3, t_4} (1 + |t_i|) \cdot |\text{cum}(X_{t_1,j_1}, X_{t_2,j_2}, X_{t_3,j_3}, X_{t_4,j_4})| < \infty, \quad \sup_{u} \sum_{t_1, t_2, t_3} (1 + |t_i|) \cdot |\text{cum}(X_{0,j_1}(u), X_{t_1,j_2}(u), X_{t_2,j_3}(u), X_{t_3,j_4}(u))| < \infty.$$  

Furthermore, if Assumption 2.1(P1, P4) or 2.3(L1, L2, L7) holds, then for $1 \leq n \leq 8$ we have

$$\sup_{t_1} \sum_{t_2,\ldots,t_n} |\text{cum}(X_{t_1,j_1}, X_{t_2,j_2}, \ldots, X_{t_n,j_n})| < \infty, \quad \sup_{u} \sum_{t_2,\ldots,t_n} |\text{cum}(X_{0,j_1}(u), X_{t_2,j_2}(u), \ldots, X_{t_n,j_n}(u))| < \infty.$$  

**PROOF.** The proof immediately follows from Lemma A.3, thus we omit the details. \hfill \Box

The following lemmas are used to bound $A_{1,1}, \ldots, B_2$.

**Lemma A.4**

Suppose that for $1 \leq n \leq 8$, the DFTs satisfy $|\text{cum}(J_{k_1,j_1}, \ldots, J_{k_n,j_n})| \leq K|\sum_{l} k_l|^{-2}$. Then for all $1 \leq j_1, j_2, j_3, j_4 \leq d$ we have

$$\frac{1}{\sqrt{T}} \sum_{k} h_k(J_{k,j_1}, J_{k+r,j_2} - \mathbb{E}(J_{k,j_1}, J_{k+r,j_2}))(\hat{f}_{j_3,j_4}(\omega_k) - \mathbb{E}(\hat{f}_{j_3,j_4}(\omega_k))) = O_p\left( \frac{1}{\sqrt{T}} + \frac{1}{\sqrt{dT}} \right). \quad (A.10)$$

26
and
\[
\frac{1}{\sqrt{T}} \sum_k h_k(J_{k,j_1} J_{k+r,j_2} - \mathbb{E}(J_{k,j_1} J_{k+r,j_2})) (\hat{f}_{j_3,j_4}(\omega_{k+r}) - \mathbb{E}(\hat{f}_{j_3,j_4}(\omega_{k+r})) = O_p \left( \frac{1}{\sqrt{T}} + \frac{1}{bT} \right) \tag{A.11}
\]
where \( \{h_k\} \) is a bounded sequence.

**PROOF.** Let
\[
D_T = \frac{1}{\sqrt{T}} \sum_k h_k(J_{k,j_1} J_{k+r,j_2} - \mathbb{E}(J_{k,j_1} J_{k+r,j_2})) (\hat{f}_{j_3,j_4}(\omega_k) - \mathbb{E}(\hat{f}_{j_3,j_4}(\omega_k))
\]
\[
= \frac{1}{\sqrt{T}} \sum_{k,l} h_k K_b(\omega_k - \omega_l)(J_{k,j_1} J_{k+r,j_2} - \mathbb{E}(J_{k,j_1} J_{k+r,j_2})) (J_{l,j_3} J_{l,j_4} - \mathbb{E}(J_{l,j_3} J_{l,j_4})).
\]

Under the stated assumptions it is straightforward to show that \( \mathbb{E}(D_T) = O(\frac{1}{\sqrt{T}}) \). We now consider var\((D_T)\). By expanding var\((D_T)\) in terms of cumulants of various orders (up to eighth order) using \( \frac{1}{T} \sum_k K_b(\omega_k - \omega_l) = O(1) \) and under the stated assumptions it is straightforward to show that var\((D_T)\) is \( O(\frac{1}{T}) \). Therefore, altogether we have \( D_T = O_p(\frac{1}{(bT)^{1/2}} + \frac{1}{T^{1/2}}) \), which gives (A.10). A similar proof is used to show (A.11). \( \square \)

**Lemma A.5**

Suppose that the vector time series \( \{X_t\} \) is \( \alpha \)-mixing, where the \( \alpha \)-mixing rate \( K|t|^{-\alpha} \) is such that there exists an \( s \) such that for \( s > 4\alpha/(\alpha - 3) \) and \( \sup_t \|X_t\|_s < \infty \). Then for all \( 1 \leq j_1,j_2 \leq d \) we have
\[
\|J_{k,j_1} J_{k+r,j_2} - \mathbb{E}(J_{k,j_1} J_{k+r,j_2})\|_2 = O(1).
\]

**PROOF.** We have
\[
J_{k,j_1} J_{k+r,j_2} - \mathbb{E}(J_{k,j_1} J_{k+r,j_2}) = \frac{1}{2\pi T} \sum_{t,\tau=1}^T \rho_{t,\tau}(X_{t,j_1} X_{t,j_2} - \mathbb{E}(X_{t,j_1} X_{t,j_2})),
\]
where \( \rho_{t,\tau} = \exp(i\omega_k(t - \tau))\exp(-i\omega_\tau) \). Now evaluating the variance we have
\[
\mathbb{E}[J_{k,j_1} J_{k+r,j_2} - \mathbb{E}(J_{k,j_1} J_{k+r,j_2})]^2 \leq \frac{1}{(2\pi)^2} (I + II + III),
\]
where
\[
I = T^{-2} \sum_{t_1,j_2=1}^T \sum_{\tau_1,\tau_2=1}^T \rho_{t_1,\tau_1} \rho_{t_2,\tau_2} \text{cov}(X_{t_1,j_1}, X_{t_2,j_1}) \text{cov}(X_{\tau_1,j_2}, X_{\tau_2,j_2})
\]
\[
II = T^{-2} \sum_{t_1,t_2=1}^T \sum_{\tau_1,\tau_2=1}^T \rho_{t_1,\tau_1} \rho_{t_2,\tau_2} \text{cov}(X_{t_1,j_1}, X_{\tau_2,j_2}) \text{cov}(X_{\tau_1,j_2}, X_{t_2,j_1})
\]
\[
III = T^{-2} \sum_{t_1,j_2=1}^T \sum_{\tau_1,\tau_2=1}^T \rho_{t_1,\tau_1} \rho_{t_2,\tau_2} \text{cum}(X_{t_1,j_1}, X_{\tau_1,j_2}, X_{t_2,j_1}, X_{\tau_2,j_2}).
\]

Since \( \sup_{t_1,j_1,j_2} |\text{cov}(X_{t_1,j_1}, X_{\tau_2,j_2})| < \infty \) and \( \sup_{t_1,j_1,j_2} \sum_{\tau_1,\tau_2} |\text{cum}(X_{t_1,j_1}, X_{\tau_1,j_2}, X_{t_2,j_1}, X_{\tau_2,j_2})| < \infty \) (see Corollary A.1) we obtain the result. \( \square \)
Lemma A.6
Suppose Assumption 2.1 or 2.3 is satisfied. Then we have \(1 \leq j_1, j_2, j_3, j_4 \leq d\) we have
\[
|\mathbb{E}(\hat{f}_{k,j_1,j_2}) - f_{k,j_1,j_2}| = O(b), \tag{A.13}
\]
\[
\|\hat{f}_{k,j_1,j_2} - \mathbb{E}(\hat{f}_{k,j_1,j_2})\|_4 = O \left( \frac{1}{\sqrt{T}} \right), \tag{A.14}
\]
and thus \(\|\hat{f}_{k,j_1,j_2} - f_{k,j_1,j_2}\|_4^2 = O\left( \frac{1}{bT} + b^2 + \frac{b^{1/2}}{T^{1/2}} \right)\).

PROOF. The proof of (A.13) is well known for spectral density estimators, hence we omit the details. To prove (A.14) we use that
\[
\|\hat{f}_{k,j_1,j_2} - \mathbb{E}(\hat{f}_{k,j_1,j_2})\|_4^4 = \text{var}(\hat{f}_{k,j_1,j_2})^2 + \text{cum}(\hat{f}_{k,j_1,j_2}, \hat{f}_{k,j_1,j_2}, \hat{f}_{k,j_1,j_2}, \hat{f}_{k,j_1,j_2}).
\]
By decomposing the above in terms of cumulants of the DFTs, using the DFT cumulant results in Lemma 3.2, Lee and Subba Rao (2011) and some long but straightforward calculations we obtain the result. \(\Box\)

Lemma A.7
Suppose that the vector time series \(\{X_t\}\), where for all \(1 \leq j_1, j_2 \leq d\) the DFTs satisfy \(\sup_{k_2,T} \sum_{k_1}^T |\text{cov}(J_{k_1,j_1}, J_{k_2,j_2})| < \infty\) and \(\sup_{k_2,T} \sum_{k_1}^T |\text{cum}(J_{k_1,j_1}, J_{k_1+1,j_2}, J_{k_2,j_2}, J_{k_2+1,j_2})| < \infty\), then we have
\[
\left\| \frac{1}{\sqrt{T}} \sum_{k=1}^T h_k (J_{k,j_1}J_{k+r,j_2} - \mathbb{E}(J_{k,j_1}J_{k+r,j_2})) \right\|_2 = O \left( \max_{1 \leq k \leq T} |h_k| \right),
\]
where \(\{h_k\}\) is a bounded sequence.

PROOF. First we observe that the above is a variance, thus it is equal to
\[
\text{var} \left( \frac{1}{\sqrt{T}} \sum_{k=1}^T h_k (J_{k,j_1}J_{k+r,j_2} - \mathbb{E}(J_{k,j_1}J_{k+r,j_2})) \right) = \frac{1}{T} \sum_{k_1,k_2=1}^T h_{k_1} h_{k_2} \text{cov}(J_{k_1,j_1}, J_{k_1+1,j_2}, J_{k_2,j_2}, J_{k_2+1,j_2})
\]
\[
= \frac{1}{T} \sum_{k_1,k_2=1}^T h_{k_1} h_{k_2} \left\{ \text{cov}(J_{k_1,j_1}, J_{k_2,j_2}) \text{cov}(J_{k_1+1,j_2}, J_{k_2+1,j_2}) + \text{cov}(J_{k_1,j_1}, J_{k_2,j_2}) \text{cov}(J_{k_1+1,j_2}, J_{k_2,j_1}) + \right. 
\]
\[ \left. \text{cum}(J_{k_1,j_1}, J_{k_1+1,j_2}, J_{k_2,j_2}, J_{k_2+1,j_2}) \right\}
\]
Therefore, under the stated conditions, the result immediately follows. \(\Box\)

Lemma A.8
(i) Suppose Assumption 2.1(P1,P2) holds. Then we have
\[
\text{cov}(J_{k,j_1}, J_{k+r,j_2}) = O \left( \frac{1}{T} \right) \quad \text{cum}(J_{k,j_1}, J_{k+r,j_2}, J_{k,j_1}, J_{k+r,j_2}) = O \left( \frac{1}{T} \right).
\]

28
(ii) Suppose Assumption 2.3(L1-L4) holds. Then we have

$$\text{cov}(J_{k_1,j_1}, J_{k_2,j_2}) = \int_0^1 f_{j_1,j_2}(u, \omega_k) \exp(-2\pi i(k_1 - k_2)u) du + O\left(\frac{1}{T}\right)$$  \hspace{1cm} (A.15)

and

$$\text{cum}(J_{k_1,j_1}, \ldots, J_{k_4,j_4}) = \frac{(2\pi)^{n/2-1}}{T^{n/2-1}} \int_0^1 f_{4,j_1,\ldots,j_4}(u; \omega_{k_2},\ldots,\omega_{k_4}) \exp(-i2\pi u(k_1 + k_2 + k_3 + k_4)) du + O\left(\frac{1}{T^{n/2}}\right)$$

where $f_{j_1,j_2}(u, \omega)$ is the joint local spectral and $f_{4,j_1,\ldots,j_4}$ the local joint tri-spectra density.

(iii) Under Assumption 2.3(L5), for $r \neq 0$ we have $|\int_0^1 f_{j_1,j_2}(u, \omega_k) \exp(-2\pi i ru) du| \leq K|r|^{-1}$ and $\int_0^1 f_{4,j_1,\ldots,j_4}(u; \omega_{k_2},\ldots,\omega_{k_4}) \exp(-2\pi i ru) du \leq K|r|^{-1}$.

PROOF. Under Assumption 2.1(P1, P2), we have $\sum_r |r| \cdot |\text{cov}(X_{0,j_1}, X_{r,j_2})| < \infty$ and $\sum_{k_1,k_2,k_3} (1+|k_1|) \cdot |\text{cum}(X_{0,j_1}, X_{k_1,j_2}, X_{k_2,j_3}, X_{k_3,j_4})| < \infty$, thus (i) follows from (Brillinger, 1981), Theorem 4.3.2, we have (i).

To prove (A.15) we expand $\text{cov}(J_{k_1,j_1}, J_{k_2,j_2})$ to give

$$\text{cov}(J_{k_1,j_1}, J_{k_2,j_2}) = \frac{1}{2\pi T} \sum_{t,\tau=1}^T \text{cov}(X_{t,T}, X_{\tau,T}) \exp(-i(t-\tau)\omega_k + \tau\omega_r).$$

Under Assumption 2.3(L2) we can replace $\text{cov}(X_{t,T}, X_{\tau,T})$ with $\text{cov}(X_{t,T}(\cdot), X_{\tau,T}(\cdot))$ to give

$$\text{cov}(J_{k_1,j_1}, J_{k_2,j_2}) = \frac{1}{2\pi T} \sum_{t,\tau=1}^T \text{cov}(X_{t,j_1}(\frac{T}{T}), X_{\tau,j_2}(\frac{T}{T})) \exp(-i(t-\tau)\omega_{k_1}) \exp(-i\tau\omega_{k_1-k_2}) + \text{Op} \left(\frac{1}{T} \sum_r |r| \cdot \kappa_2(r)\right)$$

$$= \frac{1}{T} \sum_{\tau=1}^T f_{j_1,j_2}(\frac{T}{T},\omega) \exp(-i(k_1 - k_2)\omega_{\tau}) + \text{Op} \left(\frac{1}{T}\right).$$

By replacing the sum by an integral we get (A.15). Using an identical argument we obtain (A.16), see Lee and Subba Rao (2011) for details. (iii) follows from using Assumption 2.3(L5), and integration by parts, we omit the details. \hfill \Box

Lemma A.9

Suppose that Assumption 2.1 holds, then we have

$$\sqrt{T}c_{j_1,j_2}(r, \ell) = \sqrt{T}c_{j_1,j_2}(r, \ell) + \text{Op} \left(\frac{1}{b\sqrt{T}} + b + b^2 T^{1/2}\right).$$  \hspace{1cm} (A.17)

Under Assumption 2.3 we have

$$\sqrt{T}c_{j_1,j_2}(r, \ell) = \sqrt{T}c_{j_1,j_2}(r, \ell) + \sqrt{T}S_{T,j_1,j_2}(r, \ell) + \sqrt{T}B_{T,j_1,j_2}(r, \ell) + \text{Op} \left(\frac{1}{b\sqrt{T}} + b + b^2 \sqrt{T}\right).$$  \hspace{1cm} (A.18)
PROOF. To prove the result we use the expansion (A.6)

\[
\sqrt{T}(\tilde{c}_{j_1,j_2}(r, \ell) - \bar{c}_{j_1,j_2}(r, \ell)) = A_{1,1} + A_{1,2} + O_p(A_2) + O_p(B_2) + \sqrt{T}(S_{T,j_1,j_2}(r, \ell) + B_{T,j_1,j_2}(r, \ell)),
\]

where \(A_{1,1}\) and \(A_{1,2}\) are defined in (A.4), \(A_2\) and \(B_2\) are defined in (A.7) and \(S_{T,j_1,j_2}(r, \ell)\) and \(B_{T,j_1,j_2}(r, \ell)\) are defined in (A.5). Below we derive bounds for \(A_{1,1}, \ldots, B_2\).

First we derive bounds under Assumption 2.1 (stationarity). We use Lemma A.4 to show that \(|A_{1,1}| = O_p(\frac{1}{b \sqrt{T}})\). By Lemma A.8(i) we have \(|\sqrt{T}B_{T,j_1,j_2}(r, \ell)| = O(T^{-1/2})\). From Lemma A.8(i) we see that the conditions in Lemma A.7 are satisfied, therefore \(|A_{1,2}| = O_p(b)\). By using Lemmas A.5 and A.6 we have \(|A_2| = O_p(b^{1/2} + b^2 \sqrt{T} + b)\). Finally we use Lemmas A.6, and A.8 to show that \(|\sqrt{T}S_{T,j_1,j_2}(r, \ell)| = O_p(b^2 T^{-1/2})\) and \(B_2 = O_p(\frac{1}{b T^{1/2}})\). Thus we obtain (A.17).

To prove (A.18) (under the nonstationary Assumption 2.3), we use the same procedure as above, to obtain the same bounds for \(A_{1,1}, A_{1,2}\) and \(A_2\). However, in the nonstationary set-up we do not have \(\mathbb{E}(J_k J_{k+r}) = O(T^{-1})\), thus we do not have the same bounds for \(|\sqrt{T}S_{T,j_1,j_2}(r, \ell)| + \sqrt{T}B_{T,j_1,j_2}(r, \ell)\) and \(B_2\) as in the stationary case. Instead we use Lemma A.8(ii) to obtain \(|B_2| = O_p(b^{-1} T^{-1/2})\). However, the terms \(\sqrt{T}S_{T,j_1,j_2}(r, \ell) = \sqrt{T}B_{T,j_1,j_2}(r, \ell)\) are not asymptotically negligible. Indeed it can be shown that, \(\sqrt{T}S_{T,j_1,j_2}(r, \ell) = O_p(1)\) and \(\sqrt{T}B_{T,j_1,j_2}(r, \ell) = O(b T^{1/2})\). Thus we obtain (A.18). \(\Box\)

**Proof of Theorems 2.1 and 2.4** The proof of Theorems 2.1 and 2.4 follows immediately from Lemma A.9. \(\Box\)

### A.2 Proof of Theorem 2.2

Throughout the proof, we can 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.

**Lemma A.10**

Suppose Assumption 2.1(P1,P2) is satisfied. 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

\[
TCov(\tilde{c}_{j_1,j_2}(r_1, \ell_1), \tilde{c}_{j_3,j_4}(r_2, \ell_2)) = \{\delta_{j_1,j_3}\delta_{j_2,j_4}\delta_{\ell_1,\ell_2} + \delta_{j_1,j_4}\delta_{j_2,j_3}\delta_{\ell_1,-\ell_2}\} \delta_{r_1,r_2} + \kappa^{(\ell_1,\ell_2)}(j_1, j_2, j_3, j_4) \delta_{r_1, r_2} + O\left(\frac{1}{T}\right),
\]

\[
TCov(\tilde{c}_{j_1,j_2}(r_1, \ell_1), \bar{c}_{j_3,j_4}(r_2, \ell_2)) = O\left(\frac{1}{T}\right),
\]

\[
TCov(\bar{c}_{j_1,j_2}(r_1, \ell_1), \tilde{c}_{j_3,j_4}(r_2, \ell_2)) = O\left(\frac{1}{T}\right),
\]

\[
TCov(\bar{c}_{j_1,j_2}(r_1, \ell_1), \bar{c}_{j_3,j_4}(r_2, \ell_2)) = \{\delta_{j_1,j_3}\delta_{j_2,j_4}\delta_{\ell_1,\ell_2} + \delta_{j_1,j_4}\delta_{j_2,j_3}\delta_{\ell_1,-\ell_2}\} \delta_{r_1,r_2} + \kappa^{(\ell_1,\ell_2)}(j_1, j_2, j_3, j_4) \delta_{r_1, r_2} + O\left(\frac{1}{T}\right),
\]

where \(\delta_{jk} = 1\) if \(j = k\) and \(\delta_{jk} = 0\) otherwise.
Proof. Straightforward calculations give

$$TCov(c_{j_1;j_2}(r_1, \ell_1), c_{j_3;j_4}(r_2, \ell_2))$$

$$= \frac{1}{T} \sum_{k_1,k_2=1}^{T} \sum_{s_1,s_2,s_3,s_4=1}^{d} L_{j_1s_1} (\omega_{k_1}) L_{j_2s_2} (\omega_{k_1+r_1}) L_{j_3s_3} (\omega_{k_2}) L_{j_4s_4} (\omega_{k_2+r_2})$$

$\text{Cov}(J_{k_1,s_2} J_{k_1+r_1,s_2}, J_{k_2,s_3} J_{k_2+r_2,s_4}) \exp(i\ell_1\omega_{k_1} - i\ell_2\omega_{k_2}).$

and by using the identity $\text{cum}(Z_1, Z_2, Z_3, Z_4) = \text{Cov}(Z_1Z_2, Z_3Z_4) - \text{E}(Z_1Z_3)\text{E}(Z_2Z_4) - \text{E}(Z_1Z_4)\text{E}(Z_2Z_3),$ for complex-valued and zero mean random variables $Z_1, Z_2, Z_3, Z_4$, we get

$$TCov(c_{j_1;j_2}(r_1, \ell_1), c_{j_3;j_4}(r_2, \ell_2))$$

$$= \frac{1}{T} \sum_{k_1,k_2=1}^{T} \sum_{s_1,s_2,s_3,s_4=1}^{d} L_{j_1s_1} (\omega_{k_1}) L_{j_2s_2} (\omega_{k_1+r_1}) L_{j_3s_3} (\omega_{k_2}) L_{j_4s_4} (\omega_{k_2+r_2})$$

$$\{ \mathbb{E}(J_{k_1,s_1} J_{k_2,s_2}) \mathbb{E}(J_{k_1+r_1,s_1} J_{k_2+r_2,s_2}) + \mathbb{E}(J_{k_1,s_1} J_{k_2,s_2}) \mathbb{E}(J_{k_1+r_1,s_1} J_{k_2+r_2,s_2}) \} \exp(i\ell_1\omega_{k_1} - i\ell_2\omega_{k_2})$$

$$:= I + II + III.$$
where \( \exp(-i\ell_2\omega_r) \to 1 \) 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

\[
III = \frac{1}{T} \sum_{k_1, k_2=1}^{T} \sum_{s_1, s_2, s_3, s_4=1}^{d} \left( L_{j_1 s_1}(\omega_1) L_{j_2 s_2}(\omega_{k_1+r_1}) L_{j_3 s_3}(\omega_{k_2}) L_{j_4 s_4}(\omega_{k_2+r_2}) \exp(i\ell_1\omega_1 - i\ell_2\omega_2) \right) \\
\times \left( \frac{2\pi}{T} \sum_{s_1, s_2, s_3, s_4=1}^{d} L_{j_1 s_1}(\omega_1) L_{j_2 s_2}(\omega_{k_1+r_1}) L_{j_3 s_3}(\omega_{k_2}) L_{j_4 s_4}(\omega_{k_2+r_2}) \exp(i\ell_1\omega_1 - i\ell_2\omega_2) \right)
\]

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 (2.7). In the computations for the second and the third part a \( \delta_{r_1, r_2} \) crops-up which is zero due to strictly positive lags \( r_1, r_2 \).

**PROOF of Theorem 2.2**

(i) Considering the entries of \( \mathbf{C}_T(r, \ell) \), we get

\[
\mathbb{E}(\bar{c}_{j_1, j_2}(r, \ell)) = \frac{1}{T} \sum_{k=1}^{T} \sum_{s_1, s_2=1}^{d} L_{j_1 s_1}(\omega_1) \mathbb{E} \left( J_{k, s_1} J_{k+r, s_2} \right) L_{j_2 s_2}(\omega_{r + 1}) \exp(i\omega_k \ell)
\]

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

**A.3 Proofs of Theorems 2.3 and 2.5**

**Lemma A.11**

Suppose \( \mathbf{f}(\omega) \) is a non-singular matrix and the second derivatives of the elements of \( \mathbf{f}(\omega) \) with respect to \( \omega \) are bounded. Let \( A_{j_1 j_2 j_3 j_4} \) be defined as in (A.1), then for all \( 1 \leq j_1, j_2, j_3, j_4 \leq d \) and all \( z \in [0, 2\pi] \) we have

\[
\left| \frac{\partial^2 A_{j}(\mathbf{f}(\omega)', f(\omega + z)')}{\partial \omega^2} \right| < \infty,
\]

and

\[
\sup_{z} \left| a_{j_1}^\omega(z) \right| \leq \frac{C}{|r|^2} \quad \text{and} \quad \sup_{z} \left| d_{j_2}^\omega(z) \right| \leq \frac{C}{|r|^2}
\]
where
\[
a_j(r) = \frac{1}{2\pi} \int_0^{2\pi} A_j(f(\omega)', f(\omega + z)') \exp(ir\omega) d\omega,
\]
and
\[
d_{j_1,s_1,j_2,s_2,z}(r) = \frac{1}{2\pi} \int_0^{2\pi} \nabla_{\omega,s} A_j(f(\omega)', f(\omega + z)') \exp(ir\omega) d\omega,
\]
with \( f_{\omega,z} = (f(\omega), f(\omega + z)) \) and \( C \) is a finite constant.

**PROOF.** Implicit differentiation gives
\[
\frac{\partial L_{js}(f(\omega))}{\partial \omega} = \frac{\partial f(\omega)'}{\partial \omega} \nabla f_{\omega,s}(f(\omega))
\]
and
\[
\frac{\partial^2 L_{js}(f(\omega))}{\partial \omega^2} = \frac{\partial^2 f(\omega)'}{\partial \omega^2} \nabla^2 f_{\omega,s}(f(\omega)) + \frac{\partial f(\omega)'}{\partial \omega} \nabla f_{\omega,s}(f(\omega)) \frac{\partial f(\omega)}{\partial \omega}.
\]
By using Lemma A.1 we have that \( \sup_{\omega} |\frac{\partial L_{js}(f(\omega))}{\partial \omega}| < \infty \) and \( \sup_{\omega} |\frac{\partial^2 L_{js}(f(\omega))}{\partial \omega^2}| < \infty \). Since under the stated assumptions, the derivatives \( |\frac{\partial f(\omega)'}{\partial \omega}| \) and \( |\frac{\partial^2 f(\omega)'}{\partial \omega^2}| \) are bounded, we have that \( \|\frac{\partial L_{js}(f(\omega))}{\partial \omega}\| \) and \( \|\frac{\partial^2 L_{js}(f(\omega))}{\partial \omega^2}\| \) are also bounded. Therefore, we obtain (A.19). To prove \( \sup_z |a_{j,z}(r)| \leq C|r|^{-2} \) we use (A.19) and apply twice integration by parts to obtain a bound on the Fourier coefficients (see for example, (Briggs & Henson, 1997), Theorem 6.2). We use the same method to obtain \( \sup_z |d_{j,z}(r)| \leq C|r|^{-2} \).

**Lemma A.12**

Suppose that either Assumption 2.1 (P1-P3, P5) or Assumption 2.3 (L1-L6, L8) holds (in the stationary case we let \( X_1 = X_{t,T} \)). Then we have
\[
\tilde{c}_{j_1,j_2}(r, \ell) = \frac{1}{T} \sum_{t,\tau=1}^T X_{t,T} G_{\omega}(t - \tau + \ell, X_{t,T} \exp(-i\tau\omega_T) + O_p \left( \frac{1}{T} \right),
\]
where \( G_{\omega}(k) = \int_0^1 \mathbf{L}_{j_1,\omega} \mathbf{L}_{j_2,\omega} \exp(i\omega k) d\omega = \sum_{s_1,s_2=1}^d a_{j_1,s_1,j_2,s_2,\omega}(k), |G_{\omega}(k)| \leq C/|k|^2 \).

**PROOF.** Expanding \( \tilde{c}_{j_1,j_2}(r, \ell) \) we have
\[
\tilde{c}_{j_1,j_2}(r, \ell) = \frac{1}{T} \sum_{k=1}^T X_{t,T} (\omega_{k+\ell}^T \mathbf{L}_{j_2,\omega_{k+\ell}} \mathbf{L}_{j_1,\omega_{k}}) \exp(ik\omega_{k}) = \frac{1}{T} \sum_{t,\tau=1}^T \left( \frac{1}{T} \sum_{k=1}^T \mathbf{L}_{j_2,\omega_{k+\ell}} \mathbf{L}_{j_1,\omega_{k}} \exp(i\omega_{k}(t - \tau + \ell)) \right) X_{t,T} \exp(-i\tau\omega_T).
\]
Finally, we replace \( \frac{1}{T} \sum_{t=1}^T \mathbf{L}_{j_2,\omega_{k+\ell}} \mathbf{L}_{j_1,\omega_{k}} \exp(i\omega_{k}(t - \tau)) \) with its limit \( G_{\omega}(t - \tau) \), and by using (A.20) obtain the bounds on \( G_{\omega}(k) \).

\[33\]
PROOF of Theorem 2.3
We first prove (2.11). By using Lemma A.12 we have
\[
\sqrt{T} \tilde{c}_{j_1,j_2}(r, \ell) = \sqrt{T} \tilde{c}_{j_1,j_2}(r, \ell) + o_p(1)
\]
\[
= \frac{1}{\sqrt{T}} \sum_{t,\tau=1}^{T} X'_t G_{\omega_r}(t - \tau + \ell) X_{\tau} \exp(-i\tau \omega_r) + o_p(1) := \sqrt{T} S_T + o_p(1).
\]
From the above we observe that the real and imaginary parts of \(\sqrt{T} S_T\) are quadratic forms, which satisfy the assumptions in Lee and Subba Rao (2011), Corollary 2.2. By the Cramer-Wold device, we have asymptotic normality of \(\sqrt{T} \tilde{c}_{j_1,j_2}(r, \ell)\).

To prove (2.13), we use the same argument as above for linear combinations of \(\tilde{c}_{j_1,j_2}(r, \ell)\), this together with the Cramer-Wold device gives (2.13). □

PROOF of Theorem 3.1 Follows immediately from Theorem 2.3.

PROOF of Lemma 2.1 The proof of Lemma 2.1 follows immediately from Lemma A.8(iii). □

We now derive the distribution of \(\tilde{c}_{j_1,j_2}(r, \ell)\) under the assumption of local stationarity. To prove this we expand \(f_{k,r}\) in \(S_{T,j_1,j_2}(r, \ell)\) (defined in (A.5)) to obtain
\[
S_{T,j_1,j_2}(r, \ell) = \sum_{s_1,s_2=1}^{d} \frac{1}{T} \sum_{t,\tau=1}^{T} \lambda_b(t - \tau) g(X'_t X'_\tau) \exp(i(t - \tau) \omega_r) \nabla f_{k,r} A_{j_1,s_1,j_2,s_2}(f_{k,r}) + o_p\left(1 \sqrt{T}\right)
\]
(A.22)
where the 2\(d(d + 1)\)-dimensional random vector \(g(X'_t X'_\tau)\) is defined as
\[
g(X'_t X'_\tau) = \left(\begin{array}{c}
\text{vech}(X'_t X'_\tau) \\
\text{vech}(X'_t X'_\tau) \exp(i(t - \tau) \omega_r)
\end{array}\right) - \mathbb{E}\left(\begin{array}{c}
\text{vech}(X'_t X'_\tau) \\
\text{vech}(X'_t X'_\tau) \exp(i(t - \tau) \omega_r)
\end{array}\right)
\]
and \(d_{j_1,s_1,j_2,s_2,\omega_r}(t - \tau - \ell)\) is defined in (A.20).

PROOF of Theorem 2.5 Theorem 2.4 implies that
\[
\tilde{c}_{j_1,j_2}(r, \ell) - B_{T,j_1,j_2}(r, \ell) = (\tilde{c}_{j_1,j_2}(r, \ell) - B_{T,j_1,j_2}(r, \ell)) + S_{T,j_1,j_2}(r, \ell) + o_p\left(\frac{1}{\sqrt{T}}\right).
\]
Since \(\tilde{c}_{j_1,j_2}(r, \ell)\) and \(S_{T,j_1,j_2}(r, \ell)\) can be written as quadratic forms we use Lee and Subba Rao (2011), Corollary 2.2, to prove (2.18) (we omit the details). □

PROOF of Theorem 3.2 The proof follows immediately from Theorem 2.5. □

A.4 The stationary bootstrap
In this section we prove the results in Section 4.
PROOF of Lemma 4.1. To simplify notation we prove the result for the case $d = 1$, the proof $d > 1$ is identical. By using the triangle inequality we have

$$\left| \hat{h}_n(\omega_1, \ldots, \omega_{k-1}) - f_n(\omega_1, \ldots, \omega_{k-1}) \right| \leq I + II + III + IV,$$

where

$$I = \frac{1}{(2\pi)^{n-1}} \sum_{r_1, \ldots, r_{n-1} = -T}^{T} (1 - p)^{\max(r_0, 0) - \min(r_0, 0)} \left| \mathbb{E}(\tilde{\kappa}_n(r_1, \ldots, r_{n-1}) - \kappa_n(r_1, \ldots, r_{n-1})) \right|$$

$$II = \frac{1}{(2\pi)^{n-1}} \sum_{r_1, \ldots, r_{n-1} = -T}^{T} (1 - p)^{\max(r_0, 0) - \min(r_0, 0)} \left| \mathbb{E}(\tilde{\kappa}_n(r_1, \ldots, r_{n-1}) - \kappa_n(r_1, \ldots, r_{n-1})) \right|$$

$$III = \frac{1}{(2\pi)^{n-1}} \sum_{r_1, \ldots, r_{n-1} = -T}^{T} \left| (1 - p)^{\max(r_0, 0) - \min(r_0, 0)} - 1 \right| \left| \kappa_n(r_1, \ldots, r_{n-1}) \right|$$

$$IV = \frac{1}{(2\pi)^{n-1}} \sum_{|r_1| \ldots or ..., |r_{n-1}| > T} \left| \kappa_n(r_1, \ldots, r_{n-1}) \right|.$$

Under the stated assumptions we have $\text{var}(I) = O\left(\frac{1}{T^{2(n-1)}}\right)$. It is straightforward to show that $\left| \mathbb{E}(\tilde{\kappa}_n(r_1, \ldots, r_{n-1}) - \kappa_n(r_1, \ldots, r_{n-1})) \right| \leq c^{\max(r_0, 0) - \min(r_0, 0)}$, using this gives $\left| II \right| \leq c T \sum_{r_1, \ldots, r_{n-1} = -T}^{T} (1 - p)^{\max(r_0, 0) - \min(r_0, 0)} \max\left(0, (1 - p)^{\max(r_0, 0) - \min(r_0, 0)}\right) = O\left(\frac{1}{T^{2(n-1)}}\right)$. Substituting the bound $|1 - (1 - p)^k| \leq K lp$, into $III$ gives $\left| III \right| = O(p)$. Finally, by using similar arguments to those used in Brillinger (1981), Theorem 4.3.2, we have $IV = O\left(\frac{1}{T}\right)$. Altogether this gives the required result. □

We recall that the bootstrap variance involves the DFT of the bootstrap sample, thus we start by evaluating the cumulants of the bootstrap DFT.

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

$$\text{cum}\ast(X_{t_{1},j_{1}}, X_{t_{2},j_{2}}, \ldots, X_{t_{n},j_{n}}) = \text{cum}\ast(X_{t_{1},j_{1}}, X_{t_{2},j_{2}}, \ldots, X_{t_{n},j_{n}} | L < |t_n|)P(L < |t_n|)$$

$$+ \text{cum}\ast(X_{t_{1},j_{1}}, X_{t_{2},j_{2}}, \ldots, X_{t_{n},j_{n}} | L \geq |t_n|)P(L \geq |t_n|).$$

We observe that $\text{cum}\ast(X_{t_{1},j_{1}}, X_{t_{2},j_{2}}, \ldots, X_{t_{n},j_{n}} | L < |t_n|) = 0$ (since the random variables in separate blocks are conditionally independent), $\text{cum}\ast(X_{t_{1},j_{1}}, X_{t_{2},j_{2}}, \ldots, X_{t_{n},j_{n}} | L \geq |t_n|) = \tilde{\kappa}_{j_{1},\ldots,j_{n}}(t_2 - t_1, \ldots, t_n - t_1)$ and $P(L \geq |t_n|) = (1 - p)^{|t_n|}$, thus altogether we have $\text{cum}\ast(X_{t_{1},j_{1}}, X_{t_{2},j_{2}}, \ldots, X_{t_{n},j_{n}}) = (1 - p)^{|t_n|} \tilde{\kappa}_{j_{1},\ldots,j_{n}}(t_2 - t_1, \ldots, t_n - t_1)$.

To prove (ii), we first bound the difference $\tilde{\kappa}_{j_{1},\ldots,j_{n}}(r_2, \ldots, r_n) - \tilde{\kappa}_{j_{1},\ldots,j_{n}}(r_2, \ldots, r_n)$. Without loss of generality we consider the case $1 \leq r_1 \leq r_2 \cdots \leq r_n < T$. By definition we have

$$\tilde{\mu}_{j_{1},\ldots,j_{n}}(r_2, \ldots, r_n) - \tilde{\mu}_{j_{1},\ldots,j_{n}}(r_2, \ldots, r_n) = \frac{1}{T} \sum_{t=T-r_n+1}^{T} Y_t \prod_{i=2}^{n} Y_{t+r_i},$$

therefore

$$\left\| \tilde{\mu}_{j_{1},\ldots,j_{n}}(r_2, \ldots, r_n) - \tilde{\mu}_{j_{1},\ldots,j_{n}}(r_2, \ldots, r_n) \right\|_{q/n} \leq \frac{|r_n|}{T} \sup_{t,T} \|X_{t,T}\|_{q}.$$
Substituting this bound into (4.5) gives

\[
\left\| \hat{\kappa}_{j_1,\ldots,j_n}(r_2, \ldots, r_n) - \hat{\kappa}_{j_1,\ldots,j_n}(r_2, \ldots, r_n) \right\|_{q/n} \leq C \frac{|r_n|}{T} \sup_{t,T} \|X_{iT}\|_q,
\]  

(A.23)

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

We use this result to prove (ii). Substituting (i) into \( \text{cum}^*(J_{T,s_1}^*(\omega_{j_1}), \ldots, J_{T,s_n}^*(\omega_{j_n})) \) gives

\[
\text{cum}^*(J_{T,j_1}^*(\omega_{k_1}), \ldots, J_{T,j_n}^*(\omega_{k_n}))
= \frac{1}{(2\pi T)^{n/2}} \sum_{t_1, \ldots, t_n = 1}^{T} \hat{\kappa}_{n;j_1,\ldots,j_n}(t_2 - t_1, \ldots, t_n - t_1)e^{-it_1\omega_{k_1}-\ldots-it_n\omega_{k_n}}
= \frac{1}{(2\pi T)^{n/2}} \sum_{t_1, \ldots, t_n = 1}^{T} (1 - p)^{\max_i((t_i - t_1),0) - \min_i((t_i - t_1),0)} \hat{\kappa}_{n;j_1,\ldots,j_n}(t_2 - t_1, \ldots, t_n - t_1)e^{-it_1\omega_{k_1}-\ldots-it_n\omega_{k_n}}.
\]

To simplify notation define \( g(t) = \max_i((t_i - t_1),0) - \min_i((t_i - t_1),0) \). By changing the range of the above summand, we obtain

\[
\text{cum}^*(J_{T,j_1}^*(\omega_{k_1}), \ldots, J_{T,j_n}^*(\omega_{k_n}))
= \frac{1}{(2\pi T)^{n/2}} \sum_{r_2, \ldots, r_n = -T + 1}^{T - 1} (1 - p)^{g(t)} \hat{\kappa}_{n;j_1,\ldots,j_n}(r_2, \ldots, r_n)e^{-ir_2\omega_{k_2}-\ldots-ir_n\omega_{k_n}}
\times \sum_{\min_i(r_i,0)+1 \leq t \leq T - \max_i(r_i,0)} e^{-it_1(\omega_{k_1}+\omega_{k_2}+\ldots+\omega_{k_n})}.
\]

Replacing \( \hat{\kappa}_{n;j_1,\ldots,j_n}(r_2, \ldots, r_n) \) in the above with \( \hat{\kappa}_{n;j_1,\ldots,j_n}(r_2, \ldots, r_n) \) gives

\[
\text{cum}^*(J_{T,j_1}^*(\omega_{k_1}), \ldots, J_{T,j_n}^*(\omega_{k_n}))
= \frac{1}{(2\pi T)^{n/2}} \sum_{r_2, \ldots, r_n = -T + 1}^{T - 1} (1 - p)^{g(t)} \hat{\kappa}_{n;j_1,\ldots,j_n}(r_2, \ldots, r_n)e^{-ir_2\omega_{k_2}-\ldots-ir_n\omega_{k_n}}
\times \sum_{\min_i(r_i,0)+1 \leq t \leq T - \max_i(r_i,0)} e^{-it_1(\omega_{k_1}+\omega_{k_2}+\ldots+\omega_{k_n})} + R_{1,1}.
\]

By using the inequality

\[
\sum_{1 \leq r_2 < \ldots < r_n \leq T} (1 - p)^{r_n} r_n \leq \sum_{r_n} (1 - p)^{r_n} r_n^{n-1} = \frac{1}{p^n} \quad (\text{A.24})
\]

we have \( \|R_{1,1}\|_{q/n} = O\left(\frac{p^n}{(2\pi T)^{n/2}}\right) \). Finally, replacing the range of the inner sum from \( \min_i(r_i,0) + 1 \leq t \leq T - \max_i(r_i,0) \) to \( 1 \leq t \leq T \) gives

\[
\text{cum}^*(J_{T,j_1}^*(\omega_{k_1}), \ldots, J_{T,j_n}^*(\omega_{k_n}))
= \frac{1}{(2\pi T)^{n/2}} \sum_{r_2, \ldots, r_n = -T + 1}^{T - 1} (1 - p)^{g(t)} \hat{\kappa}_{n;j_1,\ldots,j_n}(r_2, \ldots, r_n)e^{-ir_2\omega_{k_2}-\ldots-ir_n\omega_{k_n}}
\times \sum_{t=1}^{T} e^{-it(\omega_{k_1}+\omega_{k_2}+\ldots+\omega_{k_n})} + R_{1,1} + R_{1,2},
\]

36
where
\[
|R_{1,2}| \leq \left| \frac{1}{(2\pi T)^{n/2}} \sum_{r_2,\ldots,r_n=1}^T (1-p)^{\theta(t)} \hat{r}_{n;j_1,\ldots,j_n}(r_2,\ldots,r_n) e^{-i\sum_{k=1}^n \omega_k r_k} \right| \\
\times \sum_{1 \leq t \leq \min(|r_1|,0)+1} (1-p)^{\theta(t)} |\hat{r}_{n;j_1,\ldots,j_n}(r_2,\ldots,r_n)| \sum_{r_1=1}^T 1 \\
\leq \frac{1}{(2\pi T)^{n/2}} \sum_{r_2,\ldots,r_n=-T+1}^{T-1} (1-p)^{\theta(t)} |\hat{r}_{n;j_1,\ldots,j_n}(r_2,\ldots,r_n)| \sum_{1 \leq t \leq \min(|r_1|,0)+1} T_{-\max(|r_1|,0)} \leq T \\
\leq \frac{2}{(2\pi T)^{n/2}} \sum_{r_2,\ldots,r_n=-T+1}^{T-1} (1-p)^{\theta(t)} |\hat{r}_{n;j_1,\ldots,j_n}(r_2,\ldots,r_n)| \max_i (|r_i|) \\
\leq \frac{2n!}{(2\pi T)^{n/2}} \sum_{1 \leq r_2 < \ldots < r_n \leq T} (1-p)^{\max(|r_i|)} |\hat{r}_{n;j_1,\ldots,j_n}(r_2,\ldots,r_n)|.
\]

By using (A.24) and \(\|\hat{r}_{n;j_1,\ldots,j_n}(r_2,\ldots,r_n)\|_{q/n} \leq C\) we have \(\|R_{1,2}\|_{q/n} = O(\frac{n!}{T^{n/2}p^n})\). Thus we have \(\|R_1\|_{q/n} = O(\frac{n!}{T^{n/2}p^n})\) and the desired result. \(\square\)

**Corollary A.2**

Let us suppose we observe \(\{X_t\}\), and \(\sup_t \|X_t\|_s < \infty\), and \(\{X_t^*\}\) is the stationary vector bootstrap sample. Then we have

(i) \(\text{cum}^*(J^*_{T,j_1}(\omega_k_1), J^*_{T,j_2}(\omega_k_2)) = f_{2;j_1,j_2}(\omega_k_1) I(k_1 = -k_2) + O_p(\frac{1}{T^{2p}})\)

(ii) \(\text{cum}^*(J^*_{T,j_1}(\omega_k_1), J^*_{T,j_2}(\omega_k_2), J^*_{T,j_3}(\omega_k_3), J^*_{T,j_4}(\omega_k_4)) = \frac{1}{4} f_{4;j_1,\ldots,j_4}(\omega_k_1, \omega_k_2, \omega_k_3, \omega_k_4) I(k_4 = -k_1 - k_2 - k_3) + O(\frac{1}{T^{2p}})\),

where \(I\) denotes the indicator variable.

**PROOF.** The results immediately follow from Theorem 4.1(ii). \(\square\)

**Lemma A.13**

Suppose Assumption 4.1 holds, let \(\hat{f}^*\) be defined in (4.1) and \(\sup_{t,T} \|X_{t,T}\|_{2s} < \infty\). Then we have

\[
\|\text{var}^*[\hat{f}^*_{s_1,s_2}(\omega_k)]\|_s = O\left(\frac{1}{bT} + \frac{1}{(Tp^2)^2}\right), \quad (A.25)
\]

\[
\|\text{cum}_4[\text{var}^*[\hat{f}^*_{s_1,s_2}(\omega_k)]]\|_s = O\left(\frac{1}{(bT)^2} + \frac{1}{(Tp^2)^2}\right) \quad (A.26)
\]

and

\[
\|E^*[\hat{f}^*_{s_1,s_2}(\omega_k)] - f_{s_1,s_2}(\omega_k)\|_s = O\left(\frac{1}{Tp^2} + \frac{1}{T^{1/2}p} + b\right). \quad (A.27)
\]

37
PROOF To prove (A.25) we expand \( \text{var}^*(\hat{f}_{s_1,s_2}(\omega)) \) to give

\[
\|\text{var}^*(\hat{f}_{s_1,s_2}(\omega_k))\|_s = \left\| \frac{1}{T^2} \sum_{l_1,l_2} K_b(\omega_k - \omega_{l_1}) K_b(\omega_k - \omega_{l_2}) \left[ \text{cov}(J^*_{l_1,s_1}, J^*_{l_2,s_2}) \text{cov}(J^*_{l_1,s_2}, J^*_{l_2,s_1}) \right] + \text{cov}(J^*_{l_1,s_1}, J^*_{l_2,s_2}) \text{cov}(J^*_{l_1,s_2}, J^*_{l_2,s_1}) + \text{cum}(J^*_{l_1,s_1}, J^*_{l_2,s_2}, J^*_{l_2,s_1}, J^*_{l_1,s_2}) \right\|_2
\]

where the above bound is obtained by using the near uncorrelatedness properties of the bootstrap DFT given in Theorem 4.1(ii) and Lemma 4.1.

We now prove (A.26), to reduce notation we consider the case \( s := s_1 = s_2 \), expanding the cumulant we obtain

\[
\|\text{cum}_4^*(\hat{f}_{s,s}(\omega_k))\|_s = \left\| \frac{1}{T^4} \sum_{l_1,l_2,l_3,l_4} \left( \prod_{i=1}^4 K_b(\omega_k - \omega_{l_i}) \right) \text{cum}(|J^*_{l_1,s}|^2, |J^*_{l_2,s}|^2, |J^*_{l_3,s}|^2, |J^*_{l_4,s}|^2) \right\|_s = O\left( \frac{1}{(bT)^2} + \frac{1}{(Tp)^2} \right),
\]

where the above bound is derived by expanding the cumulant of products in terms of products of cumulants and applying Theorem 4.1(ii) and Lemma 4.1.

To prove (A.27), we first use the Minkowski inequality to obtain

\[
\left\| E^*(\hat{f}_{s_1,s_2}(\omega_k)) - f_{s_1,s_2}(\omega_k) \right\|_s \leq \left\| \frac{1}{T} \sum_j K_b(\omega_k - \omega_j) \left[ E^*(J^*_{j,s_1}, J^*_{j,s_2}) - \tilde{h}_{2,s_1,s_2}(\omega_j) \right] \right\|_s + \left\| \frac{1}{T} \sum_j K_b(\omega_k - \omega_j) f_{s_1,s_2}(\omega_j) \right\|_s.
\]

Finally by using Lemma 4.1 we obtain

\[
\|E^*(\hat{f}_{s_1,s_2}(\omega_k)) - f_{s_1,s_2}(\omega_k)\|_s = O\left( \frac{1}{Tp^2} + \frac{1}{T^{1/2}p} + b \right),
\]

thus giving the required result.

\[
\text{A.27}
\]

Finally by using Lemma 4.1 we obtain

\[
\|E^*(\hat{f}_{s_1,s_2}(\omega_k)) - f_{s_1,s_2}(\omega_k)\|_s = O\left( \frac{1}{Tp^2} + \frac{1}{T^{1/2}p} + b \right),
\]

thus giving the required result.

\[
\text{A.27}
\]

Analogous to \( \hat{C}_{T}(r, \ell) \), direct analysis of the variance of \( \hat{C}_{T}^*(r, \ell) \) with respect to the bootstrap measure is extremely difficult because of \( \hat{L}^* \) in \( \hat{C}_{T}^*(r, \ell) \). However, analysis of \( \hat{C}_{T}^*(r, \ell) \) is much easier, therefore to show that the bootstrap variance converges to the true variance we will show that var\(*(\hat{C}_{T}^*(r, \ell)) \) can be replaced with var\(*(\hat{C}_{T}^*(r, \ell)) \), this allows to asymptotically study the properties of \( \hat{C}_{T}^*(r, \ell) \). To proof this result we require the following definitions.

\[
\bar{c}_{j_1,j_2}^*(r, \ell) = \frac{1}{T} \sum_{k=1}^T \sum_{s_1,s_2=1}^d A_{j_1,s_1,j_2,s_2} (\hat{f}_{k,r}^*) J^*_{k,s_1} \bar{J}^*_{k+r,s_2} \exp(i\ell \omega_k)
\]

\[
\bar{c}_{j_1,j_2}^+(r, \ell) = \frac{1}{T} \sum_{k=1}^T \sum_{s_1,s_2=1}^d A_{j_1,s_1,j_2,s_2} (\text{E}^*(\hat{f}_{k,r}^*)) J^*_{k,s_1} \bar{J}^*_{k+r,s_2} \exp(i\ell \omega_k)
\]

and

\[
\bar{c}_{j_1,j_2}^+(r, \ell) = \frac{1}{T} \sum_{k=1}^T \sum_{s_1,s_2=1}^d A_{j_1,s_1,j_2,s_2} (\hat{f}_{k,r}) J^*_{k,s_1} \bar{J}^*_{k+r,s_2} \exp(i\ell \omega_k).
\]

(A.28)
Lemma A.14
Suppose that Assumptions 2.1 and 4.1 hold. Then we have

\[
\begin{align*}
T & \left( E^* \left[ c_{j_1,j_2}^*(r, \ell_1) \right] E^* \left[ c_{j_3,j_4}^*(r, \ell_2) \right] - E^* \left[ c_{j_1,j_2}^*(r, \ell_1) \right] E^* \left[ c_{j_3,j_4}^*(r, \ell_2) \right] \right) = O_p \left( a(T, b, p) \right) \tag{A.29} \\
T & \left( E^* \left[ c_{j_1,j_2}^*(r, \ell_1) \right] E^* \left[ c_{j_3,j_4}^*(r, \ell_2) \right] - E^* \left[ c_{j_1,j_2}^*(r, \ell_1) \right] E^* \left[ c_{j_3,j_4}^*(r, \ell_2) \right] \right) = O_p \left( a(T, b, p) \right) \tag{A.30} \\
T & \left( E^* \left[ c_{j_1,j_2}^*(r, \ell_1) \right] c_{j_3,j_4}^*(r, \ell_2) - E^* \left[ c_{j_1,j_2}^*(r, \ell_1) \right] c_{j_3,j_4}^*(r, \ell_2) \right) = O_p \left( a(T, b, p) \right) \tag{A.31} \\
T & \left( E^* \left[ c_{j_1,j_2}^*(r, \ell_1) \right] c_{j_3,j_4}^*(r, \ell_2) - E^* \left[ c_{j_1,j_2}^*(r, \ell_1) \right] c_{j_3,j_4}^*(r, \ell_2) \right) = O_p \left( a(T, b, p) \right) \tag{A.32}
\end{align*}
\]

where \( a(T, b, p) = \frac{1}{bT_p^2} + \frac{1}{T_p^2} + \frac{1}{T} + b + \frac{1}{T_p^2} + \frac{1}{T^{1/2} p} \).

Proof. To simplify notation we consider the case \( d = 1 \), this avoids the use of the subscripts \( j_1, \ldots, j_4 \), the proof for other cases is the same. In addition without loss of generality we consider the case \( \ell_1 = \ell_2 = \ell \) and \( r_1 = r_2 = r \). We first prove (A.29). Recalling that the only difference between \( \hat{c}^*(r, \ell) \) and \( \hat{c}^*(r, \ell) \) is that \( A(f_{k,r}^*) \) is replaced with \( A(E^*(f_{k,r}^*)) \), the difference between their expectations squared (with respect to the stationary bootstrap measure) is

\[
T \left| \mathbb{E}^*(\hat{c}^*(r, \ell)) \right|^2 - \left| \mathbb{E}^*(\hat{c}^*(r, \ell)) \right|^2 = \frac{1}{T} \sum_{k_1,k_2} \left( \mathbb{E}^* \left[ A(E^*(f_{k_1,r}^*))J_{k_1}^*J_{k_1+r}^* \right] \mathbb{E}^* [A(E^*(f_{k_2,r}^*))J_{k_2}^*J_{k_2+r}^*] \right) - A(\mathbb{E}^*(\hat{f}_{k,r}^*))A(\mathbb{E}(f_{k,r}^*)) \mathbb{E}^*(J_{k_1}^*J_{k_1+r}^*) \mathbb{E}^*(J_{k_2}^*J_{k_2+r}^*)
\]

\[
= \frac{1}{T} \sum_{k_1,k_2} \left( \mathbb{E}^* \left[ a_{k_1} I_{k_1,r}^* \right] \mathbb{E}^* [a_{k_2} I_{k_2,r}^*] - \hat{a}_{k_1} \hat{a}_{k_2} \mathbb{E}[I_{k_1,r}^*] \mathbb{E}[I_{k_2,r}^*] \right), \tag{A.33}
\]

where

\[
a_k^* = A(\hat{f}_{k,r}^*), \quad \hat{a}_k = A(E(\hat{f}_{k,r}^*)), \quad \hat{f}_{k,r}^* = \hat{f}_{k,r}^* - (E^*(\hat{f}_{k,r}^*)) \text{ and } I_{k,r}^* = J_{k}^*J_{k+r}^*. \tag{A.34}
\]

To bound the above we use the Taylor expansion

\[
A(\hat{f}_{k,r}^*) = A(E^*(\hat{f}_{k,r}^*)) + (\hat{f}_{k,r}^* - E^*(\hat{f}_{k,r}^*))\nabla A(E^*(\hat{f}_{k,r}^*)) + \frac{1}{2} (\hat{f}_{k,r}^* - E^*(\hat{f}_{k,r}^*))\nabla^2 A(E^*(\hat{f}_{k,r}^*)) (\hat{f}_{k,r}^* - E^*(\hat{f}_{k,r}^*)),
\]

where \( \nabla \) and \( \nabla^2 \) denotes the first and second partial derivative with respect to \( \hat{f}_{k,r}^* \) and \( \hat{f}_{k,r}^* \) lies between \( \hat{f}_{k,r}^* \) and \( E^*(\hat{f}_{k,r}^*) \). To reduce cumbersome notation (and with a slight loss of accuracy, since it will not effect the calculation) we shall ignore that \( \hat{f}_{k,r}^* \) is a vector and use (A.34) to rewrite the Taylor expansion as

\[
a_k^* = \hat{a}_k + \hat{f}_{k} \frac{\partial \hat{a}_k}{\partial f} + \hat{f}_{k}^2 \frac{1}{2} \frac{\partial^2 \hat{a}_k}{\partial f^2}, \tag{A.35}
\]

where \( \hat{a}_k = \nabla^2 A(\hat{f}_{k,r}^*) \). Substituting (A.35) into (A.33) we obtain the decomposition

\[
T \left( \left| \mathbb{E}^*(\hat{c}^*(r, \ell)) \right|^2 - \left| \mathbb{E}^*(\hat{c}^*(r, \ell)) \right|^2 \right) = \sum_{i=1}^{8} I_i,
\]

39
where the terms \( \{ I_i \}_{i=1}^8 \) are

\[
I_1 = \frac{1}{T} \sum_{k_1, k_2} \tilde{a}_{k_1} \frac{\partial^2 \tilde{a}_{k_2}}{\partial f^2} \mathbb{E}^* (I_{k_1, r}) \mathbb{E}^* (I_{k_2, r}) e^{i\ell(\omega_{k_1} - \omega_{k_2})} \\
= \frac{1}{T} \sum_{k_1, k_2} \tilde{a}_{k_1} \frac{\partial a_{k_2}}{\partial f} \mathbb{E}^* (I_{k_1, r}) \mathbb{E}^* (I_{k_2, r})
\]

\[
I_2 = \frac{1}{T} \sum_{k_1, k_2} \frac{\partial a_{k_1}}{\partial f} \mathbb{E}^* (I_{k_1, r}) \mathbb{E}^* (\tilde{I}_{k_2, r})
\]

\[
I_3 = \frac{1}{2T} \sum_{k_1, k_2} \tilde{a}_{k_2} \mathbb{E}^* \left[ I_{k_1, r} \tilde{f}_{k_1} \frac{\partial^2 \tilde{a}_{k_1}}{\partial f^2} \right] \mathbb{E}^* (I_{k_2, r})
\]

\[
I_8 = \frac{1}{4T} \sum_{k_1, k_2} \mathbb{E}^* \left[ I_{k_1, r} \tilde{f}_{k_1} \frac{\partial^2 \tilde{a}_{k_1}}{\partial f^2} \right] \mathbb{E}^* \left[ I_{k_2, r} \tilde{f}_{k_2} \frac{\partial^2 \tilde{a}_{k_2}}{\partial f^2} \right]
\]

(with \( \tilde{I}_{k, r} = I_{k, r}^* - \mathbb{E}^* (I_{k, r}) \)) and \( I_4, I_5, I_6, I_7 \) are defined similarly. We first bound \( I_1 \). Writing \( \tilde{f}_k = \frac{1}{T} \sum_{j=1}^T K_b(\omega_k - \omega_j) \tilde{f}_{k, j} \) gives

\[
I_1 = \frac{1}{T} \sum_{k_1, k_2, j} K_b(\omega_{k_2} - \omega_j) \frac{\partial a_{k_1}}{\partial f} \mathbb{E}^* (I_{k_1, r}) \text{cov}^*(\tilde{I}_{k_2, r}, \tilde{I}_{j, 0}).
\]

By using uniform convergence result in Lemma A.1 we have

\[
|I_1| = O_p(1) \frac{1}{T} \sum_{k_1, k_2, j} |K_b(\omega_{k_2} - \omega_j)| |a_{k_1}| \frac{\partial a_{k_2}}{\partial f} \mathbb{E}^* (I_{k_1, r}) \text{cov}^*(\tilde{I}_{k_2, r}, \tilde{I}_{j, 0}),
\]

where \( a_k = A(f_{k, r}) \) and \( \frac{\partial a_{k_1}}{\partial f} = \nabla_f A(f_{k, r}) \). By expanding

\[
\text{cov}^*(\tilde{I}_{k_2, r}, \tilde{I}_{j, 0}) = \text{cum}^*(J_{k_2, r}^*, J_{j, 0}^*) + \text{cum}^*(J_{k_2, r}^*, J_{j, 0}^*) \text{cum}^*(J_{k_2, r}^*, J_{j, 0}^*) + \text{cum}^*(J_{k_2, r}^*, J_{j, 0}^*) + \text{cum}^*(J_{k_2, r}^*, J_{j, 0}^*)
\]

using Theorem 4.1 and taking expectations inside the above sum we have

\[
|I_2| = O_p(1) \frac{1}{T} \sum_{k_1, k_2, j_1, j_2} |K_b(\omega_{k_2} - \omega_{j_1})| |K_b(\omega_{k_2} - \omega_{j_2})| |\frac{\partial a_{k_1}}{\partial f}| |\frac{\partial a_{k_2}}{\partial f}| \mathbb{E}^* (I_{k_1, r}) \text{cov}^*(\tilde{I}_{k_2, r}, \tilde{I}_{j_1, 0}) \mathbb{E}^* (\tilde{I}_{k_2, r}, \tilde{I}_{j_2, 0})
\]

\[
= O_p(\frac{1}{T^2 p^2})
\]

To bound \( I_3 \) we use the Cauchy-Schwarz inequality and Assumption 4.1(B1)

\[
|I_3| \leq \frac{1}{T} \sum_{k_1, k_2} |\tilde{a}_{k_2}| \mathbb{E}^* (\tilde{f}_{k_1}^{1/2} \mathbb{E}^* I_{k_1, r}^{1/2} \mathbb{E}^* (\frac{\partial^2 \tilde{a}_{k_1}}{\partial f^2})^{1/3} \mathbb{E}^* (I_{k_2, r})|
\]

\[
= O_p(1) \frac{1}{T} \sum_{k_1, k_2} |\tilde{a}_{k_2}| \frac{\partial a_{k_1}}{\partial f} \mathbb{E}^* (\tilde{f}_{k_1}^{1/2} \mathbb{E}^* I_{k_1, r}^{1/2} \mathbb{E}^* (I_{k_2, r})|
\]

40
thus by using Lemma A.13 we obtain $|I_3| = O_p\left(\frac{1}{T^2}\right)$ Using a similar method we obtain

$$|I_8| \leq O_p\left(\frac{1}{T}\right) \sum_{k_1,k_2} \frac{1}{T} \sum_{k_1,k_2} \left| \frac{\partial^2 \tilde{a}_{k_1}}{\partial f^2} \right| \left| \frac{\partial^2 \tilde{a}_{k_2}}{\partial f^2} \right| E^*|f_{k_1}^*|^{1/2} B^* |f_{k_2}^*|^{1/2} E^*|f_{k_1}^*|^{1/2} B^* |f_{k_2}^*|^{1/2} = O\left(\frac{1}{b^2 T}\right).$$

Similar arguments can be used to obtain the same bounds for $I_4, \ldots, I_7$, which altogether gives (A.29).

To bound (A.30) we note that

$$\tilde{a}_k = a_k + f_k \frac{\partial a_k}{\partial f} + f_k \frac{\partial^2 \tilde{a}_k}{\partial f^2}, \quad (A.36)$$

where $a_k = A(f_{k,r})$ and $f_k = f_{k,r} - f_{k,r}$. Substituting (A.36) into (A.30) gives

$$T \left| \left| \mathbf{E}^*(\hat{c}^*(r, \ell)) \right|^2 - \left| \mathbf{E}^*(\hat{c}^*(r, \ell)) \right|^2 \right| = \sum_{i=1}^{8} II_i,$$

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

$$II_1 = \frac{2}{T} \sum_{k_1,k_2} \tilde{a}_{k_1} \frac{\partial \tilde{a}_{k_1}}{\partial f} f_{k_2} E^*[I_{k_1,r}^*] E^*[I_{k_2,r}^*] e^{i(\omega_{k_1} - \omega_{k_2})},$$

$$II_2 = \frac{1}{T} \sum_{k_1,k_2} \frac{\partial \tilde{a}_{k_1}}{\partial f} \frac{\partial \tilde{a}_{k_2}}{\partial f} f_{k_1} f_{k_2} E^*[I_{k_1,r}^*] E^*[I_{k_2,r}^*],$$

$$II_3 = \frac{2}{T} \sum_{k_1,k_2} \tilde{a}_{k_2} f_{k_1} \frac{\partial^2 \tilde{a}_{k_1}}{\partial f^2} E^*[I_{k_1,r}^*] E^*[I_{k_2,r}^*],$$

$$II_8 = \frac{1}{T} \sum_{k_1,k_2} f_{k_1} f_{k_2} \frac{\partial^2 \tilde{a}_{k_1}}{\partial f^2} \frac{\partial \tilde{a}_{k_2}}{\partial f} E^*[I_{k_1,r}^*] E^*[I_{k_2,r}^*],$$

and $II_4, II_5, II_6$ and $II_7$ are defined similarly. By using Theorem 4.1(ii) on each term we obtain (A.30).

To bound (A.31) we note that by making a Taylor expansion we have

$$\tilde{a}_k^*, \tilde{a}_k^* = \tilde{a}_k + \frac{\partial a_k}{\partial f} + \frac{\partial^2 \tilde{a}_k}{\partial f^2} + \frac{1}{2} \frac{\partial^2 \tilde{a}_k}{\partial f} + \frac{1}{2} \frac{\partial^2 \tilde{a}_k}{\partial f}$$

Substituting the above into $E^*|\hat{c}^*(r, \ell)|^2 - |E^*|\hat{c}^*(r, \ell)|^2$ gives

$$T \left( E^*|\hat{c}^*(r, \ell)|^2 - |E^*|\hat{c}^*(r, \ell)|^2 \right) = \sum_{i=1}^{3} III_i, \quad (A.37)$$

where

$$III_1 = \frac{2}{T} \sum_{k_1,k_2} \tilde{a}_{k_1} \frac{\partial \tilde{a}_{k_2}}{\partial f} E^*(I_{k_1,r}^* I_{k_2,r} \bar{f}_{k_2}),$$

$$III_2 = \frac{1}{T} \sum_{k_1,k_2} \tilde{a}_{k_1} E^*(I_{k_1,r}^* I_{k_2,r} \bar{f}_{k_2} f_{k_2} \partial \bar{a}_{k_2}),$$

$$III_3 = \frac{1}{T} \sum_{k_1,k_2} E^*(I_{k_1,r}^* I_{k_2,r} \bar{f}_{k_2} f_{k_2} \partial \bar{a}_{k_2}).$$
We first bound $III_1$, expanding $E^*(I^*_{k_1,r} I^*_{k_2,r} f_{k_2}^*)$ gives $III_1 = III_{1,1} + III_{1,2}$, where

$$III_{1,1} = \frac{1}{T} \sum_{k_1, k_2} \tilde{a}_{k_1} \frac{\partial \tilde{a}_{k_2}}{\partial f} c_{k_1}^*(I^*_{k_1,r}, I^*_{k_2,r}, f_{k_2}^*)$$

$$III_{1,2} = \frac{1}{T} \sum_{k_1, k_2} \left( \tilde{a}_{k_1} \frac{\partial \tilde{a}_{k_2}}{\partial f} E^*(I^*_{k_1,r}) c_{k_2}^*(I^*_{k_2,r}, f_{k_2}^*) + \tilde{a}_{k_1} \frac{\partial \tilde{a}_{k_2}}{\partial f} E^*(I^*_{k_2,r}) c_{k_1}^*(I^*_{k_1,r}, f_{k_2}^*) \right).$$

By using the same method to bound $I_1$ we can obtain similar bounds for $III_{1,1}$ and $III_{1,2}$, which altogether gives $III_1 = O_p(T^{-1})$. By using the Cauchy-Schwarz inequality we have $III_2 = O_p(\frac{1}{T^2} + \frac{1}{T^3})$, and similarly $III_3 = O_p(\frac{1}{T^2} + \frac{1}{T^3})$, thus we obtain (A.31).

Finally to bound (A.32) we use a similar decomposition to (A.37) to give

$$T \left( |E^* c^*(r, \ell)|^2 - |E^* \tilde{c}^*(r, \ell)|^2 \right) = \sum_{i=1}^3 IV_i,$$

where

$$IV_1 = \frac{2}{T} \sum_{k_1, k_2} a_{k_1} \frac{\partial a_{k_2}}{\partial f} \tilde{f}_{k_2} E^*(I^*_{k_1,r} I^*_{k_2,r})$$

$$IV_2 = \frac{1}{T} \sum_{k_1, k_2} a_{k_1} \frac{\partial a_{k_2}}{\partial f^2} \tilde{f}_{k_2}^2 E^*(I^*_{k_1,r} I^*_{k_2,r})$$

$$IV_3 = \frac{1}{T} \sum_{k_1, k_2} \tilde{f}_{k_1} \tilde{f}_{k_2} \frac{\partial a_{k_1}}{\partial f} \frac{\partial a_{k_2}}{\partial f} E^*(I^*_{k_1,r} I^*_{k_2,r}).$$

In order to bound the above we note that by decomposing $E^*(I^*_{k_1,r} I^*_{k_2,r})$ as the product of cumulants and using Theorem 4.1(ii) we have $\|E^*(I^*_{k_1,r} I^*_{k_2,r})\|_2 = O(I(k_1 = k_2) + \frac{1}{T^2})$. Therefore by using the the Cauchy-Schwarz inequality, Assumption 4.1 and Lemma A.13, equation (A.27) we obtain $IV_i = O_p(b + \frac{1}{T^2} + \frac{1}{T^3})$, and thus (A.32).

**PROOF of Lemma 4.2** We observe that

$$T \left| \text{cov}^*(\tilde{c}_{j_1,j_2}^*(r, \ell_1), \tilde{c}_{j_3,j_4}^*(r, \ell_2)) - \text{cov}^*(\tilde{c}_{j_1,j_2}^*(r, \ell_1), \tilde{c}_{j_3,j_4}^*(r, \ell_2)) \right|$$

$$\leq T \left( E^*(\tilde{c}_{j_1,j_2}^*(r, \ell_1) \tilde{c}_{j_3,j_4}^*(r, \ell_2)) - E^*(\tilde{c}_{j_1,j_2}^*(r, \ell) \tilde{c}_{j_3,j_4}^*(r, \ell_2)) \right)$$

$$+ T \left( E^*(\tilde{c}_{j_1,j_2}^*(r, \ell)) \tilde{E}^*(\tilde{c}_{j_3,j_4}^*(r, \ell_2)) - E^*(\tilde{c}_{j_1,j_2}^*(r, \ell)) \tilde{E}^*(\tilde{c}_{j_3,j_4}^*(r, \ell_2)) \right).$$

Substituting (A.29)-(A.32) into the above gives the bound $O_p(a(T, b, p))$. Using a similar method we can show

$$T \left| \text{cov}^*(\tilde{c}_{j_1,j_2}^*(r, \ell_1), \tilde{c}_{j_3,j_4}^*(r, \ell_2)) - \text{cov}^*(\tilde{c}_{j_1,j_2}^*(r, \ell_1), \tilde{c}_{j_3,j_4}^*(r, \ell_2)) \right| = O_p(a(T, b, p)).$$

Together, these two results give the bounds in Lemma 4.2.

In the above lemma we have shown that the bootstrap variance of $\tilde{c}_{j_1,j_2}^*(r, \ell)$ is asymptotically equivalent to the bootstrap variance of $\tilde{c}_{j_1,j_2}^*(r, \ell)$. Therefore in the following lemma we obtain an asymptotic expression for the bootstrap variance of $\tilde{c}_{j_1,j_2}^*(r, \ell)$ (which we have shown above is equivalent to the asymptotic variance of $\tilde{c}_{j_1,j_2}^*(r, \ell)$).
Lemma A.15 Let us suppose that Assumptions 2.1, where the mixing rate $\alpha$ and moments $r$ (where $\|X_t\|_r < \infty$) are such that $\alpha(1 - 8/r) > 7$. Then we have

$$T \text{cov}^*(\widetilde{c}_{j_1,j_2}(r, \ell_1), \widetilde{c}_{j_3,j_4}(r, \ell_2)) = \delta_{j_1,j_3} \delta_{j_2,j_4} \delta_{\ell_1,\ell_2} + \delta_{j_1,j_4} \delta_{j_2,j_3} \delta_{\ell_1,-\ell_2} + \kappa(\ell_1, \ell_2)(j_1, j_2, j_3, j_4) + O_p \left( \frac{1}{T p^6} \right)$$

$$T \text{cov}^*(\widetilde{c}_{j_1,j_2}(r, \ell_1), \widetilde{c}_{j_3,j_4}(r, \ell_2)) = \delta_{j_1,j_3} \delta_{j_2,j_4} \delta_{\ell_1,\ell_2} + \delta_{j_1,j_4} \delta_{j_2,j_3} \delta_{\ell_1,-\ell_2} + \kappa(\ell_1, \ell_2)(j_1, j_2, j_3, j_4) + O_p \left( \frac{1}{T p^6} \right)$$

**Proof.** 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 4.2 to obtain an expression for the covariance between the DFTs gives

$$T \text{cov}^*(\widetilde{c}_{j_1,j_2}(r, \ell_1), \widetilde{c}_{j_3,j_4}(r, \ell_2)) = \delta_{j_1,j_3} \delta_{j_2,j_4} \delta_{\ell_1,\ell_2} + \delta_{j_1,j_4} \delta_{j_2,j_3} \delta_{\ell_1,-\ell_2} + \kappa(\ell_1, \ell_2)(j_1, j_2, j_3, j_4) + O_p \left( \frac{1}{T p^6} \right)$$

**Proof of Theorem 4.2** The proof follows immediately from Lemma 4.2 and Lemma A.15.

**Proof of Theorem 4.2** By Slutsky’s lemma we have $\hat{S}^*(r) \xrightarrow{p} S_n$ and thus Theorem 4.2.
References


Table 1: Stationary case: Actual size of $T_{m,1,d}^*$ for several bandwidths $b$ and (approximated) expected block lengths $1/p$ and of $T_{m,1,dG}$ for several bandwidths $b$. The tests have been executed for samples of sizes $T = 50, 100, 200$ from Model I and based on lags $r = 1, \ldots, m$ for different values $m$. The results are reported for $\alpha = 1\%$ and $\alpha = 5\%$. 

<table>
<thead>
<tr>
<th>Model I</th>
<th>$n = 50$</th>
<th>$n = 100$</th>
<th>$n = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>$b$</td>
<td>$m \backslash \frac{1}{p}$</td>
<td>$G$</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>1</td>
<td>9.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>16.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>19.3</td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
<td>7.3</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>9.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>11.0</td>
</tr>
<tr>
<td>0.3</td>
<td>1</td>
<td>3.7</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>7.7</td>
</tr>
<tr>
<td>5.0</td>
<td>0.1</td>
<td>1</td>
<td>18.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>29.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>30.7</td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
<td>14.7</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>17.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>17.3</td>
</tr>
<tr>
<td>0.3</td>
<td>1</td>
<td>11.0</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>15.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>16.3</td>
</tr>
</tbody>
</table>
Figure 1: Stationary linear case: Bivariate realisations from stationary VAR(1) Models I and II and from stationary VMA(1) Models III and IV (from top to bottom) of sample size $T = 100$. First component (solid line) and second component (dashed line) are reported.
Figure 2: Stationary non-linear case: Bivariate realisations from stationary Models V and VI based on two independent GARCH(1,1) models (from top to bottom) of sample size $T = 100$. First component (solid line) and second component (dashed line) are reported.
Figure 3: Non-stationary case: Bivariate realisations from non-stationary tvVAR(1) Model VII, change point VAR(1) Model VIII and random walk Model IX (from top to bottom) of sample size $T = 100$. First component (solid line) and second component (dashed line) are reported.
Figure 4: Stationary case: Averages of squared entries of DFT covariances $T|\hat{C}_{11}(r,0)|^2$ (solid, black), $T|\hat{C}_{21}(r,0)|^2$ (dashes, black) and $T|\hat{C}_{22}(r,0)|^2$ (dotted, black) based on $M = 300$ time series from stationary Gaussian VAR(1) Model I (first row) and VMA(1) Model III (second row). The blue curves are the corresponding empirical 0.95-quantiles. The red curve is the 0.95-quantile of the $\chi^2$ distribution with two degrees of freedom. The results are reported for sample sizes $T = 50, 100, 200$ (left to right), where bandwidths $b = 0.3, 0.2, 0.1$ (left to right) have been used, respectively.
Figure 5: Non-stationary case: Averages of squared entries of DFT covariances $T|\hat{C}_{11}(r, 0)|^2$ (solid, black), $T|\hat{C}_{21}(r, 0)|^2$ (dashes, black) and $T|\hat{C}_{22}(r, 0)|^2$ (dotted, black) based on $M = 300$ time series from non-stationary Models VII, VIII and IX (top to bottom). The blue curves are the corresponding empirical 0.95-quantiles. The red curve is the 0.95-quantile of the $\chi^2$ distribution with two degrees of freedom. The results are reported for sample sizes $T = 50, 100, 200$ (left to right), where bandwidth $b = 0.3, 0.2, 0.1$ (left to right) have been used, respectively.
\[ T_{m,1,d} \]

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( b )</th>
<th>( m )</th>
<th>( \frac{1}{p} )</th>
<th>( n = 50 )</th>
<th>( n = 100 )</th>
<th>( n = 200 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>1</td>
<td>5.7</td>
<td>3.0</td>
<td>0.7</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>10.7</td>
<td>6.0</td>
<td>2.0</td>
<td>2.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>14.0</td>
<td>8.7</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
<td>1</td>
<td>4.7</td>
<td>0.7</td>
<td>0.0</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>7.7</td>
<td>3.7</td>
<td>0.7</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>11.3</td>
<td>4.7</td>
<td>1.7</td>
<td>5.7</td>
</tr>
<tr>
<td>0.3</td>
<td>1</td>
<td>1</td>
<td>4.7</td>
<td>0.0</td>
<td>0.0</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>7.7</td>
<td>1.3</td>
<td>0.0</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>8.3</td>
<td>2.7</td>
<td>0.3</td>
<td>6.7</td>
</tr>
<tr>
<td>5.0</td>
<td>0.1</td>
<td>1</td>
<td>13.0</td>
<td>8.3</td>
<td>3.0</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>20.7</td>
<td>10.3</td>
<td>5.7</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>24.0</td>
<td>13.3</td>
<td>6.7</td>
<td>5.3</td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
<td>1</td>
<td>10.7</td>
<td>4.0</td>
<td>1.3</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>18.0</td>
<td>9.0</td>
<td>4.3</td>
<td>7.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>20.3</td>
<td>7.7</td>
<td>5.7</td>
<td>11.3</td>
</tr>
<tr>
<td>0.3</td>
<td>1</td>
<td>8.7</td>
<td>2.3</td>
<td>0.7</td>
<td>6.0</td>
<td>11.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>11.0</td>
<td>4.3</td>
<td>1.7</td>
<td>9.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>12.3</td>
<td>4.3</td>
<td>2.3</td>
<td>10.0</td>
</tr>
</tbody>
</table>

Table 2: Stationary case: Actual size of \( T_{m,1,d}^* \) for several bandwidths \( b \) and (approximated) expected block lengths \( 1/p \) and of \( T_{m,1,d}$ for several bandwidths \( b \). The tests have been executed for samples of sizes \( T = 50, 100, 200 \) from Model II and based on lags \( r = 1, \ldots, m \) for different values \( m \). The results are reported for \( \alpha = 1\% \) and \( \alpha = 5\% \).
Table 3: Stationary case: Actual size of $T_{m, 1, d}^*$ for several bandwidths $b$ and (approximated) expected block lengths $1/p$ and of $T_{m, 1, d, G}$ for several bandwidths $b$. The tests have been executed for samples of sizes $T = 50, 100, 200$ from Model III and based on lags $r = 1, \ldots, m$ for different values $m$. The results are reported for $\alpha = 1\%$ and $\alpha = 5\%$. 

<table>
<thead>
<tr>
<th>Model III</th>
<th>$n = 50$</th>
<th>$n = 100$</th>
<th>$n = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$1$</td>
<td>$2$</td>
<td>$3$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$b$</td>
<td>$m\backslash \frac{1}{p}$</td>
<td>$G$</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>1</td>
<td>1.0  0.7  0.3  0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>3.0  0.3  0.3  0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>3.0  0.7  0.7  0.0</td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
<td>1</td>
<td>0.3  0.0  0.0  0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>2.0  0.0  0.0  0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>1.7  0.3  0.0  0.0</td>
</tr>
<tr>
<td>0.3</td>
<td>1</td>
<td>1</td>
<td>0.0  0.0  0.0  0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>0.0  0.0  0.0  0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>0.3  0.0  0.0  0.3</td>
</tr>
<tr>
<td>5.0</td>
<td>0.1</td>
<td>1</td>
<td>6.7  3.3  3.7  0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>9.0  3.0  2.7  0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>9.0  2.3  2.0  0.0</td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
<td>1</td>
<td>4.7  1.7  1.3  2.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>4.7  2.7  1.7  0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>6.3  2.0  2.0  0.3</td>
</tr>
<tr>
<td>0.3</td>
<td>1</td>
<td>1</td>
<td>1.3  2.3  0.3  0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>2.0  0.3  0.0  0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>4.3  0.3  0.3  0.3</td>
</tr>
</tbody>
</table>
Table 4: Stationary case: Actual size of $T_{m,1,d}^*$ for several bandwidths $b$ and (approximated) expected block lengths $1/p$ and of $T_{m,1,d,\mathcal{G}}$ for several bandwidths $b$. The tests have been executed for samples of sizes $T = 50, 100, 200$ from Model IV and based on lags $r = 1, \ldots, m$ for different values $m$. The results are reported for $\alpha = 1\%$ and $\alpha = 5\%$. 

<table>
<thead>
<tr>
<th>Model IV</th>
<th>$n = 50$</th>
<th>$n = 100$</th>
<th>$n = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>$b$</td>
<td>$m \ 1/p$</td>
<td>$G$</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
<td>0.3</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.7</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.7</td>
<td>0.3</td>
</tr>
<tr>
<td>0.3</td>
<td>1</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.3</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>5.0</td>
<td>0.1</td>
<td>1</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.0</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.3</td>
<td>5.3</td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
<td>3.7</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.7</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.3</td>
<td>1.7</td>
</tr>
<tr>
<td>0.3</td>
<td>1</td>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.7</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Table 5: Stationary case: Actual size of $\mathcal{T}^*_{m,1,d}$ for several bandwidths $b$ and (approximated) expected block lengths $1/p$ and of $\mathcal{T}_{m,1,d,G}$ for several bandwidths $b$. The tests have been executed for samples of sizes $T = 50, 100, 200$ from Model V and based on lags $r = 1, \ldots, m$ for different values $m$. The results are reported for $\alpha = 1\%$ and $\alpha = 5\%$.

Table 6: Stationary case: Actual size of $\mathcal{T}^*_{m,1,d}$ for several bandwidths $b$ and (approximated) expected block lengths $1/p$ and of $\mathcal{T}_{m,1,d,G}$ for several bandwidths $b$. The tests have been executed for samples of sizes $T = 50, 100, 200$ from Model VI and based on lags $r = 1, \ldots, m$ for different values $m$. The results are reported for $\alpha = 1\%$ and $\alpha = 5\%$. 

56
Table 7: Non-stationary case: Power of $T_{m,1,d}^*$ for several bandwidths $b$ and (approximated) expected block lengths $1/p$ and of $T_{m,1,d,G}$ for several bandwidths $b$. The tests have been executed for samples of sizes $T = 50, 100, 200$ from Model VII and based on lags $r = 1, \ldots, m$ for different values $m$. The results are reported for $\alpha = 1\%$ and $\alpha = 5\%$. 

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$b$</th>
<th>$m\frac{1}{p}$</th>
<th>$n = 50$</th>
<th>$n = 100$</th>
<th>$n = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 2 3  $G$</td>
<td>1 2 3  $G$</td>
<td>1 2 3  $G$</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>1</td>
<td>4.3 2.3 0.7 0.7</td>
<td>2.7 0.7 0.0 2.3</td>
<td>2.7 1.3 1.0 5.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>38.0 11.7 4.7 8.3</td>
<td>86.0 34.7 8.3 82.7</td>
<td>100 96.0 71.7 100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>36.0 9.7 4.0 5.7</td>
<td>71.0 21.3 5.3 65.0</td>
<td>99.7 77.0 35.0 100</td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
<td>4.3 0.3 0.3 1.0</td>
<td>1.3 0.7 0.3 2.0</td>
<td>2.7 0.3 1.3 3.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>35.0 4.7 1.3 12.3</td>
<td>84.3 29.3 8.3 93.0</td>
<td>100 88.0 56.3 100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>30.3 1.7 0.7 10.7</td>
<td>68.7 16.0 3.7 82.0</td>
<td>98.0 54.3 23.7 100</td>
</tr>
<tr>
<td>0.3</td>
<td>1</td>
<td>1.7 0.3 0.0 1.7</td>
<td>1.7 0.0 0.3 2.0</td>
<td>0.7 0.3 0.3 5.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>17.0 2.7 1.0 12.3</td>
<td>75.0 14.0 4.3 90.7</td>
<td>99.7 84.7 48.3 100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>11.0 1.7 0.7 9.3</td>
<td>51.0 7.3 2.3 73.7</td>
<td>97.7 46.3 15.3 100</td>
</tr>
<tr>
<td>5.0</td>
<td>0.1</td>
<td>14.0 7.0 3.3 3.0</td>
<td>9.0 4.0 2.0 7.0</td>
<td>7.0 3.0 2.3 11.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>63.0 30.7 14.3 21.0</td>
<td>95.3 68.7 28.7 94.7</td>
<td>100 99.3 94.7 100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>54.7 24.7 11.0 15.0</td>
<td>88.7 48.3 15.0 88.0</td>
<td>100 95.7 67.3 100</td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
<td>11.7 2.3 2.3 3.3</td>
<td>7.7 2.7 2.3 8.3</td>
<td>5.3 3.3 3.0 8.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>57.3 16.3 6.3 33.0</td>
<td>94.7 59.3 30.3 99.0</td>
<td>100 98.3 89.0 100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>48.7 11.3 3.7 28.7</td>
<td>85.7 35.3 14.3 93.0</td>
<td>99.7 85.7 56.0 100</td>
</tr>
<tr>
<td>0.3</td>
<td>1</td>
<td>6.7 3.7 0.3 5.0</td>
<td>7.7 1.7 0.3 7.3</td>
<td>3.3 3.0 1.3 13.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>41.0 13.7 5.3 30.3</td>
<td>93.3 45.3 21.0 98.3</td>
<td>99.7 98.3 87.0 100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>28.3 11.3 2.0 19.7</td>
<td>75.0 23.0 10.0 94.0</td>
<td>100 81.0 49.0 100</td>
</tr>
<tr>
<td>Model VIII</td>
<td>$n = 50$</td>
<td>$n = 100$</td>
<td>$n = 200$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>---------</td>
<td>---------</td>
<td>--------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$b$</td>
<td>$m\backslash\frac{1}{p}$</td>
<td>$G$</td>
<td>$1$</td>
<td>$2$</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>1</td>
<td>11.3</td>
<td>2.3</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>18.7</td>
<td>5.0</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>24.0</td>
<td>8.7</td>
<td>2.0</td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
<td>6.3</td>
<td>2.0</td>
<td>0.3</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>11.0</td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>17.3</td>
<td>1.0</td>
<td>0.3</td>
</tr>
<tr>
<td>0.3</td>
<td>1</td>
<td>5.3</td>
<td>1.3</td>
<td>0.0</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>9.0</td>
<td>1.7</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>11.7</td>
<td>2.0</td>
<td>0.3</td>
</tr>
<tr>
<td>5.0</td>
<td>0.1</td>
<td>1</td>
<td>26.7</td>
<td>10.3</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>37.7</td>
<td>15.0</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>39.3</td>
<td>19.3</td>
<td>7.3</td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
<td>17.0</td>
<td>11.7</td>
<td>2.7</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>24.3</td>
<td>10.3</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>31.7</td>
<td>9.7</td>
<td>3.7</td>
</tr>
<tr>
<td>0.3</td>
<td>1</td>
<td>21.7</td>
<td>7.7</td>
<td>1.7</td>
<td>5.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>21.0</td>
<td>5.0</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>22.7</td>
<td>4.0</td>
<td>1.3</td>
</tr>
</tbody>
</table>

Table 8: Non-stationary case: Power of $T_{m,1,d}^*$ for several bandwidths $b$ and (approximated) expected block lengths $1/p$ and of $T_{m,1,d,G}$ for several bandwidths $b$. The tests have been executed for samples of sizes $T = 50, 100, 200$ from Model VIII and based on lags $r = 1, \ldots, m$ for different values $m$. The results are reported for $\alpha = 1\%$ and $\alpha = 5\%$. 

58
| $\alpha$ | $b$ | $m \frac{1}{p}$ | $n = 50$ | | | | $n = 100$ | | | | $n = 200$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | $\mathcal{G}$ | 1 | 2 | 3 | $\mathcal{G}$ | 1 | 2 | 3 | $\mathcal{G}$ |
| 1.0 | 0.1 | 1 | 59.3 | 32.7 | 6.3 | 32.7 | 80.7 | 56.0 | 26.7 | 63.3 | 88.3 | 70.7 | 61.0 | 78.7 |
| | | | 3 | 72.3 | 52.0 | 26.7 | 50.3 | 86.0 | 72.0 | 51.3 | 72.3 | 91.7 | 78.0 | 71.7 | 82.7 |
| | | | 5 | 75.3 | 59.7 | 34.3 | 54.0 | 89.0 | 74.7 | 56.3 | 76.3 | 92.0 | 79.3 | 74.7 | 84.3 |
| | 0.2 | 1 | 57.7 | 13.3 | 1.0 | 30.3 | 77.3 | 41.0 | 21.7 | 63.3 | 83.0 | 67.3 | 51.0 | 82.7 |
| | | | 3 | 72.0 | 37.7 | 12.7 | 47.3 | 84.7 | 56.3 | 46.0 | 69.7 | 89.0 | 75.0 | 60.3 | 87.3 |
| | | | 5 | 77.3 | 47.3 | 20.3 | 55.7 | 85.7 | 59.7 | 50.0 | 73.3 | 90.0 | 77.0 | 63.7 | 88.7 |
| | 0.3 | 1 | 51.3 | 7.0 | 0.3 | 17.0 | 73.0 | 36.3 | 17.0 | 56.7 | 84.3 | 61.7 | 49.3 | 76.3 |
| | | | 3 | 65.0 | 21.3 | 3.7 | 35.3 | 80.3 | 53.3 | 34.7 | 64.3 | 90.3 | 71.3 | 59.0 | 84.0 |
| | | | 5 | 69.0 | 33.7 | 1.3 | 40.3 | 83.0 | 57.7 | 43.3 | 66.7 | 90.7 | 72.0 | 62.0 | 84.0 |
| | 5.0 | 0.1 | 1 | 69.3 | 52.3 | 25.0 | 45.0 | 86.0 | 70.7 | 46.3 | 74.3 | 89.3 | 76.3 | 69.7 | 82.7 |
| | | | 3 | 77.7 | 64.0 | 40.3 | 59.0 | 90.7 | 76.3 | 60.0 | 79.3 | 93.3 | 80.0 | 76.0 | 85.3 |
| | | | 5 | 81.7 | 66.3 | 50.3 | 61.7 | 93.7 | 79.3 | 65.0 | 80.3 | 93.7 | 82.3 | 78.7 | 86.7 |
| | 0.2 | 1 | 72.0 | 33.0 | 9.3 | 48.0 | 84.0 | 53.3 | 41.7 | 71.7 | 86.7 | 73.3 | 62.3 | 87.0 |
| | | | 3 | 79.7 | 50.3 | 25.3 | 57.7 | 87.3 | 63.7 | 53.3 | 76.3 | 92.3 | 78.3 | 66.7 | 90.3 |
| | | | 5 | 83.3 | 55.7 | 33.0 | 63.3 | 89.0 | 66.7 | 56.0 | 77.7 | 92.3 | 79.7 | 70.7 | 90.3 |
| | 0.3 | 1 | 63.0 | 26.7 | 4.7 | 38.3 | 81.7 | 54.0 | 38.3 | 68.3 | 91.3 | 71.0 | 62.3 | 83.0 |
| | | | 3 | 73.0 | 39.7 | 17.7 | 49.7 | 86.3 | 65.3 | 50.7 | 72.3 | 93.0 | 76.0 | 66.0 | 88.0 |
| | | | 5 | 74.3 | 48.0 | 25.7 | 52.3 | 87.7 | 67.0 | 53.3 | 74.3 | 94.0 | 77.0 | 68.7 | 87.0 |

Table 9: Non-stationary case: Power of $T^*_m,1,d$ for several bandwidths $b$ and (approximated) expected block lengths $1/p$ and of $T_{m,1,d,G}$ for several bandwidths $b$. The tests have been executed for samples of sizes $T = 50, 100, 200$ from Model IX and based on lags $r = 1, \ldots, m$ for different values $m$. The results are reported for $\alpha = 1\%$ and $\alpha = 5\%$. 

59