FREQUENCY DOMAIN TESTS OF SEMIPARAMETRIC HYPOTHESES FOR LOCALLY STATIONARY PROCESSES

MARIOS SERGIDES AND EFSTATHIOS PAPARODITIS

Abstract

Many time series in applied sciences obey a time-varying spectral structure. In this article, we focus on locally stationary processes and develop tests of the hypothesis that the time-varying spectral density has a semiparametric structure, including the interesting case of a time-varying autoregressive moving-average model. The test introduced is based on a L_2-distance measure of a kernel smoothed version of the local periodogram rescaled by the time-varying spectral density of the estimated semiparametric model. The asymptotic distribution of the test statistic under the null hypothesis is derived. As an interesting special case, we focus on the problem of testing for the presence of a time-varying autoregressive model. A semiparametric bootstrap procedure to approximate more accurately the distribution of the test statistic under the null hypothesis is proposed. Some simulations illustrate the behavior of our testing methodology in finite sample situations.

Date: January 19, 2009.

2000 *Mathematics Subject Classification.* Primary 62M10, 62M15; secondary 62G09.

Key words and phrases. Bootstrap, Local periodogram, Local spectral density, Nonparametric estimation, Semiparametric models.

1. Introduction

Most existing models in time series analysis assume that the underlying process is second-order stationary. Although this assumption is attractive from a theoretical point of view because it allows for the development of statistical inference procedures with good asymptotic properties, it seems rather restrictive in applications. A more realistic framework in time series analysis is one which allows for the dependence structure of the underlying stochastic process and more specifically for its second order properties, to vary smoothly over time. Developing a useful approach of statistical inference in such a context requires, however, that some restrictions have to be imposed on the deviations from stationarity which are allowed.

There is meanwhile a large body on statistical literature dealing with different aspects of the analysis of stochastic processes that obey time-varying spectral characteristics. One of the first attempts was Priestley (1965) who considered stochastic processes with a time-varying spectral representation similar to that of stationary process; see also Priestley (1981). Statistical inference problems for time-varying stochastic processes has attracted considerable interest during the last decades. To mention only few of the different approaches proposed in this context we refer to Dahlhaus (1997) on locally stationary processes, to Nason et al.(2000) and Ombao et al. (2005) on wavelet processes and to Davis et al. (2006) on piecewise stationary processes. One way to investigate properties of statistical inference procedures for time-varying stochastic processes, is to allow for the amount of local information available to increase to infinity as the sample size increases. Such a nonparametric-type framework for the development of an asymptotic theory of statistical inference for nonstationary processes has been developed by Dahlhaus (1997) who introduced the concept of locally stationary processes.

Locally stationary processes are stochastic processes that obey a time-varying spectral representation which generalizes the Cramér representation of stationary processes by assuming a time-varying amplitude function. Interesting subclasses of locally stationary processes are obtained by parameterizing in a proper way the associated time-varying amplitude function and consequently the underlying time-varying spectral density. Such an interesting subclass of locally processes is for instance, that of time-varying, autoregressive moving-average (tvARMA) models. tvARMA models are autoregressive moving-average model the parameters of which vary smoothly over time. The concept of local stationarity has been extended respectively modified in several directions. Nason et al. (2000) replaced the spectral representation and the Fourier basis associated with a locally stationary process by a representation with respect to a wavelets basis; see also Ombao et al. (2002) and Ombao et al. (2005). The corresponding model of a locally stationary wavelet process allows for a time-scale representation of a stochastic process.

Estimation procedures for locally stationary processes have been considered by many authors under different settings and assumptions. We mention here among others the contributions by Neumann and von Sachs (1997), Dahlhaus et al. (1999), Chang and Morettin (1999), van Bellegem and Dahlhaus (2006), Dahlhaus and Polonik (2006) and van Bellegem and von Sachs (2008). Forecasting problems for non-stationary time series have been considered by Fryzlewicz et al. (2003). An overview on some of the different developments can be found in Dahlhaus (2003). However, the important problem of testing for the presence of a parametric or semiparametric structure of the underlying locally stationary process, has attracted less attention in the literature. Testing for the presence of such a structure is important because it allows for the use of efficient, i.e., model-based estimation and forecasting procedures. For Gaussian locally stationary processes, Sakiyama and Taniguchi (2003) proposed likelihood ratio, Wald and Lagrange multiplier tests of the null hypothesis that the time-varying spectral density depends on a finite dimensional, real-valued parameter vector against a real-valued parametric alternative. However, the class of parametric time-varying spectral densities allowed in this context, is rather restrictive in that it does not include for instance the important case of testing for the presence of a semiparametric tvARMA structure against an unspecified, locally stationary alternative.

In this paper, we address the important problem of testing whether a locally stationary process belongs to a semiparametric class of time-varying processes. The semiparametric class considered under the null is large enough to include several interesting processes. The test statistic developed, evaluates over all frequencies and over an increasing set of time points, a L_2-type distance between the sample local spectral density (local periodogram) and the time-varying spectral density of the fitted semiparametric model postulated under the null. The asymptotic distribution of the test statistic proposed under the null hypothesis is derived and it is shown that this distribution is a Gaussian distribution with the nice feature that its parameters do not depend on characteristics of the underlying process. As an interesting special case we focus on the problem of testing for the presence of a semiparametric, time-varying autoregressive model. In this context, a bootstrap procedure is proposed to approximate more accurately the distribution of the test statistic under the null hypothesis. Theoretical properties of the bootstrap procedure are investigated and its asymptotic validity is established. It is demonstrated by means of numerical examples that in the testing set-up considered in this paper, the bootstrap is a very powerful and valuable tool to obtain critical values in finite sample situations.

The paper is organized as follows. In Section 2 we describe in detail our testing procedure and derive the asymptotic distribution of the test statistic proposed. In Section 3 we focus on the problem of testing for the presence of a time-varying autoregressive model and introduce the bootstrap procedure used to approximate the distribution of the test statistic under the null. Section 4 contains some simulations investigating the performance of the bootstrap and the size and power behavior of the test in finite sample situations. Section 5 concludes our findings while the proofs of the main theorems and of some auxiliary lemmas are deferred to the Appendix.

2. The Testing procedure

2.1. **The set-up.** Following Dahlhaus (1997) we consider triangular arrays $\{\mathbf{X}_T\}_{T\in\mathbb{N}}$ $\mathbf{X}_T = \{X_{t,T}, t = 1, \ldots, T\}$ of stochastic processes satisfying the following conditions.

Assumption 2.1. For all $T \in \mathbb{N}$, $X_{t,T}$ has the representation

$$X_{t,T} = \int_{-\pi}^{\pi} A^0_{t,T}(\lambda) e^{i\lambda t} d\xi(\lambda), \tag{1}$$

where

(i) $\xi(\lambda)$ is a Gaussian stochastic process on $[-\pi, \pi]$ with $\overline{\xi(\lambda)} = \xi(-\lambda)$, $E\{\xi(\lambda)\} = 0$ and

$$E\left\{d\xi(\lambda_1)\overline{d\xi(\lambda_2)}\right\} = \eta(\lambda_1 + \lambda_2)\, d\lambda_1 d\lambda_2, \tag{2}$$

where $\eta(\lambda) = \sum_{j=-\infty}^{\infty} \delta(\lambda - 2\pi j)$ is the period 2π extension of Dirac's Delta function.

(ii) There exists a constant K and a Lipshitz continuous function $A(u, \lambda)$ on $[0, 1] \times (-\pi, \pi]$ which is 2π periodic in λ, with $\overline{A(u, \lambda)} = A(u, -\lambda)$, such that for all T,

$$\sup_{t,\lambda} |A^0_{t,T}(\lambda) - A(t/T, \lambda)| \le K/T. \tag{3}$$

The uniquely defined function

$$f(u, \lambda) = \frac{1}{2\pi} |A(u, \lambda)|^2, \qquad (u, \lambda) \in [0, 1] \times (-\pi, \pi], \tag{4}$$

is called the time-varying spectral density of $\{\mathbf{X}_T\}_{T\in\mathbb{N}}$; see Dahlhaus (1996a),

The aim of this paper is to develop tests of the hypothesis that the time-varying local spectral density $f(u, \lambda)$ has a semiparametric structure. To elaborate on the kind of null and alternative hypothesis considered, let $\mathcal{F}_{\mathcal{LS}}$ be the set of local spectral densities of processes satisfying Assumption 2.1 and denote by $\mathcal{F}_{\mathcal{PLS}} \subseteq \mathcal{F}_{\mathcal{LS}}$ a semiparametric model class of local spectral densities, i.e.,

$$\mathcal{F}_{\mathcal{PLS}} = \{f(u,\lambda) = f(u,\lambda;\vartheta(u)), \quad \vartheta(u) = (\vartheta_1(u),\ldots,\vartheta_m(u)), m \in \mathbb{N}, u \in [0,1], \lambda \in \mathbb{R}\},$$

where $\vartheta_i(\cdot) : [0,1] \to \mathbb{R}$, $i = 1,2,\ldots,m$, are appropriately defined real-valued functions. We assume that in the set $\mathcal{F}_{\mathcal{PLS}}$, the time-varying local spectral density $f(u,\lambda,\vartheta(u))$ is fully determined by the unknown functions $\vartheta_i(\cdot)$, $i = 1,2,\ldots,m$, and as we will see in the sequel, we impose some rather mild assumptions on $\vartheta(\cdot)$ allowing for several interesting classes of semiparametric models.

To give one important example which fits in the above set-up, consider the case where $\mathcal{F}_{\mathcal{PLS}}$ is the semiparametric class of local spectral densities possessed by the class of time-varying autoregressive moving-average (tvARMA) models. Recall that a locally stationary process $\{X_{t,T}\}$ satisfying Assumption 2.1 has a tvARMA(p,q) representation if $X_{t,T}$ is generated by the equation

$$X_{t,T} + \sum_{j=1}^{p} a_j(t/T)X_{t-j,T} = \varepsilon_{t,T} + \sum_{j=1}^{q} b_j(t/T)\varepsilon_{t-j,T} \tag{5}$$

where $a_0(\cdot) \equiv b_0(\cdot) \equiv 1$, the ε_t's are i.i.d. $N(0,\sigma^2(t/T))$ distributed random variables, $a_p(u) \neq 0$ and $b_q(u) \neq 0$ a.e. in $[0,1]$. Notice that the above class of time varying ARMA models are not always locally stationary; cf. Dahlhaus (1996) for a discussion. Now, if the functions $\alpha_j(\cdot)$, $\beta_j(\cdot)$ and $\sigma^2(\cdot)$ are continuous on $\mathbb{R}$ and $\sum_{j=0}^{p}\alpha_j(u)z^j \neq 0$ for all $|z| \leq 1+\delta$ for some $\delta > 0$ uniformly in $u \in [0,1]$, then model (5) belongs to the locally stationary process class described in Assumption 2.1; see Dahlhaus(1996), Theorem 2.3. Notice that model (5) possesses a time-varying spectral density given by

$$f(u,\lambda;\vartheta(u)) = \frac{\sigma^2(u)}{2\pi}\left|\sum_{j=0}^{q} b_j(u)e^{i\lambda j}\right|^2 \Big/ \left|\sum_{j=0}^{p} a_j(u)e^{i\lambda j}\right|^2,$$

where $\vartheta(u) = (a_1(u),\ldots,a_p(u),b_1(u),\ldots,b_q(u),\sigma^2(u))$.

Based on the above discussion, the testing problem considered in this paper is described by

$$H_0 : f(\cdot,\cdot) \in \mathcal{F}_{\mathcal{PLS}} \quad \text{vs} \quad H_1 : f(\cdot,\cdot) \in \mathcal{F}_{\mathcal{LS}} \setminus \mathcal{F}_{\mathcal{PLS}}. \tag{6}$$

The specific case where $\vartheta(u)$ is a constant function of the time variable u, that is where $\vartheta(u) = (\vartheta_1,\ldots,\vartheta_m) \in \Theta \subset \mathbb{R}^m$ for all $u \in (0,1)$, is also allowed by (6). Such a case occurs for instance if one is interested in testing the null hypothesis that the underlying stochastic process is a parametric stationary process against the alternative of a time-varying locally stationary process.

2.2. **The Test statistic.** We start our construction of the test statistic by first considering the local periodogram defined for $N < T$, $N \in \mathbb{N}$, by

$$I_N(u,\lambda) = \frac{1}{2\pi H_{2,N}(0)} |d_N(u,\lambda)|^2, \tag{7}$$

where

$$d_N(u,\lambda) = \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right) X_{[uT]-N/2+s+1} e^{-i\lambda s},$$

$h : [0,1] \to \mathbb{R}$ is a taper function and

$$H_{k,N}(\lambda) = \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right) e^{-i\lambda s}.$$

Recall that the local periodogram is the periodogram of a segment of length N of consecutive observations around the time point $[uT]$, $u \in (0,1)$. $I_N(u,\lambda)$ is commonly computed at the Fourier frequencies $\lambda_j = 2\pi j/N$, $\;j = -[(N-1)/2], \ldots, [N/2]$.

To introduce, the basic statistic used, suppose first for simplicity that the parametric curves $\vartheta(u)$ determining the local spectral density $f(u,\lambda;\vartheta(u))$ under the null hypothesis are known, that is that $\vartheta(u) = \vartheta_0(u)$. Consider then the random variables

$$Y(u,\lambda_j) = \frac{I_N(u,\lambda_j)}{f(u,\lambda_j;\vartheta_0(u))}, \qquad j = -[(N-1)/2], \ldots, [N/2].$$

It is easy to see that if the null hypothesis is true, then

$$E[Y(u,\lambda_j)] \;=\; 1 \;+\; O(N/T + 1/N),$$

for all $u \in [0,1]$ and $\lambda_j \in (-\pi,\pi]$. Furthermore, if the alternative hypothesis is true, i.e., if $f(u,\lambda_j) \neq f(u,\lambda_j;\vartheta_0(u))$, then

$$E[Y(u,\lambda_j)] = \frac{f(u,\lambda_j)}{f(u,\lambda_j;\vartheta_0(u))} + O(N/T + 1/N),$$

where the function $f(\cdot,\cdot)/f(\cdot,\cdot;\vartheta_0(\cdot))$ is different from the unit function on $[0,1]\times(-\pi,\pi]$.

Motivated by the above observations the idea used to obtain a test statistic for the null hypothesis that $f(u,\lambda) = f(u,\lambda,\vartheta_0(u))$, is to estimate non-parametrically the mean function

$$q(u,\lambda) = E[Y(u,\lambda) - 1]$$

and then to evaluate its distance from the zero function using an appropriate L_2-distance measure. To elaborate on, for given $u \in (0,1)$ and $\lambda \in [0,\pi]$, we use the kernel estimator

$$\hat{q}(u,\lambda) = \frac{1}{N}\sum_{j} K_b(\lambda-\lambda_j)\left(\frac{I_N(u,\lambda_j)}{f(u,\lambda_j;\vartheta_0(u))}-1\right) \tag{8}$$

to estimate non-parametrically the unknown mean function $q(u,\lambda)$. Here $K_b(\cdot) = b^{-1}K(\cdot/b)$ where $K(\cdot)$ is an appropriate defined kernel and b a smoothing bandwidth satisfying certain conditions; see Assumption 2.2 below.

To proceed with the construction of the test statistic proposed, we calculate $\hat{q}(u_j,\lambda)$ for different instants of time u_j by using the local periodogram $I_N(u_j,\lambda)$ for segments of observations having midpoints $u_j = t_j/T$. Here we choose $t_j := S(j-1)+N/2$ for $j = 1,\ldots,M$, where the constant S denotes the shift from segment to segment and M refers to the total number of time points in the interval (0,1) considered. Note that by the above construction we have $T = S(M-1)+N$. Now, using a L_2-measure to evaluate the distance of the so estimated mean function $\hat{q}(u_j,\lambda)$ from the zero function and averaging over all time points $u_j = t_j/T$ and integrating over all frequencies λ considered, we end-up with the test statistic

$$Q_{0,T} = \frac{1}{M}\sum_{s=1}^{M}\int_{-\pi}^{\pi}\Big(\hat{q}(u_s,\lambda)\Big)^2 d\lambda. \tag{9}$$

It can be shown that under some rather standard assumptions to be discussed later and if $M \to \infty$ as $T \to \infty$, then, in probability,

$$Q_{0,T} \;\to\; \begin{cases} 0 & \text{if } H_0 \text{ is true} \\ \int_0^1\int_{-\pi}^{\pi}\left(\frac{f(u,\lambda)}{f(u,\lambda,\vartheta_0)}-1\right)^2 d\lambda du & \text{if } H_1 \text{ is true.}\end{cases}$$

This behavior of $Q_{0,T}$ justifies its use for testing the null hypothesis of interest.

Recall that in order to derive the test statistic (9) we have assumed that the parameterizing functions $\vartheta(u)$ are known. This corresponds to the case of testing a simple hypothesis, that is a hypothesis where the local spectral density under the null is fully specified. To extend the testing procedure proposed to the more interesting case of testing a composite hypotheses, that is to the case where the functions $\vartheta(u)$ determining the local spectral density are unknown, we replace $\vartheta(\cdot)$ in (8) by $\sqrt{N}$-consistent estimators. Let $\hat{\vartheta}(\cdot) = (\hat{\vartheta}_1(\cdot),\ldots,\hat{\vartheta}_m(\cdot))'$ be such an estimator of $\vartheta(\cdot) = (\vartheta_1(\cdot),\ldots,\vartheta_m(\cdot))'$; see among others, Dahlhaus and Giraitis (1998), Dahlhaus (2000), Dahlhaus and Neumann (2001) and van Bellegem and Dahlhaus (2006) for different proposals for estimating $\vartheta(\cdot)$.

Analogously to (9), the test statistic used in this case is given by

$$Q_T = \frac{1}{M}\sum_{s=1}^{M}\int_{-\pi}^{\pi}\left\{\frac{1}{N}\sum_{j=-M_N}^{M_N} K_b(\lambda-\lambda_j)\left(\frac{I_N(u_s,\lambda_j)}{f(u_s,\lambda_j;\hat{\vartheta}(u_s))}-1\right)\right\}^2 d\lambda. \qquad (10)$$

Notice that $f(u_s,\lambda_j;\hat{\vartheta}(u_s))$ appearing in the denominator above, is the semiparametric local spectral density obtained by substituting $\vartheta(\cdot)$ appearing in $f(u,\lambda_j;\vartheta(u))$ by its estimator $\hat{\vartheta}(\cdot)$.

2.3. **Asymptotic distribution under the null hypothesis.** We first establish a basic theorem which deals with the asymptotic distribution of the test statistic (10) under the null hypothesis in (6). For this the following set of assumptions is imposed.

Assumption 2.2.

(i) K is a bounded, symmetric, nonnegative kernel function on $(-\infty,\infty)$ with support $[-\pi,\pi]$ such that $(2\pi)^{-1}\int_{-\infty}^{\infty}K(x)dx = 1$.

(ii) The window length N satisfies $N \sim T^{\delta}$ for some $1/5 < \delta < 4/5$. Furthermore, $S = [N/\kappa]$ where κ is a fixed, small positive integer, $1 \leq \kappa < N$, which is independent of N and S.

(iii) The smoothing bandwidth b satisfies $b \sim N^{-\lambda}$, where

$$\max\{0, \frac{9\delta-7}{\delta}\} < \lambda < \min\{\frac{5\delta-1}{3\delta}, \frac{1}{2}, \frac{1-\delta}{\delta}\}.$$

(iv) The taper function h is of bounded variation and vanishes outside the interval [0,1].

(v) $\sqrt{N}(\hat{\theta}(u)-\theta(u)) = O_p(1)$ where the $O_p(\cdot)$ term does not depend on u.

Some remarks concerning the above assumptions are in order. Note that the constant κ appearing in (ii) determines the degree of overlapping between the segments used. We consider the case $\kappa \geq 1$ only, since for $\kappa < 1$ the shift from segment to segment described by S is greater than the segment length N. In the later case, a loss of efficiency is expected due to the fact that some observations are omitted. If $\kappa = 1$ then the observed series is partitioned in nonoverlapping segments of length N while if $\kappa > 1$ then the segments considered overlap. Notice that for N and S known, the total number of time points M considered is given by $M = 1 + (T-N)/S$. Concerning the rate at which the segment length N is allowed to increase to infinity given in (ii) and the rate at which the bandwidth b is allowed to converge to zero given in (iii), we mention that they are controlled in a way that leads to simple expressions for the mean and for the variance of the limiting distribution of Q_T under H_0. Notice that the range of values of N and of b is large enough allowing for a flexibility in choosing these parameters in practice. Figure 1 describes the allowed range of λ according to the values of δ. Assumption 2.2(v) is general enough and allows for different estimators of $\vartheta(u)$.

Please insert Figure 1 about here.

The following theorem establishes the asymptotic distribution of Q_T when the null hypothesis is true.

Theorem 2.1. *Under Assumption 2.1 and 2.2 and if H_0 is true, then, as $T \to \infty$,*

$$N\sqrt{Mb}(Q_T - \mu_T) \Rightarrow N(0, \tau^2),$$

where

$$\mu_T = \frac{(tap(1))^{1/2}}{Nb} \int_{-\pi}^{\pi} K^2(x)dx + \frac{(tap(1))^{1/2}}{4\pi N} \int_{-\pi}^{\pi} \int_{-2\pi}^{2\pi} K(x)K(x-u)dxdu,$$

$$\tau^2 = tap(\kappa)\frac{2}{\pi} \int_{-2\pi}^{2\pi} \left(\int K(u)K(u+x)du \right)^2 dx$$

and for $s \in \{1, 2, \ldots, m\}$

$$tap(s) = \frac{\sum_{|m|<s} \left(\int_0^{1-|m|/s} h^2(u)h^2(u+|m|/s)du \right)^2}{\left(\int_0^1 h^2(x)dx \right)^4}.$$

According to the above theorem, an attractive feature of the test statistic Q_T, is that its limiting distribution under the null hypothesis does not depend on unknown parameters or characteristics of the underlying locally stationary process $\{X_{t,T}\}$. Furthermore, and based on this theorem, an asymptotically α-level test is obtained by rejecting the null hypothesis if

$$Q_T \geq \mu_T + \frac{\tau}{N\sqrt{Mb}} z_\alpha,$$

where z_α denotes the $100(1-\alpha)\%$ percentile of the standard Gaussian distribution.

3. Testing for a time-varying autoregressive structure

3.1. **Preliminaries.** A special case of the testing problem (6) and which commonly arises in many situations, is that of testing for the presence of a time-varying autoregressive (tvAR) model. Recall that the locally stationary process (1) obeys a time-varying autoregressive representation of order p if $X_{t,T}$ is generated by the equation

$$X_{t,T} = \sum_{j=1}^{p} \beta_j(t/T) X_{t-j,T} + \varepsilon_{t,T}, \tag{11}$$

where the $\varepsilon_{t,T}$'s are i.i.d., $N(0, \sigma^2(t/T))$ random variables, $\beta_p(u) \neq 0$ for all $u \in [0,1]$, the functions $\beta_j(\cdot)$ as well as the variance function $\sigma^2(\cdot)$ are of bounded variation and $\sum_{j=1}^{p} \beta_j(u) z^j \neq 0$ for all $u \in [0,1]$ and all $0 < |z| \leq 1 + \delta$, for some $\delta > 0$. Although the results of this paper can be easily adapted to cover other special types of semiparametric locally stationary processes i.e. tvARMA(p,q) or tvMA(q), we concentrate on the class of time-varying autoregressive process because these processes provide, due to their simplicity, easy implementation and interpretation, a very interesting subclass of semiparametric time-varying processes. Now let $\mathcal{F}_{tvAR(p)}$ be the set of local spectral densities of time-varying autoregressive processes of order p. The testing problem considered in this section is then described by the following pair of null and alternative hypothesis

$$H_0: \; f(\cdot,\cdot) \in \mathcal{F}_{tvAR(p)} \quad \text{vs} \quad H_1: \; f(\cdot,\cdot) \in \mathcal{F}_{\mathcal{LS}} \setminus \mathcal{F}_{tvAR(p)}. \tag{12}$$

Note that the set $\mathcal{F}_{\mathcal{LS}} \setminus \mathcal{F}_{tvAR(p)}$ contains also all locally stationary autoregressive processes with an autoregressive order different from p.

3.2. **Consistency.** We first discuss a consistency property of our test. For this, suppose that the true spectral density $f(u,\lambda)$ lies in the alternative and measure for $u \in [0,1]$ the distance between $f(u,\lambda)$ and $f(u,\lambda;\vartheta(u))$ by the function

$$\mathcal{L}(u,\vartheta(u)) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\log[f(u,\lambda;\vartheta(u))] + \frac{f(u,\lambda)}{f(u,\lambda;\vartheta(u))} \right) d\lambda. \tag{13}$$

Let $\overline{\vartheta}(u)$ be the value of $\vartheta(u)$ which minimizes $\mathcal{L}(u,\vartheta(u))$ and let $\hat{\vartheta}(u)$ be the estimator of $\overline{\vartheta}(u)$ which is obtained by minimizing the local Whittle likelihood, i.e., $\overline{\vartheta}(u) = \arg\min \mathcal{L}_N(u,\vartheta(u))$, where

$$\mathcal{L}_N\{u,\vartheta(u)\} = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\log[f(u,\lambda;\vartheta(u))] + \frac{I_N(u,\lambda)}{f(u,\lambda;\vartheta(u))} \right) d\lambda.$$

Notice that

$$\frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \left(\frac{\log[f(u,\lambda;\vartheta(u))]}{f(u,\lambda)} + \frac{f(u,\lambda)}{f(u,\lambda;\vartheta(u))} - 1 \right) d\lambda du$$

is the asymptotic Kullback-Leibler information divergence between two Gaussian locally stationary processes with time-varying spectral densities $f(u,\lambda;\vartheta(u))$ and $f(u,\lambda)$ respectively; see Theorem 3.4 of Dahlhaus (1996). The curve $\overline{\vartheta}(u)$ obtained by minimizing

(13) is that leading to the best time-varying autoregressive fit, that is to the p-th order autoregressive fit which minimizes the Kullback-Leibler information divergence (13).

Assumption 3.1. Let $\nabla = (\partial/\partial\vartheta_1, \ldots, \partial/\partial\vartheta_m)'$ be the gradient with respect to ϑ.

(i) $\nabla\mathcal{L}_N(u, \hat{\vartheta}(u)) = 0$, $\nabla\mathcal{L}(u, \overline{\vartheta}(u)) = 0$ for all u and N.

(ii) The derivatives $\partial^2 A(u,\lambda)/\partial u\partial\lambda$ and $\partial^3 A(u,\lambda)/\partial u^3$ are uniformly bounded in $(u,\lambda) \in [0,1] \times [-\pi,\pi]$.

(iii) The derivatives

$$\frac{\partial^3}{\partial\vartheta_{i_1}\partial\vartheta_{i_2}\partial\vartheta_{i_3}} f^{-1}(u,\lambda;\vartheta(u)), \quad \frac{\partial^3}{\partial\vartheta_{i_1}\partial\vartheta_{i_2}\partial\vartheta_{i_3}} f(u,\lambda;\vartheta(u)), \quad \frac{\partial^2}{\partial\lambda^2}\frac{\partial}{\partial\vartheta_{i_1}} f^{-1}(u,\lambda;\vartheta(u))$$

are bounded for $1 \le i_1, i_2, i_3 \le p$ uniformly in $(u,\lambda,\vartheta) \in [0,1] \times [-\pi,\pi] \times \Theta$, where Θ is an open convex subset of $\mathbb{R}^p$.

(iv) $\sup_{0\le u\le 1, \vartheta\in\Theta} ||\nabla^2\mathcal{L}^{-1}(u,\hat{\vartheta}(u))||_{sp}$ where $||\cdot||_{sp}$ denotes the spectral norm of a matrix.

We first state the following result which deals with the limiting properties of Q_T when the alternative hypothesis is true.

Theorem 3.1. *Under Assumptions 2.1, 2.2 and 3.1 and if* $f(\cdot,\cdot) \in \mathcal{F}_{\mathcal{LS}} \setminus \mathcal{F}_{tvAR(p)}$*, then as* $T \to \infty$,

$$Q_T \to D^2 = \int_0^1 \int_{-\pi}^{\pi} \left(\frac{f(u,\lambda)}{f(u,\lambda;\overline{\vartheta}(u))} - 1 \right)^2 d\lambda du,$$

in probability.

Notice that the limit D^2 given above is a L_2-distance measure between the true local spectral density $f(u,\lambda)$ and its best parametric fit $f(u,\lambda;\overline{\vartheta}(u))$. Theorem 3.1 implies then that under the assumptions made and if H_1 is true, then $\lim_{T\to\infty} P(N\sqrt{Mb}(Q_T - \mu_T)/\tau \ge z_\alpha) = 1$, that is the test Q_T is consistent against any alternative for which $D^2 > 0$.

3.3. **Bootstrapping the test statistic.** To obtain critical values of the test, Theorem 2.1 enables us to approximate the unknown distribution of $N\sqrt{Mb}(Q_T - \mu_T)/\tau$ by that of a standard Gaussian distribution. We experienced, however, that the quality of this approximation is rather poor in finite sample situations and very large to huge sample sizes are required in order for this approximation to be valuable in practice; see Section 4 for a numerical illustration of this point. To improve upon the large sample Gaussian approximation of Theorem 2.1, we propose here, an alternative, bootstrap-based procedure, which leads in finite sample situations to more accurate estimates of the distribution of Q_T under the null. The procedure proposed works by generating pseudo-observations $X^+_{1,T}, X^+_{2,T}, \ldots, X^+_{T,T}$ using the fitted tvAR(p) process and calculating the test statistic Q_T of interest using the so generated pseudo-observations.

To elaborate on, we first fit locally to the time series the pth order time-varying autoregressive process postulated under the null hypothesis. Denote by $\hat{\beta}_u(p)' = (\hat{\beta}_1(u), \ldots, \hat{\beta}_p(u))$ and by $\hat{\sigma}^2_N(u)$ the estimators of the time varying autoregressive parameters and of the error variance respectively. Such an estimator can be for instance obtained using for $[uT] \in \{N/2, N/2+1, \ldots, T-N/2\}$ the local Yule-Walker equations

$$\hat{R}_u(p)\hat{\beta}_u(p) = \hat{r}_u(p),$$

where

$$\hat{R}_u(p) = \hat{c}_N(u, i-j)_{i,j=1,\ldots,p}, \qquad \hat{r}_u(p) = (\hat{c}_N(u,1), \ldots, \hat{c}_N(u,p))'$$

and

$$\hat{c}_N(u,j) = \frac{1}{N}\sum_{\substack{k,l=0\\ k-l=j}}^{N-1} X_{[uT]-N/2+k+1,T}X_{[uT]-N/2+l+1,T}.$$

The corresponding estimator of the variance function $\sigma^2(u)$ of the errors is given by

$$\hat{\sigma}^2_N(u) = \hat{c}_N(u,0) + \hat{\beta}'_u(p)\hat{r}_u(p).$$

Properties of local Yule-Walker estimators have been investigated by Dahlhaus and Giraitis (1998); see also Sergides and Paparoditis (2008). Other estimators of the local autoregressive parameters have been investigated by, e.g., Dahlhaus et al. (1999) and van Bellegem and Dahlhaus (2006).

The bootstrap algorithm proposed to approximate the distribution of Q_T under the null hypothesis of a tvAR(p) process consists then of the following four Steps:

STEP 1: Fit locally the time-varying autoregressive model of order p to the observations $X_{1,T}, X_{2,T}, \ldots, X_{T,T}$ and calculate the estimated parameters $\hat{\beta}_{t/T}(p)' = (\hat{\beta}_1(t/T), \ldots, \hat{\beta}_p(t/T))$ and $\hat{\sigma_p}^2(t/T)$.

STEP 2: Generate bootstrap observations $X^+_{1,T}, X^+_{2,T}, \ldots, X^+_{T,T}$ using the fitted local autoregressive model, that is,

$$X^+_{t,T} = \sum_{j=1}^{p} \hat{\beta}_j(\frac{t}{T})X^+_{t-j,T} + \hat{\sigma}_p(\frac{t}{T}) \cdot \varepsilon^+_t,$$

where $X^+_{j,T} = X_{j,T}$ for $j = 1, 2, \ldots, p$ and ε^+_t are i.i.d random variables with $\varepsilon^+_t \backsim N(0,1)$.

STEP 3: Compute the local periodogram $I_N^+(u,\lambda)$ over segments of length N of the bootstrap pseudo-observations $X_{t,T}^+$, i.e., compute

$$I_N^+(u,\lambda) = \frac{1}{2\pi H_{2,N}(0)} |d_N^+(u,\lambda)|^2 \tag{14}$$

where

$$d_N^+(u,\lambda) = \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right) X_{[uT]-N/2+s+1}^+ e^{-i\lambda s}.$$

STEP 4: The bootstrapped test statistic is then defined by

$$Q_T^+ = \frac{1}{M}\sum_{i=1}^{M}\int_{-\pi}^{\pi}\left\{\frac{1}{N}\sum_{j=-M_N}^{M_N} K_b(\lambda-\lambda_j)\left(\frac{I_N^+(u_i,\lambda_j)}{f(u_i,\lambda_j;\hat{\vartheta})}-1\right)\right\}^2 d\lambda$$

Notice that we could have in STEP 4 rescaled the local bootstrap periodogram $I_N^+(u,\lambda)$ by $f(u_i,\lambda_j;\hat{\vartheta}^+)$ instead by $f(u_i,\lambda_j;\hat{\vartheta})$, where $\hat{\vartheta}(\cdot)^+$ denotes the estimator of the autoregressive parameter functions $\vartheta(\cdot)$ obtained using the bootstrap pseudo-series $X_{1,T}^+$, $X_{2,T}^+$, $\ldots, X_{T,T}^+$. The specification of Q_T^+ used is, however, preferred because besides of being computationally more convenient, it is also justified theoretically by the fact that the limiting distribution of the test statistic Q_T under the null is not affected if the unknown $\vartheta(\cdot)$ is replaced by a $\sqrt{N}$-consistent estimator $\hat{\vartheta}(\cdot)$.

To ensure that the estimated local autoregressive process generating the bootstrap observations $X_{1,T}^+, \ldots, X_{T,T}^+$ is locally stationary, we assume that the parameter estimators used in Step 1 of the bootstrap algorithm satisfy the following assumption.

Assumption 3.2. $T_0 \in \mathbb{N}$ exists such that for $T \geq T_0$ the estimated coefficients $\widehat{\beta}_j(u)$ and $\widehat{\sigma}_N^2(u)$ are Lipschitz continuous in u with the Lipschitz constant independent of T, and, $1-\sum_{j=1}^{p}\widehat{\beta}_j(u)z^j \neq 0$ for $|z| \leq 1+c$ with $c>0$ independent of T and uniformly in u.

The following theorem shows that the bootstrap procedure proposed leads to an asymptotically valid approximation of the distribution of the test statistic Q_T under the null hypothesis of a tvAR(p) process. We follow Bickel and Freedman (1981) and measure the distance between distributions F and G by the Mallow metric d_2 defined as $d_2(F,G) = \inf\{E|X-Y|^2\}^{1/2}$, where the infimum is taken over all pairs of random variables X and Y having marginal distributions F and G respectively. We adopt the convention that where random variables appear as arguments of d_2 these represent the corresponding distributions.

Theorem 3.2. *Let Assumptions 2.1, 2.2, 3.1 and 3.2 be satisfied. Then, conditionally on* $X_{1,T}, X_{2,T}, \ldots, X_{T,T}$*, we have as* $T \to \infty$*,*

$$d_2\Big(N\sqrt{Mb}(Q_T^+ - \mu_T), N\sqrt{Mb}(Q_T - \mu_T)\Big) \;\to\; 0,$$

in probability, where μ_T *is defined in Theorem 2.1.*

4. Applications

4.1. **Some remarks on choosing the testing parameters.** From the previous discussion it is clear that implementation of the testing procedure proposed, requires essentially the selection of two parameters: the time window width N and the smoothing bandwidth b. Although a through investigation of this problem is beyond the scope of this paper, we give in what follows a rather heuristic discussion on how to select this parameters in practice.

Concerning the value of the time window width N, we mention that the selection of this parameter is inherit to any statistical inference procedure for locally stationary process which is based on segments of observations. Choosing N too large will induce a large bias since a large N is associated with a loss of information on the local structure of the underlying process. On the other hand, choosing N too small will lead to an increase of the variance of the estimators involved due to the small number of observations used. Any approach to select N should therefore be guided by the requirement that N should be large enough to allow for reasonable local estimation but not too large to avoid a 'smoothing out' the interesting local characteristics of the process. Based on this observation and depending on the overall size n of the time series at hand, the choices $N = 64$ or $N = 128$ are convenient in most situations.

Concerning the choice of the smoothing parameter b, one way to proceed is to select this parameter using a local version of a cross-validation criterion like the one proposed by Beltrão and Bloomfield (1987). To elaborate on, notice first that our aim is to obtain a "good" estimate of the function $q(u, \lambda) = f(u, \lambda)/f(u, \lambda; \vartheta)$. To stress the dependence of this function on the estimated parametric curves, we write $q(u, \lambda, \vartheta(u))$ in the sequel. Using as a starting point the function

$$\sum_{i=1}^{M}\sum_{j=1}^{N}\left\{\log q(u_i, \lambda_j; \hat{\vartheta}) + \frac{I_N(u_i, \lambda_j)/f(u_i, \lambda_j, \hat{\vartheta})}{q(u_i, \lambda_j; \hat{\vartheta})}\right\}, \tag{15}$$

a leave-one-out estimator of $q(u,\lambda_j;\vartheta)$ is given by

$$\hat{q}_{-j}(u,\lambda_j;\hat{\vartheta}) = \frac{1}{N}\sum_{j\in N_j} K_h(\lambda_j-\lambda_s)\frac{I_N(u,\lambda_j)}{f(u,\lambda_j;\hat{\vartheta})} \tag{16}$$

where $N_j = \{s : -M_N \le s \le M_N \text{ and } j-s \neq \pm j \mod M_N\}$. Notice that $\hat{q}_{-j}$ is a kernel estimator of q obtained by ignoring the jth ordinate of the local periodogram $I_N(u,\lambda_j)$. Now, substituting $\hat{q}_{-j}(u,\lambda_j;\hat{\vartheta})$ for $q(u,\lambda_j;\hat{\vartheta})$ in (15) leads to the function

$$CV(b) = \sum_{i=1}^{M}\sum_{j=1}^{N}\left\{\log \hat{q}_{-j}(u_i,\lambda_j;\hat{\vartheta}) + \frac{I_N(u_i,\lambda_j)/f(u_i,\lambda_j;\hat{\vartheta})}{\hat{q}_{-j}(u_i,\lambda_j;\hat{\vartheta})}\right\}, \tag{17}$$

which can be used as a cross-validation-type criterion to select b.

4.2. **Simulations.**

4.2.1. *Bootstrap Approximations.* We first, illustrate the advantages of using the bootstrap procedure proposed by comparing its performance in approximating the distribution of Q_T under the null with that of the limiting Gaussian approximation. For this purpose, observations $\{X_{t,T}, t = 1,\ldots,T\}$ from the first order, time-varying autoregressive model

$$X_{t,T} = \phi(\frac{t}{T})X_{t-1,T} + \varepsilon_t \tag{18}$$

have been generated, where $\phi(t/T) = 0.9\cos(1.5-\cos(4\pi(t/T)))$ and the ε_t's are i.i.d. random variables with $\varepsilon_t \sim N(0,1)$. To estimate the exact distribution of the test statistic Q_T we generate 1000 series of length $T = 1024$ and for each of these series we calculated Q_T using the Bartlett-Priestley kernel, $K(x) = \mathbf{1}_{[-\pi,\pi]}(x)3(4\pi)^{-1}(1-(x/\pi)^2)$ and the bandwidth $b = 0.2$. The window width N has been set equal to $N = 128$ and two different shifts, $S = 128$ and $S = 64$, have been considered. Notice that for $S = 128$ we have $\kappa = 1$, while for $S = 64$, $\kappa = 2$.

To investigate the performance of the bootstrap method, we choose randomly 21 series from the generated 1000 replications of process (18) and for each of the selected series we apply the bootstrap produce proposed using 300 bootstrap replications. Based on the bootstrap replications, we estimated for each series the density $\hat{g}^*$ of the corresponding bootstrap approximation of the distribution of Q_T. We also estimated the density of the exact distribution of Q_T based on the 1000 replications of process (18). The so estimated density is denoted by $\hat{g}$. The density estimates $\hat{g}^*$ and $\hat{g}$ have been obtained using standard SPlus smoothing routines. We then compare the estimated exact density $\hat{g}$ with the Gaussian approximation given in Theorem 2.1 and with the median bootstrap approximation. The median bootstrap approximations is that for which $\sum_{x_i}|\hat{g}^*(x_i) -$

$\hat{g}(x_i)|$ takes its median value over the 21 series used. Figure 1 shows the estimated densities of the exact, the asymptotic Gaussian and the median bootstrap approximation. As it is clearly seen from these exhibits, the bootstrap performs much better compared to the Gaussian approximation and estimates very accurately the exact distribution of interest.

Please insert Figure 2 about here.

4.2.2. *Size and power performance of the test.* We next investigate the size and the power performance of the test in finite sample situations by means of a small simulation study. For this, we consider realizations of length $T = 512$ and $T = 1024$ of the time-varying AR(2) model

$$X_{t,T} = 0.9\cos(1.5 - \cos(4\pi t/T))X_{t-1,T} - \phi_2 X_{t-2,T} + \varepsilon_t \tag{19}$$

where the ε_t's are independent, standard Gaussian distributed random variables. The null hypothesis is that the underlying process is a time-varying first order autoregressive process. Different values of the parameter ϕ_2 have been considered corresponding to validity of the null ($\phi_2 = 0$) and of the alternative hypothesis ($\phi_2 \neq 0$). In each case we fit a time-varying AR(1) model using a local least squares estimator and compute the test statistic Q_T using the Bartlett-Priestley kernel and different values of the bandwidth parameter b. We also apply the test proposed for different segment lengths N and shifts S. In all cases the critical values of the test have been obtained using B=500 replications of the bootstrap procedure described in Section 3. The results obtained over 500 replications are summarized in Table 1.

Please insert Table 1 about here.

As Table 1 shows, estimating the critical values of the test using the bootstrap procedure proposed, leads to a very good size and power behavior of the test. Notice that for both sample sizes and all combinations of bandwidth values, segment lengths and shifts considered, the empirical size of the test is very close to the nominal level of 5%. Furthermore, under the alternative, the test has power even for small deviations from the null and the power of the test increases rapidly approaching unity as the deviations from the null and/or the sample size become larger.

We conclude this section with an investigation of the power behavior of the test in a situation which is not covered by our theoretical analysis, namely that of a sudden change in the model structure of the underlying process. In particular, we generate observations from the piecewise stationary process

$$X_{t,T} = \begin{cases} 0.9X_{t-1,T} + \varepsilon_t & \text{for } \ t = 1, 2, \ldots, [T/2] \\ \\ -05X_{t-1,T} + \varepsilon_t & \text{for } \ t = [T/2]+1, [T/2]+2, \ldots, T, \end{cases}$$

where the ε_t's are independent, standard Gaussian distributed random variables. We fit to the generated time series a first order, time-varying autoregressive process and investigate the power of our test is detecting this erroneous model fit. Table 2 summarizes the results obtained for two sample sizes, $T = 512$ and $T = 1024$ observations. Notice that the critical values of the test have been calculated using $B = 500$ bootstrap replications. As this table shows, the test proposed seems to have considerable power in detecting model misspecifications even in the case of sudden changes in the stochastic structure of the underlying process.

Please insert Table 2 about here.

5. Conclusions

In this paper we have introduced and investigated properties of a test of the hypothesis that the time-varying spectral density of a locally stationary process has a semiparametric structure. Our approach is general enough and allows for testing the interesting case of a time-varying autoregressive moving-average model. The test introduced is based on a L_2-distance measure of a kernel smoothed version of the local periodogram rescaled by the time-varying spectral density of the estimated semiparametric model under the null. The asymptotic distribution of the test statistic under the null hypothesis has been derived and it has been shown that this distribution is Gaussian with parameters that do not depend on characteristics of the underlying stochastic process. As an interesting special case, we focused then on the problem of testing for the presence of a time-varying autoregressive model structure. A semiparametric bootstrap procedure to approximate more accurately the distribution of the test statistic under the null hypothesis has been proposed and its asymptotic validity has been established. The favorable size and power properties of our test in finite sample situations have been demonstrated by means of some simulations and numerical examples.

Appendix: Auxiliary results and proofs

A useful tool for handling taper data, is the periodic extension (with period 2π) of the function $L_T(\alpha) : \mathbb{R} \to \mathbb{R}$, with

$$L_T(\alpha) = \begin{cases} T, & |\alpha| \le 1/T \\ 1/|\alpha|, & 1/T \le |\alpha| \le \pi. \end{cases} \tag{20}$$

Lemma 5.1 and Lemma 5.2 below can be proved as in Dahlhaus (1997), Lemma A.5 and A.6.

Lemma 5.1.

a. $\int_\Pi L_T^k(\alpha) \le KT^{k-1}$ *for all* $k > 1$.

b. $\int_\Pi L_T(\alpha) \le K \log(T)$

c. $|\alpha| L_T(\alpha) \le K$

d. $\int_\Pi L_T(\beta-\alpha) L_S(\alpha+\gamma) \le K \max\{\log(T), \log(S)\} L_{\min\{T,S\}}(\beta+\gamma)$

For a complex-valued function f define $H_N(f(\cdot),\lambda) := \sum_{s=0}^{N-1} f(s)e^{-i\lambda s}$ and let

$$H_{k,N}(\lambda) = H_N(h^k(\frac{\cdot}{N}),\lambda)$$

and

$$H_N(\lambda) = H_{1,N}(\lambda).$$

Straightforward calculation gives

$$\sum_j H_{k,N}(\alpha-\lambda_j) H_{\ell,N}(\lambda_j-\beta) = 2\pi N H_{k+\ell,N}(\alpha-\beta)$$

where the sum extends over all Fourier frequencies $\lambda_j = 2\pi j/N,\ \ j = -[(N-1)/2], \ldots, [N/2]$. Under Assumption 2.2 (iv) there is a constant C independent of T and λ such that

$$|H_{k,N}(\lambda)| \le C L_N(\alpha) \tag{21}$$

and

$$K_b(\lambda) \le C b L_{1/b}^2(\lambda). \tag{22}$$

Lemma 5.2.

(i) *Let* $N, T \in \mathbb{N}$. *Suppose that the data taper* h *satisfies Assumption 2.2 (iv) and* $\psi : [0,1] \to \mathbb{R}$ *is Lipshitz continuous. Then we have for* $0 \le t \le N$, *that,*

$$H_N\left(\psi\left(\frac{\cdot}{T}\right) h\left(\frac{\cdot}{N}\right),\lambda\right) = \psi\left(\frac{t}{T}\right) H_N(\lambda) + O\left(\frac{N}{T} L_N(\lambda)\right).$$

The same holds, if $\psi(\cdot/T)$ *on the left side is replaced by numbers* $\psi_{s,T}$ *with* $\sup_s |\psi_{s,T} - \psi(s/T)| = O(T^{-1})$

(ii) *Let $t_j = S(j-1) + N/2$, $u_j = t_j/T$ with N, M, S and T satisfying Assumption 2.2 and $\psi : [0,1] \to \mathbb{R}$ be Lipshitz continuous. Then*

$$\left| \sum_{j=1}^{M} \psi(u_j) e^{i\lambda S j} \right| \leq K L_M(S\lambda).$$

Before proceeding with the next lemma we use for simplicity the notation $f_{\vartheta_0}(u,\lambda)$ for $f(u,\lambda;\vartheta_0(u))$ and $f_{\hat{\vartheta}}(u,\lambda)$ for $f(u,\lambda;\hat{\vartheta}(u))$.

Lemma 5.3. *Under Assumptions 2.1, 2.2 and if H_0 is true, then*

$$E(N\sqrt{Mb}Q_{0,T}) = \mu_T + o(1)$$

where

$$\mu_T = \frac{M^{1/2} c_{tap}}{b^{1/2}} \int_{-\pi}^{\pi} K^2(x)dx + \frac{M^{1/2} b^{1/2} c_{tap}}{4\pi} \int_{-\pi}^{\pi} \int_{-2\pi}^{2\pi} K(x)K(x-u)dxdu$$

and $c_{tap} = \int_0^1 h^4(x)/(\int_0^1 h^2(x))^2$.

Proof: First note that

$$\begin{aligned}
E(N\sqrt{Mb}Q_{0,T}) &= \frac{b^{1/2}}{M^{1/2}N} \sum_{m=1}^{M} \int_{-\pi}^{\pi} \sum_j \sum_s \frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)}{f_{\vartheta_0}(u_i,\lambda_j) f_{\vartheta_0}(u_i,\lambda_s)} cum\left(I_N(u_m,\lambda_j), I_N(u_m,\lambda_s)\right) d\lambda \\
&\quad + O(\sqrt{Mb}N^5/T^4 + \sqrt{Mb}\log^2(N)/N) \\
&= \frac{b^{1/2}}{4\pi^2 M^{1/2} N H_{2,N}^2(0)} \sum_{m=1}^{M} \int_{-\pi}^{\pi} \sum_j \sum_s \frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)}{f_{\vartheta_0}(u_m,\lambda_j) f_{\vartheta_0}(u_m,\lambda_s)} d\lambda \\
&\quad \times \Big(cum\left(d_N(u_m,\lambda_j), d_N(u_m,\lambda_s)\right) cum\left(d_N(u_m,-\lambda_j), d_N(u_m,-\lambda_s)\right) \\
&\quad + cum\left(d_N(u_m,\lambda_j), d_N(u_m,-\lambda_s)\right) cum\left(d_N(u_m,-\lambda_j), d_N(u_m,\lambda_s)\right) \Big) + o(1) \\
&= \mu_{1,T} + \mu_{2,T} + o(1)
\end{aligned}$$

with an obvious notation for $\mu_{i,T}$, $i = 1,2$. Recall the definition of $d_N(u,\lambda)$ to see that

$$\begin{aligned}
& cum\left(d_N(u_m,\lambda_j), d_N(u_m,\lambda_s)\right) cum\left(d_N(u_m,-\lambda_j), d_N(u_m,-\lambda_s)\right) \\
= & \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} H_N(A^0_{t_m-N/2+1+\cdot,T}(\mu_1) h(\frac{\cdot}{N}), \lambda_j - \mu_1) H_N(A^0_{t_m-N/2+1+\cdot,T}(-\mu_1) h(\frac{\cdot}{N}), -\lambda_s + \mu_1) \\
& \times H_N(A^0_{t_m-N/2+1+\cdot,T}(\mu_2) h(\frac{\cdot}{N}), -\lambda_j - \mu_2) H_N(A^0_{t_m-N/2+1+\cdot,T}(-\mu_2) h(\frac{\cdot}{N}), \lambda_s + \mu_2) d\mu_1 d\mu_2.
\end{aligned}$$

Substituting $A^0_{t_m-N/2+1+\cdot,T}(\mu_2)$ by $A(t/T,\mu_2)$ on the above expression, using (3) and the fact that $A(\cdot,\cdot)$ is Lipshitz continuous, we get that the above term is equal to

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_{\vartheta_0}(u_m,\mu_1)f_{\vartheta_0}(u_m,\mu_2)H_N(\lambda_j-\mu_1)H_N(-\lambda_s+\mu_1)H_N(-\lambda_j-\mu_2)H_N(\lambda_s+\mu_2)d\mu_1 d\mu_2 \tag{23}$$

plus a remainder term R_m depending on the difference $|A^0_{t_m-N/2+1+\cdot,T}(\mu_2)-A(t/T,\mu_2)|$ which satisfies

$$\begin{aligned}
&\left|\frac{b^{1/2}}{M^{1/2}N^3}\sum_{m=1}^{M}\sum_{j=-J_N}^{J_N}\sum_{s=-J_N}^{J_N}\frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)}{f_{\vartheta_0}(u_m,\lambda_j)f_{\vartheta_0}(u_m,\lambda_s)}R_m\right| \\
\leq\ & \frac{N}{T}\frac{b^{1/2}M^{1/2}}{N^3}\log^2(N)b\sum_{j=-J_N}^{J_N}\sum_{s=-J_N}^{J_N}L^2_{1/b}(\lambda_j-\lambda_s)L^2_N(\lambda_j-\lambda_s) \\
=\ & O(\frac{NM^{1/2}\log^2(N)}{b^{1/2}T}).
\end{aligned} \tag{24}$$

Using the bound (24) and replacing $f_{\vartheta_0}(u_m,\mu_1)$ and $f_{\vartheta_0}(u_m,\mu_2)$ by $f_{\vartheta_0}(u_m,\lambda_j)$ and $f_{\vartheta_0}(u_m,\lambda_s)$ respectively, we get that the term $\mu_{1,T}$ is equal to

$$\begin{aligned}
&\frac{b^{1/2}M^{1/2}}{NH^2_{2,N}(0)}\int_{-\pi}^{\pi}\sum_{j=-J_N}^{J_N}\sum_{s=-J_N}^{J_N}K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)|H_N(\lambda_j-\lambda_s)|^2 d\lambda \\
=\ & \frac{M^{1/2}c_{tap}}{b^{1/2}}\int_{-\pi}^{\pi}K^2(x)dx+O(\frac{\log(N)M^{1/2}}{Nb^{3/2}})
\end{aligned}$$

plus a remainder term R_m which depends on the difference $|f_{\vartheta_0}(u_m,\mu_2)-f_{\vartheta_0}(u_m,\lambda_j)|$ and which satisfies

$$\begin{aligned}
&\left|\frac{b^{1/2}}{M^{1/2}N^3}\sum_{m=1}^{M}\sum_{j=-J_N}^{J_N}\sum_{s=-J_N}^{J_N}\frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)}{f_{\vartheta_0}(u_m,\lambda_j)f_{\vartheta_0}(u_m,\lambda_s)}R_m\right| \\
\leq\ & \frac{b^{1/2}M^{1/2}}{N^3}\int_{-\pi}^{\pi}\sum_{j=-J_N}^{J_N}\sum_{s=-J_N}^{J_N}K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)L_N(\lambda_s-\lambda_j)d\lambda \\
=\ & O(\frac{b^{1/2}M^{1/2}\log^2(N)}{N}).
\end{aligned}$$

Similar arguments yield that the second term $\mu_{2,T}$ is equal to

$$
\begin{aligned}
& \frac{b^{1/2}M^{1/2}}{NH_{2,N}^2(0)}\int_{-\pi}^{\pi}\sum_{j=-J_N}^{J_N}\sum_{s=-J_N}^{J_N} K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)|H_N(\lambda_j+\lambda_s)|^2 d\lambda + o(1)\\
= \ & \frac{M^{1/2}b^{1/2}c_{tap}}{4\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} K(x)K(x-u)dxdu + o(1)
\end{aligned}
$$

■

Lemma 5.4. *Under Assumptions 2.1 and 2.2 and if H_0 is true, then*

$$
Var(N\sqrt{Mb}Q_{0,T}) = \tau^2 + o(1)
$$

where τ^2 is defined in Theorem 2.1.

Proof: First note that

$$
\begin{aligned}
& Var(N\sqrt{Mb}Q_{0,T})\\
= \ & \frac{b}{MN^2}\sum_{m_1=1}^{M}\sum_{m_2=1}^{M}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,s,k,l=-J_N}^{J_N} \frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)}\\
& \Big(cum(I_N(u_{m_1},\lambda_j),I_N(u_{m_2},\lambda_k))cum(I_N(u_{m_1},\lambda_s),I_N(u_{m_2},\lambda_l))\\
& +cum(I_N(u_{m_1},\lambda_j),I_N(u_{m_2},\lambda_l))cum(I_N(u_{m_1},\lambda_s),I_N(u_{m_2},\lambda_k))\\
& +cum(I_N(u_{m_1},\lambda_j),I_N(u_{m_2},\lambda_k),I_N(u_{m_1},\lambda_s),I_N(u_{m_2},\lambda_l))\Big)d\lambda d\mu + o(1)\\
= \ & V_{1,T}+V_{2,T}+V_{3,T}
\end{aligned}
$$

with an obvious notation for $V_{i,T}$ $i = 1,2,3$. From (7) and using the fact that $\xi(\lambda)$ is Gaussian and that by Isserlis theorem $cum(Z_1Z_2,Z_3Z_4) = cum(Z_1,Z_3)cum(Z_2,Z_4) + cum(Z_1,Z_4)cum(Z_2,Z_3)$, for Z_i Gaussian random variables, we get that the term $V_{1,T}$ can be further decomposed as the sum of four terms, that is we can write $V_{1,T} = \sum_{i=1}^{4} V_{j,T}^{(i)}$.

The first term in this decomposition, $V_{1,T}^{(1)}$, equals

$$
\begin{aligned}
V_{1,T}^{(1)} \ = \ & \frac{b}{MN^2}\sum_{m_1,m_2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,k,l,s}\frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)}\\
& cum\left(d_N(u_{m_1},\lambda_j),d_N(u_{m_2},-\lambda_k)\right)cum\left(d_N(u_{m_1},-\lambda_j),d_N(u_{m_2},\lambda_k)\right)\\
& \times cum\left(d_N(u_{m_1},\lambda_s),d_N(u_{m_2},-\lambda_l)\right)cum\left(d_N(u_{m_1},-\lambda_s),d_N(u_{m_2},\lambda_l)\right)
\end{aligned}
$$

Using arguments similar to those used in the proof of Lemma 5.3 we have that the term

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\mu_1)h(\frac{\cdot}{N}),\lambda_j-\mu_1)H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(-\mu_1)h(\frac{\cdot}{N}),-\lambda_k+\mu_1)$$
$$\times H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\mu_2)h(\frac{\cdot}{N}),-\lambda_j-\mu_2)H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(-\mu_2)h(\frac{\cdot}{N}),\lambda_k+\mu_2)$$
$$\times H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\mu_3)h(\frac{\cdot}{N}),\lambda_s-\mu_3)H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(-\mu_3)h(\frac{\cdot}{N}),-\lambda_l+\mu_3)$$
$$\times H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\mu_4)h(\frac{\cdot}{N}),-\lambda_s-\mu_4)H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(-\mu_4)h(\frac{\cdot}{N}),\lambda_l+\mu_4)$$
$$\times \exp\{i(\mu_1+\mu_2+\mu_3+\mu_4)(t_{m_1}-t_{m_2})\}\, d\mu_1 d\mu_2 d\mu_3 d\mu_4$$

is equal to

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} A(u_{m_1},\mu_1)A(u_{m_2},-\mu_1)A(u_{m_1},\mu_2)A(u_{m_2},-\mu_2)A(u_{m_1},\mu_3)A(u_{m_2},-\mu_3)A(u_{m_1},\mu_4)$$
$$\times A(u_{m_2},-\mu_4)H_N(\lambda_j-\mu_1)H_N(-\lambda_k+\mu_1)H_N(-\lambda_j-\mu_2)H_N(\lambda_k+\mu_2)H_N(\lambda_s-\mu_3)H_N(-\lambda_l+\mu_3)$$
$$\times H_N(-\lambda_s-\mu_4)H_N(\lambda_l+\mu_4)\exp\{i(\mu_1+\mu_2+\mu_3+\mu_4)(t_{m_1}-t_{m_2})\}\, d\mu_1 d\mu_2 d\mu_3 d\mu_4$$
$$+R_1(m_1,m_2) \tag{25}$$

where $R_1(m_1,m_2)$ satisfies

$$\frac{b}{MN^2}\sum_{m_1,m_2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,k,s,l}\frac{K_b(\lambda-\lambda_j)K_b(\mu-\lambda_s)K_b(\lambda-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)}R_1(m_1,m_2)d\lambda d\mu$$
$$\leq \frac{K_1}{H_N^4}\frac{b}{MN^2}\frac{N}{T}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,k,s,l}K_b(\lambda-\lambda_j)K_b(\mu-\lambda_s)K_b(\lambda-\lambda_k)K_b(\mu-\lambda_l)$$
$$\times L_N(\lambda_j-\mu_1)L_N(-\lambda_k+\mu_1)L_N(-\lambda_j-\mu_2)L_N(\lambda_k+\mu_2)L_N(\lambda_s-\mu_3)H_N(-\lambda_l+\mu_3)$$
$$\times L_N(-\lambda_s-\mu_4)L_N(\lambda_l+\mu_4)L_M^2(S(\mu_1+\mu_2+\mu_3+\mu_4))d\mu_1 d\mu_2 d\mu_3 d\mu_4 d\lambda d\mu \tag{26}$$

since by Lemma A.6 of Dahlhaus (1997) we have

$$\sum_{m=1}^{M}\frac{1}{f_{\vartheta_0}(u_m,\lambda_k)f_{\vartheta_0}(u_m,\lambda_l)}e^{i((\mu_1+\mu_2+\mu_3+\mu_4)(Sm)} = O(L_M(S(\mu_1+\mu_2+\mu_3+\mu_4))).$$

Now using Lemma A.4(e) and Lemma A.4(j) of Dahlhaus (1997), expression (26) can be bounded by

$$\frac{K_1}{H_N^4}\frac{b}{MN^2}\frac{N}{T}\frac{NM}{S}\log^3(N)\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,s,k,l}K_b(\lambda-\lambda_j)K_b(\mu-\lambda_s)K_b(\lambda-\lambda_k)K_b(\mu-\lambda_l)L_N^2(\lambda_j-\lambda_k)$$
$$\times L_N^2(\lambda_s-\lambda_l)d\lambda d\mu = O(\frac{\log^3(N)N^2}{ST}).$$

Furthermore, replacing $A(u_i,\mu_1)$ by $A(u_i,\lambda_j)$ in the first term of (25) we get that this term can be written as

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} A(u_{m_1},\lambda_j)A(u_{m_2},-\mu_1)A(u_{m_1},\mu_2)A(u_{m_2},-\mu_2)A(u_{m_1},\mu_3)A(u_{m_2},-\mu_3)A(u_{m_1},\mu_4)$$
$$\times A(u_{m_2},-\mu_4)H_N(\lambda_j-\mu_1)H_N(-\lambda_k+\mu_1)H_N(-\lambda_j-\mu_2)H_N(\lambda_k+\mu_2)H_N(\lambda_s-\mu_3)H_N(-\lambda_l+\mu_3)$$
$$\times H_N(-\lambda_s-\mu_4)H_N(\lambda_l+\mu_4)\exp\{i(\mu_1+\mu_2+\mu_3+\mu_4)(t_{m_1}-t_{m_2})\}\,d\mu_1 d\mu_2 d\mu_3 d\mu_4$$
$$+R_2(m_1,m_2) \tag{27}$$

where the remainder term $R_2(m_1,m_2)$ satisfies

$$\frac{b}{MN^2H_N^4}\sum_{m_1,m_2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,k,s,l}\frac{K_b(\lambda-\lambda_j)K_b(\mu-\lambda_s)K_b(\lambda-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)}R_2(m_1,m_2)d\lambda d\mu$$
$$\leq \frac{K_1}{H_N^4}\frac{b}{MN^2}\frac{MN}{S}\log^3(N)\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,k,s,l}K_b(\lambda-\lambda_j)K_b(\mu-\lambda_s)K_b(\lambda-\lambda_k)K_b(\mu-\lambda_l)$$
$$\times L_N(\lambda_j-\lambda_k)L_N^2(\lambda_l-\lambda_s)d\lambda d\mu = O(\frac{\log^3(N)}{S}).$$

From (27) we get that

$$V_{1,T}^{(1)} = \frac{b}{16\pi^4MN^2H_N^4b^4}\sum_{m_1,m_2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\left[\sum_{j,k}K(\frac{\lambda-\lambda_j}{b})K(\frac{\mu-\lambda_k}{b})\sum_{s_1,s_2,s_3,s_4}h(\frac{s_1}{N})h(\frac{s_2}{N})h(\frac{s_3}{N})h(\frac{s_4}{N})\right.$$
$$\left.\times e^{-i[s_1\lambda_j-s_2\lambda_k-s_3\lambda_j+s_4\lambda_k]}\int_{-\pi}^{\pi}e^{i[\mu_1(s_1-s_2)+\mu_1S(m_1-m_2)]}d\mu_1\int_{-\pi}^{\pi}e^{i[\mu_2(s_3-s_4)+\mu_2S(m_1-m_2)]}d\mu_2\right]^2 d\lambda d\mu$$
$$+o(1)$$
$$= \frac{b}{N^2H_N^4b^4}\sum_{|m|<N/S}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\left[\sum_{j,k}K(\frac{\lambda-\lambda_j}{b})K(\frac{\mu-\lambda_k}{b})\right.$$
$$\left.\times\left|\sum_{s_1=0}^{N-1-Sm}h(\frac{s_1}{N})h(\frac{s_1+Sm}{N})e^{-i[s_1(\lambda_j-\lambda_k)]}\right|^2\right]^2 d\lambda d\mu + o(1),$$

which by straightforward calculations yield

$$V_{1,T}^{(1)} = \frac{\sum_{|m|<\kappa}\left(\int_0^{1-m/\kappa}h^2(u)h^2(u+m/\kappa)du\right)^2}{2\pi\left(\int_0^1 h^2(x)\right)^4}\int_{-2\pi}^{2\pi}\left(\int K(u)K(u+x)du\right)^2 dx + O(\frac{\log^2(N)}{N^2b^4}).$$

The terms $V_{1,T}^{(j)}$, $j = 2, 3, 4$ are handled similarly and we get

$$V_{1,T}^{(2)} = \frac{\sum_{|m|<\kappa} \left(\int_0^{1-m/\kappa} h^2(u) h^2(u+m/\kappa) du\right)^2}{2\pi \left(\int_0^1 h^2(x)\right)^4} \int_{-2\pi}^{2\pi} \left(\int K(u)K(u-x)du\right)^2 dx + o(1),$$

$$\begin{aligned} V_{1,T}^{(3)} &= \frac{b}{MN^2} \sum_{m_1,m_2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sum_{j,k,l,s} \frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)} \\ &\quad \times cum\left(d_N(u_{m_1},\lambda_j), d_N(u_{m_2},\lambda_k)\right) cum\left(d_N(u_{m_1},-\lambda_j), d_N(u_{m_2},-\lambda_k)\right) \\ &\quad \times cum\left(d_N(u_{m_1},\lambda_s), d_N(u_{m_2},-\lambda_l)\right) cum\left(d_N(u_{m_1},-\lambda_s), d_N(u_{m_2},\lambda_l)\right) = O(b) \end{aligned}$$

and

$$\begin{aligned} V_{1,T}^{(4)} &= \frac{b}{MN^2} \sum_{m_1,m_2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sum_{j,k,l,s} \frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)} \\ &\quad \times cum\left(d_N(u_{m_1},\lambda_j), d_N(u_{m_2},-\lambda_k)\right) cum\left(d_N(u_{m_1},-\lambda_j), d_N(u_{m_2},\lambda_k)\right) \\ &\quad \times cum\left(d_N(u_{m_1},\lambda_s), d_N(u_{m_2},\lambda_l)\right) cum\left(d_N(u_{m_1},-\lambda_s), d_N(u_{m_2},-\lambda_l)\right) \\ &= O(b). \end{aligned}$$

The term $V_{2,T}$ has the same structure as the term $V_{1,T}$ and converges, therefore, to the same limit. Finally,

$$\begin{aligned} V_{3,T} &= \frac{b}{MN^2} \sum_{m_1=1}^{M} \sum_{m_2=1}^{M} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sum_{j,s,k,l=-J_N}^{J_N} \frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)} \\ &\quad \times cum(d_N(u_{m_1},\lambda_j)d_N(u_{m_1},-\lambda_j), d_N(u_{m_2},\lambda_k)d_N(u_{m_2},-\lambda_k), d_N(u_{m_1},\lambda_s)d_N(u_{m_1},-\lambda_s), \\ &\quad d_N(u_{m_2},\lambda_l)d_N(u_{m_2},-\lambda_l))d\lambda d\mu \end{aligned}$$

and using properties of the cumulants it can be shown by cumbersome but straightforward calculations that this term converges to zero.

To handle this term notice that using the product theorem of cumulants, see Brillinger (1981), we have to sum over all indecomposable partitions $P_1, \ldots, P_m$ of the scheme

$$\begin{matrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \\ a_4 & b_4 \end{matrix}$$

where a_1 stands for the position of $d_N(u_{m_1},\lambda_j)$, b_1 for the position of $d_N(u_{m_1},-\lambda_j)$, etc. Following the notation of Dahlhaus (1997), let $P_i=\{c_1,\ldots,c_k\}$, $\overline{P}_i:=\{c_1,\ldots,c_{k-1}\}$, $\beta_{P_i}:=(\beta_{c_1},\ldots,\beta_{c_{k-1}})$ and $\beta_{c_k}=-\sum_{j=1}^{k-1}\beta_{c_j}$. Also, let m be the size of the corresponding partition and $\beta=(\beta_{\overline{P}_i},\ldots,\beta_{\overline{P}_m})$. We then get

$$\begin{aligned}
V_{3,T} &= \frac{b}{MN^2H_N^4}\sum_{ip}\sum_{m_1=1}^{M}\sum_{m_2=1}^{M}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,s,k,l}\frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)} \\
&\quad \int_{\Pi^{8-m}} H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\beta_{a_1})h(\tfrac{\cdot}{N}),\lambda_j-\beta_{a_1})H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\beta_{b_1})h(\tfrac{\cdot}{N}),-\lambda_j-\beta_{b_1}) \\
&\quad \times H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(\beta_{a_2})h(\tfrac{\cdot}{N}),\lambda_k-\beta_{a_2})H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(\beta_{b_2})h(\tfrac{\cdot}{N}),-\lambda_k-\beta_{b_2}) \\
&\quad \times H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\beta_{a_3})h(\tfrac{\cdot}{N}),\lambda_s-\beta_{a_3})H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\beta_{b_3})h(\tfrac{\cdot}{N}),-\lambda_s-\beta_{b_3}) \\
&\quad \times H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(\beta_{a_4})h(\tfrac{\cdot}{N}),\lambda_l-\beta_{a_4})H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(\beta_{b_4})h(\tfrac{\cdot}{N}),-\lambda_l-\beta_{b_4}) \\
&\quad \prod_{\nu=1}^{m} g_{|P_\nu|}(\beta_{\overline{P}_\nu})\exp\{i(t_{m_1}(\beta_{a_1}+\beta_{b_1}+\beta_{a_3}+\beta_{b_3}) \\
&\qquad\qquad + t_{m_2}(\beta_{a_2}+\beta_{b_2}+\beta_{a_4}+\beta_{b_4}))\}\,d\beta d\lambda d\mu \qquad (28)
\end{aligned}$$

Now replace in (28) the terms $H_N(A^0_{t_{m_i}-N/2+1+\cdot,T}(\beta)h(\frac{\cdot}{N}),-\lambda_k-\beta)$ by $A(u_{m_i},\beta)H_N(-\lambda_k-\beta)$ to get

$$\begin{aligned}
V_{3,T} &= \frac{b}{MN^2H_N^4}\sum_{ip}\sum_{m_1=1}^{M}\sum_{m_2=1}^{M}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,s,k,l}\frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)} \\
&\quad \int_{\Pi^{8-m}} A(u_{m_1},\beta_{a_1})H_N(\lambda_j-\beta_{a_1})A(u_{m_1},\beta_{b_1})H_N(-\lambda_j-\beta_{b_1})A(u_{m_2},\beta_{a_2})H_N(\lambda_k-\beta_{a_2}) \\
&\quad \times A(u_{m_2},\beta_{b_2})H_N(-\lambda_k-\beta_{b_2})A(u_{m_1},\beta_{a_3})H_N(\lambda_s-\beta_{a_3})A(u_{m_1},\beta_{b_3})H_N(-\lambda_s-\beta_{b_3}) \\
&\quad \times H_N(A(u_{m_2},\beta_{a_4})H_N(\lambda_l-\beta_{a_4})A(u_{m_2},\beta_{b_4})H_N(-\lambda_l-\beta_{b_4})\prod_{\nu=1}^{m} g_{|P_\nu|}(\beta_{\overline{P}_\nu}) \\
&\quad \times\exp\{i(t_{m_1}(\beta_{a_1}+\beta_{b_1}+\beta_{a_3}+\beta_{b_3})+t_{m_2}(\beta_{a_2}+\beta_{b_2}+\beta_{a_4}+\beta_{b_4}))\}\,d\beta d\lambda d\mu+E_T,(29)
\end{aligned}$$

where due to the indecomposability of the partitions considered, the following upper bound is true for the error term E_T

$$\frac{b}{MN^2H_N^4}\frac{N}{T}\sum_{ip}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,s,k,l=-J_N}^{J_N}K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l)$$
$$\int_{\Pi^{8-m}}L_N(\lambda_j-\beta_{a_1})L_N(-\lambda_j-\beta_{b_1})L_N(\lambda_k-\beta_{a_2})L_N(-\lambda_k-\beta_{b_2})L_N(\lambda_s-\beta_{a_3})L_N(-\lambda_s-\beta_{b_3})$$
$$\times L_N(\lambda_l-\beta_{a_4})L_N(-\lambda_l-\beta_{b_4})L_M(S(\beta_{a_1}+\beta_{b_1}+\beta_{a_3}+\beta_{b_3}))L_M(S(\beta_{a_2}+\beta_{b_2}+\beta_{a_4}+\beta_{b_4}))d\beta d\lambda d\mu$$
$$\leq \frac{b\log^4(N)}{MN^2}\frac{N}{T}\sum_{ip}\int_{\Pi^{8-m}}L_N(-\beta_{a_1}-\beta_{b_1})L_N(-\beta_{a_2}-\beta_{b_2})L_N(-\beta_{a_3}-\beta_{b_3})L_N(-\beta_{a_4}-\beta_{b_4})$$
$$\times L_M(S(\beta_{a_1}+\beta_{b_1}+\beta_{a_3}+\beta_{b_3}))L_M(S(\beta_{a_2}+\beta_{b_2}+\beta_{a_4}+\beta_{b_4}))d\beta.$$

Therefore, and because $-\beta_{a_i}-\beta_{b_i}\neq 0\ \ \forall i$, $\beta_{a_1}+\beta_{b_1}+\beta_{a_3}+\beta_{b_3}\neq 0$ and $\beta_{a_2}+\beta_{b_2}+\beta_{a_4}+\beta_{b_4}\neq 0$, we get that

$$\frac{b\log^4(N)}{MN^2}\frac{N}{T}\sum_{ip}\int_{\Pi^{8-m}}L_N(-\beta_{a_1}-\beta_{b_1})L_N(-\beta_{a_2}-\beta_{b_2})L_N(-\beta_{a_3}-\beta_{b_3})L_N(-\beta_{a_4}-\beta_{b_4})$$
$$\times L_M(S(\beta_{a_1}+\beta_{b_1}+\beta_{a_3}+\beta_{b_3}))L_M(S(\beta_{a_2}+\beta_{b_2}+\beta_{a_4}+\beta_{b_4}))d\beta$$
$$\leq \frac{b\log^4(N)}{N^2M}\frac{N}{T}\frac{N^4}{S^3}\log^3(M)\log^3(S)\to 0.$$

Similarly the first term on the right hand side of (29) is bounded by

$$\frac{b\log^4(N)}{N^2M}\frac{N^4}{S^3}\log^3(M)\log^3(S)\to 0,$$

which shows that $V_{3,T}\to 0$ as $T\to\infty$. ■

Lemma 5.5. *Under Assumptions 2.1 and 2.2 and if H_0 is true, we have for every $\ell\geq 3$ that*

$$N^\ell M^{\ell/2}h^{\ell/2}cum_\ell(Q_{0,T})=o(1)$$

Proof: Let $\Pi=(-\pi,\pi]$ and $\mu=(\mu_1,\ldots,\mu_\ell)$. We then have

$$N^\ell M^{\ell/2}b^{\ell/2}cum_\ell(Q_{0,T})$$
$$= N^{-\ell}M^{-\ell/2}b^{\ell/2}\sum_{m_1,\ldots,m_l=1}^{M}\sum_{j_{1,1},\ldots,j_{1,\ell}=-J_N}^{J_N}\sum_{j_{2,1},\ldots,j_{2,\ell}=-J_N}^{J_N}\int_{\Pi^l}\prod_{\nu=1}^{\ell}\frac{K_b(\mu_\nu-\lambda_{j_{1,\nu}})K_b(\mu_\nu-\lambda_{j_{2,\nu}})}{f_{\vartheta_0}(u_{m_\nu},\lambda_{j_{2,\nu}})f_{\vartheta_0}(u_{m_\nu},\lambda_{j_{2,\nu}})}$$
$$cum\{\prod_{k=1}^{2}\left(I_N(u_{m_1},\lambda_{j_{k,1}})-f_{\vartheta_0}(u_{m_1},\lambda_{j_{k,1}})\right),\ldots,\prod_{k=1}^{2}\left(I_N(u_{m_\ell},\lambda_{j_{k,\ell})}-f_{\vartheta_0}(u_{m_\ell},\lambda_{j_{k,\ell}})\right)d\mu_1\ldots d\mu_\ell.$$

Using the product theorem for cumulants, we have that

$$cum\{\prod_{k=1}^{2}\left(I_N(u_{m_1},\lambda_{j_{k,1}})-f_{\vartheta_0}(u_{m_1},\lambda_{j_{k,1}})\right),\ldots,\prod_{k=1}^{2}\left(I_N(u_{m_\ell},\lambda_{j_{k,\ell})}-f_{\vartheta_0}(u_{m_\ell},\lambda_{j_{k,\ell}})\right)$$
$$=\sum_{i.p.}\prod_{s=1}^{n}cum\{\left(I_N(u_{m_p},\lambda_{j_{q,p}})-f_{\vartheta_0}(u_{m_p},\lambda_{j_{q,p}})\right),(p,q)\in P_s\}$$

where the sum is over all indecomposable partitions $\{P_1,\ldots,P_n\}$ of the table

$$\begin{matrix}(1,1) & (1,2)\\ \vdots & \vdots\\ (\ell,1) & (\ell,2).\end{matrix}$$

We consider the sum $\sum_{i.p.1}$ over all partitions with $|P_i|>1$. That is,

$$N^{-\ell}H_{2,N}^{-2\ell}(0)M^{-\ell/2}b^{\ell/2}\sum_{i.p.1}\sum_{m_1,\ldots,m_l=1}^{M}\sum_{j_{1,1},\ldots,j_{1,\ell}=-J_N}^{J_N}\sum_{j_{2,1},\ldots,j_{2,\ell}=-J_N}^{J_N}\int_{\Pi^l}\prod_{\nu=1}^{\ell}\frac{K_b(\mu_\nu-\lambda_{j_{1,\nu}})K_b(\mu_\nu-\lambda_{j_{2,\nu}})}{f_{\vartheta_0}(u_{m_\nu},\lambda_{j_{2,\nu}})f_{\vartheta_0}(u_{m_\nu},\lambda_{j_{2,\nu}})}$$
$$\prod_{s=1}^{n}cum\{d_N(u_{m_p},\lambda_{j_{q,p}})d_N(u_{m_p},-\lambda_{j_{q,p}}),(p,q)\in P_s\}d\mu_1\ldots d\mu_\ell \qquad (30)$$

Using the product theorem of cumulants, see Brillinger (1981), we have to sum over all indecomposable partitions $\{Q_{s,1},\ldots,Q_{s,m}\}$ of the table

$$\begin{matrix}a_{p_{s_1},q_{s_1}} & b_{p_{s_1},q_{s_1}}\\ \vdots & \vdots\\ a_{p_{s_{|P_s|}},q_{s_{|P_s|}}} & b_{p_{s_{|P_s|}},q_{s_{|P_s|}}}\end{matrix}$$

for all sets $P_s=\{(p_{s_1},q_{s_1}),\ldots,(p_{s_{|P_s|}},q_{s_{|P_s|}})\}$. Note that $a_{p_{s_r},q_{s_r}}$ and $b_{p_{s_r},q_{s_r}}$ stand for the position of $d_N(u_{m_p^{(r)}},\lambda_{j_{q,p}^{(r)}})$ and $d_N(u_{m_p^{(r)}},-\lambda_{j_{q,p}^{(r)}})$ respectively where (r) denotes the position of $d_N(u_{m_p^{(r)}},-\lambda_{j_{q,p}^{(r)}})$ in a fixed order. For simplicity we use the notation $a_{p_{s_r},q_{s_r}}:=a_{s,r}$ and $b_{p_{s_r},q_{s_r}}:=b_{s,r}$. Furthermore, if $Q_{s,i}=\{c_{s,1},\ldots,c_{s,k}\}$ we set $\overline{Q}_{s,i}=\{c_{s,1},\ldots,c_{s,k-1}\}$, $\beta_{\overline{Q}_{s,i}}:=(\beta_{c_{s,1}},\ldots,\beta_{c_{s,k-1}})$, $\beta_{c_{s,k}}=-\sum_{j=1}^{k-1}\beta_{c_{s,j}}$ and $\beta^{(s)}:=(\beta_{\overline{Q}_{s,1}},\ldots,\beta_{\overline{Q}_{s,m}})$. We then get that (30) is equal to

$$\left(\frac{b}{N^2 H_{2,N}^4(0) M}\right)^{\ell/2} \sum_{i.p.1}\sum_{i.p.^*}\sum_{m_1,\ldots,m_l=1}^{M}\sum_{j_{1,1},\ldots,j_{1,\ell}}\sum_{j_{2,1},\ldots,j_{2,\ell}}\int_{\Pi^l}\prod_{\nu=1}^{\ell}\frac{K_b(\mu_\nu-\lambda_{j_{1,\nu}})K_b(\mu_\nu-\lambda_{j_{2,\nu}})}{f_{\vartheta_0}(u_{m_\nu},\lambda_{j_{2,\nu}})f_{\vartheta_0}(u_{m_\nu},\lambda_{j_{2,\nu}})}$$
$$\times\prod_{s=1}^{n}\int_{\Pi^{2|P_s|-k}}\Bigg\{\prod_{\substack{r=1\\(p,q)\in P_s}}^{|P_s|} H_N(A^0_{t^{(r)}_{m_p}-N/2+1+\cdot,T}(\beta_{a_{s,r}})h(\frac{\cdot}{N}),\lambda_{j^{(r)}_{q,p}}-\beta_{a_{s,r}})$$
$$\times H_N(A^0_{t^{(r)}_{m_p}-N/2+1+\cdot,T}(\beta_{b_{s,r}})h(\frac{\cdot}{N}),-\lambda_{j^{(r)}_{q,p}}-\beta_{b_{s,r}})\Bigg\}\Big\{\prod_{r=1}^{m} g_{|Q_{s,r}|}(\beta_{\overline{|Q_{s,r}|}})\Big\}$$
$$\times\exp\left\{i\sum_{r=1}^{|P_s|}t^{(r)}_{m_p}(\beta_{a_{s,r}}+\beta_{b_{s,r}})\right\}d\beta^{(1)}\ldots d\beta^{(n)}d\mu_1\ldots d\mu_\ell$$

Replace the terms $H_N(A^0_{t^{(r)}_{m_p}-N/2+1+\cdot,T}(\beta)h(\frac{\cdot}{N}),\lambda-\beta)$ by the terms $A(u^{(r)}_{m_p},\beta)H_N(\lambda-\beta)$ to get that the above expression is equal to

$$\left(\frac{b}{N^2 H_{2,N}^4(0) M}\right)^{\ell/2} \sum_{i.p.1}\sum_{i.p.^*}\sum_{m_1,\ldots,m_l=1}^{M}\sum_{j_{1,1},\ldots,j_{1,\ell}}\sum_{j_{2,1},\ldots,j_{2,\ell}}\int_{\Pi^l}\prod_{\nu=1}^{\ell}\frac{K_b(\mu_\nu-\lambda_{j_{1,\nu}})K_b(\mu_\nu-\lambda_{j_{2,\nu}})}{f_{\vartheta_0}(u_{m_\nu},\lambda_{j_{2,\nu}})f_{\vartheta_0}(u_{m_\nu},\lambda_{j_{2,\nu}})}$$
$$\times\prod_{s=1}^{n}\int_{\Pi^{2|P_s|-k}}\Bigg\{\prod_{\substack{r=1\\(p,q)\in P_s}}^{|P_s|} A(u^{(r)}_{m_p},\beta_{a_{s,r}})H_N(\lambda_{j^{(r)}_{q,p}}-\beta_{a_{s,r}})A(u^{(r)}_{m_p},\beta_{b_{s,r}})H_N(-\lambda_{j^{(r)}_{q,p}}-\beta_{b_{s,r}})\Bigg\}$$
$$\times\Big\{\prod_{r=1}^{m} g_{|Q_{s,r}|}(\beta_{\overline{|Q_{s,r}|}})\Big\}\exp\left\{i\sum_{r=1}^{|P_s|}t^{(r)}_{m_p}(\beta_{a_{s,r}}+\beta_{b_{s,r}})\right\}d\beta^{(1)}\ldots d\beta^{(n)}d\mu_1\ldots d\mu_\ell+E_T \quad (31)$$

where the error term E_T is bounded by

$$\left(\frac{b}{N^2 H_{2,N}^4(0) M}\right)^{\ell/2}\frac{N}{T}\sum_{i.p.1}\sum_{i.p.^*}\sum_{j_{1,1},\ldots,j_{1,\ell}=-J_N}^{J_N}\sum_{j_{2,1},\ldots,j_{2,\ell}=-J_N}^{J_N}\int_{\Pi^l}\prod_{\nu=1}^{\ell}K_b(\mu_\nu-\lambda_{j_{1,\nu}})K_b(\mu_\nu-\lambda_{j_{2,\nu}})$$
$$\prod_{s=1}^{n}\Bigg\{\prod_{\substack{r=1\\(p,q)\in P_s}}^{|P_s|}\int_{\Pi^{2|P_s|-m}}L_N(\lambda_{j^{(r)}_{q,p}}-\beta_{a_{s,r}})L_N(-\lambda_{j^{(r)}_{q,p}}-\beta_{b_{s,r}})L_M(S(\beta_{a_{s,r}}+\beta_{b_{s,r}}+\beta_{a_{x,y}}+\beta_{b_{x,y}}))\Bigg\}$$
$$d\beta^{(1)}\ldots d\beta^{(n)}d\mu_1\ldots d\mu_\ell$$

for some $(x, y) \in \{1, \ldots, n\} \times \{1, \ldots, |P_x|\}$ with $x \neq s$. Integration over all $\beta_{a_{s,r}}$ and $\beta_{b_{s,r}}$ gives that expression (31) is bounded by

$$\left(\frac{b}{N^2 H_{2,N}^4(0) M}\right)^{\ell/2} \frac{N}{T} \frac{M^{\ell-1} N}{S} N^{2\ell} \to 0,$$

which completes the proof. ■

Lemma 5.6. *Under Assumptions 2.1 and 2.2 and if H_0 is true, we have that*

$$N\sqrt{Mb}(Q_T - \mu_T) = N\sqrt{Mb}(Q_{0,T} - \mu_T) + o_p(1)$$

Proof:

$$\begin{aligned} Q_T &= Q_{0,T} + \frac{1}{MN^2} \sum_{i=1}^{M} \int_{-\pi}^{\pi} \left\{ \sum_{j=-J_N}^{J_N} K_b(\lambda - \lambda_j) \left(\frac{I_N(u_i, \lambda_j)}{f_{\hat{\vartheta}}(u_i, \lambda_j)} - \frac{I_N(u_i, \lambda_j)}{f_{\vartheta_0}(u_i, \lambda_j)} \right) \right\}^2 d\lambda \\ &+ \frac{2b^{1/2}}{NM^{1/2}} \sum_{i=1}^{M} \int_{-\pi}^{\pi} \sum_{j,s} K_b(\lambda - \lambda_j) K_b(\lambda - \lambda_s) \left(\frac{I_N(u_i, \lambda_j)}{f_{\hat{\vartheta}}(u_i, \lambda_j)} - \frac{I_N(u_i, \lambda_j)}{f_{\vartheta_0}(u_i, \lambda_j)} \right) \left(\frac{I_N(u_i, \lambda_s)}{f_{\vartheta_0}(u_i, \lambda_s)} - 1 \right) d\lambda \\ &= Q_{0,T} + Y_{1,T} + Y_{2,T} \end{aligned}$$

with an obvious notation for $Y_{1,T}$ and $Y_{2,T}$. The term $Y_{1,T}$ is bounded by

$$\begin{aligned} |Y_{1,T}| &\leq \sup_{u,j} \left(\frac{f_{\vartheta_0}(u, \lambda_j) - f_{\hat{\vartheta}}(u, \lambda_j)}{f_{\hat{\vartheta}}(u, \lambda_j)} \right)^2 \frac{b^{1/2}}{M^{1/2} N} \sum_{i=1}^{M} \int_{-\pi}^{\pi} \left\{ \sum_j K_b(\lambda - \lambda_j) \frac{I_N(u_i, \lambda_j)}{f_{\vartheta_0}(u_i, \lambda_j)} \right\}^2 d\lambda \\ &= O_p\left(\frac{b^{1/2}}{M^{1/2}} \right). \end{aligned}$$

For the second term we have

$$\begin{aligned} Y_{2,T} &= \frac{b^{1/2}}{NM^{1/2}} \sum_{i=1}^{M} \int_{-\pi}^{\pi} \sum_{j,s} K_b(\lambda - \lambda_j) K_b(\lambda - \lambda_s) \left(\frac{I_N(u_i, \lambda_j)}{f_{\vartheta_0}(u_i, \lambda_j)} - 1 \right) \left(\frac{f_{\hat{\vartheta}}(u_i, \lambda_j) - f_{\vartheta_0}(u_i, \lambda_j)}{f_{\hat{\vartheta}}(u_i, \lambda_j)} \right) \\ &\times \left(\frac{I_N(u_i, \lambda_s)}{f_{\vartheta_0}(u_i, \lambda_s)} - 1 \right) d\lambda \\ &+ \frac{b^{1/2}}{NM^{1/2}} \sum_{i=1}^{M} \int_{-\pi}^{\pi} \sum_{j,s} K_b(\lambda - \lambda_j) K_b(\lambda - \lambda_s) \left(\frac{f_{\hat{\vartheta}}(u_i, \lambda_j) - f_{\vartheta_0}(u_i, \lambda_j)}{f_{\hat{\vartheta}}(u_i, \lambda_j)} \right) \left(\frac{I_N(u_i, \lambda_s)}{f_{\vartheta_0}(u_i, \lambda_s)} - 1 \right) d\lambda \\ &= W_{1,T} + W_{2,T} \end{aligned}$$

with an obvious notation for $W_{1,T}$ and $W_{2,T}$.

By a standard Taylor series argument, for fixed u, we have that for $\hat{\vartheta}(u) = (\hat{\vartheta}_1(u), \hat{\vartheta}_2(u), \ldots, \hat{\vartheta}_p(u))'$ and $\vartheta_0(u) = (\vartheta_1(u), \vartheta_2(u), \ldots, \vartheta_p(u))'$, $\tilde{\vartheta}_0(u) = (\tilde{\vartheta}_1(u), \tilde{\vartheta}_2(u), \ldots, \tilde{\vartheta}_p(u))'$ with $||\tilde{\vartheta}(u) - \vartheta_0(u)|| \leq ||\hat{\vartheta}(u) - \vartheta_0(u)||$ exists such that

$$\frac{f_{\hat{\vartheta}}(u,\lambda) - f_{\vartheta_0}(u,\lambda)}{f_{\hat{\vartheta}}(u,\lambda)} =$$
$$O_p(1)\left\{\sum_{m=1}^{p}(\hat{\vartheta}_m(u) - \vartheta_m(u)) f_T^{(1)}(\vartheta_m, \lambda) + \frac{1}{2}\sum_{m=1}^{p}\sum_{l=1}^{p}(\hat{\vartheta}_m(u) - \vartheta_m(u))(\hat{\vartheta}_l(u) - \vartheta_l(u)) f_T^{(2)}(\tilde{\vartheta}_m, \tilde{\vartheta}_l, \lambda)\right\} \tag{32}$$

where $f_T^{(1)}(\vartheta_m, \lambda)$ and $f_T^{(2)}(\tilde{\vartheta}_m, \tilde{\vartheta}_l, \lambda)$ denote the first and second second partial derivatives of f with respect to ϑ_m and ϑ_l and ϑ_m respectively, and evaluated at ϑ_m and $(\tilde{\vartheta}_m, \tilde{\vartheta}_l)$. Notice that the $O_p(1)$ term appear in (32) is due to the fact that $|1/f_{\hat{\vartheta}}(u,\lambda)| = O_p(1)$. Using (32) we get

$$\begin{aligned}
W_{1,T} &= O_p(1)\frac{b^{1/2}}{NM^{1/2}}\sum_{i=1}^{M}\sum_{m=1}^{p}(\hat{\vartheta}_m(u_i) - \vartheta_m(u_i))\int_{-\pi}^{\pi}\sum_{j,s}K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)f^{(1)}(\vartheta_m,\lambda_j) \\
&\quad \times\left(\frac{I_N(u_i,\lambda_j)}{f_{\vartheta_0}(u_i,\lambda_j)} - 1\right)\left(\frac{I_N(u_i,\lambda_s)}{f_{\vartheta_0}(u_i,\lambda_s)} - 1\right)d\lambda \\
&\quad +O_p(1)\frac{b^{1/2}}{NM^{1/2}}\sum_{i=1}^{M}\sum_{l=1}^{p}\sum_{m=1}^{p}(\hat{\vartheta}_l(u_i) - \vartheta_l(u_i))(\hat{\vartheta}_m(u_i) - \vartheta_m(u_i))\int_{-\pi}^{\pi}\sum_{j,s}K_b(\lambda-\lambda_j) \\
&\quad \times K_b(\lambda-\lambda_s)f^{(2)}(\tilde{\vartheta}_m,\tilde{\vartheta}_l,\lambda_j)\left(\frac{I_N(u_i,\lambda_j)}{f_{\vartheta_0}(u_i,\lambda_j)} - 1\right)\left(\frac{I_N(u_i,\lambda_s)}{f_{\vartheta_0}(u_i,\lambda_s)} - 1\right)d\lambda \\
&= O_p(N^{-1/2}) + O_p(b^{1/2}).
\end{aligned}$$

The $O_p(N^{-1/2})$ term is due to the fact that $\sup_u |\hat{\vartheta}_m(u) - \vartheta_m(u)| = O_p(N^{-1/2})$ and that

$$\frac{b^{1/2}}{NM^{1/2}}\sum_{i=1}^{M}\int_{-\pi}^{\pi}\sum_{j,s}K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)f^{(1)}(\vartheta_m,\lambda_j)\left(\frac{I_N(u_i,\lambda_j)}{f_{\vartheta_0}(u_i,\lambda_j)} - 1\right)\left(\frac{I_N(u_i,\lambda_s)}{f_{\vartheta_0}(u_i,\lambda_s)} - 1\right)d\lambda$$

can be handled as $Q_{0,T}$. Similarly we can show that $W_{2,T} = o_p(1)$ which completes the proof. ■

Proof of Theorem 2.1: By Lemma 5.3, 5.4 and 5.5 we have that the cumulants of all orders of $Q_{0,T}$ converge to the corresponding cumulants of the limiting Gaussian distribution. The assertion of the theorem follows then by Lemma 5.6. ■

Proof of Theorem 3.1 : Follow the same steps as in the proof of Lemma 5.6 substituting $\overline{\vartheta}$ for ϑ_0 in $f_{\vartheta_0}(u_i, \lambda_j)$ and using the property that under the alternative hypothesis, $\hat{\vartheta}$ is a $\sqrt{N}$-consistent estimator of $\overline{\vartheta}$. ■

Proof of Theorem 3.2: First notice that for $T \geq T_0$ and by Theorem 2.3 of Dahlhaus (1996), $\{X^+_{t,T}\}$ is locally stationary with

$$X^+_{t,T} = \int_{-\pi}^{\pi} \hat{A}^0_{t,T}(\lambda) e^{i\lambda t} d\xi^+(\lambda) \tag{33}$$

where

(i) $\xi^+(\lambda)$ is a Gaussian stochastic process on $(-\pi, \pi]$ and

$$cov^+\{\xi^+(\lambda_k), \xi^+(\lambda_j)\} = \delta(k,j) d\lambda_k \tag{34}$$

(ii) There exists a constant K and a function $\hat{A}(u, \lambda)$ on $[0,1] \times (-\pi, \pi)$ such that for all T,

$$\sup_{t,\lambda} |\hat{A}^0_{t,T} - \hat{A}(t/T, \lambda)| \leq K/T$$

(iii) Furthermore,

$$f_{\hat{\vartheta}(u)}(u, \lambda) = \frac{1}{2\pi} |\hat{A}(u, \lambda)|^2 \tag{35}$$

where $\hat{\vartheta}(u) = (\hat{\beta}_1(u), \ldots, \hat{\beta}_p(u), \hat{\sigma}^2(u))$ and the function $1/f(u, \lambda; \hat{\vartheta})$ is bounded in probability.

Now, following the same steps as in the proof of Lemma 5.3, 5.4 and 5.5, we get that the limits of all cumulants of the bootstrap test statistic $N\sqrt{Mb}(Q^+_T - \mu_T)$ converge to the cumulants of the limiting Gaussian distribution given in Theorem 3.2. ■

Acknowledgments. The authors are very grateful to the Editor, the Associate Editor and two referees for teir valuable and insightful comments leading to a considerable improvement of the paper.

References

Beltrão, K. I. and Bloomfield, P. (1987). Determining the bandwidth of a kernel spectrum estimate. *Journal of Time Series Analysis*, **8**, 21-38.

Bickel, P. J. and Freedman, D. A. (1981). Some Asymptotic Theory for the Bootstrap. *Annals of Statistics*, **9**, 1196-1217.

Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory.* Holden-Day, San Francisco.

Chang, C. and Morettin, P. (1999). Estimation of time-varying linear systems. *Statist. Inference Stoch. Proc.*, **2**, 253-285.

Dahlhaus, R. (1996). On the Kullback-Leibler information diverenge of locally stationary processes. *Stochastic Processes and their Applications*, **62**, 139-168.

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *The Annals of Statistics*, **25**, 1-37.

Dahlhaus, R. (2000). A likelihood approximation for locally stationary processes. *The Annals of Statistics*, **28**, 1762-1794.

Dahlhaus, R. (2003). Curve estimation for locally stationary time series models. In: *Recent Advances and Trends in Nonparametric Statistics* (ed. by M. G. Akritas and D. N. Politis) Elservier Science B.V., 451-466.

Dahlhaus, R. and Giraitis, L. (1998). On optimal segment length for parameter estimators for locally stationary time series. *Journal of Time Series Analysis*, **19**, 629-655.

Dahlhaus, R., Neumann, M.H. and von Sachs, R. (1999). Nonlinear wavelet estimation of time-varying autoregressive processes . *Bernoulli*, **5**, 873-906.

Dahlhaus, R. and Neumann, M.H.(2001). Locally adaptive fitting of semiparametric models to nonstationary time series. *Stochastic Processes and their Applications*, **91**, 277-308.

Dahlhaus, R. and Polonik, W. (2006). Nonparametric quasi-maximum likelihood estimation for Gaussian locally stationary processes. *Annals of Statistics*, **34**, 2790-2824.

Davis, R. A., Lee, T. C. M. and Rodriguez-Yam, G. A. (2006). Structural break estimation for nonstationary time series models. *J. Am. Statist. Assoc.* **101**, 223-239.

Fryzlewicz, P. Van Bellegem and von Sachs, R.(2003) Forecasting non-stationary time series by wavelet process modelling. *Ann. Inst. Statist. Math.* **55**, 737-764.

Nason, G. P., von Sachs, R. and Kroisandt, G. (2000). Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. *J. R. Stat. Soc. Ser. B Stat. Methol.*, **62**, 271-292.

Neumann, M. and von Sachs, R (1997). Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra. *The Annals of Statistics*, **25**, 38-76.

Ombao, H., Raz, J., von Sachs, R. and Guo, W. (2002) The SLEX model of a non-stationary random process. Nonparametric approach to time series analysis. *Ann. Inst. Statist. Math.* **54**, 171-200.

Ombao, H., von Sachs, R. and Guo, W. (2005) SLEX analysis of multivariate nonstationary time series. *J. Amer. Stat. Assoc.* **100**, 519-531.

Priestley, M. B. (1965). Evolutionary spectra and non-stationary processes. *J. R. Stat. Soc. Ser. B*, **62**, 204-237.

Priestley, M. B. (1981). *Spectral Analysis and Time Series.* Academic Press, New York.

Sakiyama, S. and Taniguchi, M. (2003). Testing composite hypotheses for locally stationary processes. *Journal of Time Series Analysis*, **24**, 483-504.

Sergides, M. and Paparoditis, E. (2008). Bootstrapping the local periodogram of locally stationary processes. *Journal of Time Series Analysis*,29, 264-299.

Van Bellegem, S. and Dahlhaus, R. (2006). Semiparametric estimation by model selection for locally stationary processes. *J. R. Stat. Soc. Ser. B Stat. Methol.*, **68**, 721-746.

Van Bellegem, S. and von Sachs, R. (2008). Locally adaptive estimation of evolutionary wavelet spectra. *Annals of Statistics*, **36**, 1879-1924.

	-a-						-b-					
	$b = 0.3$			$b = 0.2$			$b = 0.3$			$b = 0.2$		
T=512, N=64	$a = 0.01$	$a = 0.05$	$a = 0.1$	$a = 0.01$	$a = 0.05$	$a = 0.1$	$a = 0.01$	$a = 0.05$	$a = 0.1$	$a = 0.01$	$a = 0.05$	$a = 0.1$
$\phi_2 = 0.0$												
$\kappa = 1$	0.008	0.038	0.084	0.008	0.042	0.078	0.008	0.040	0.074	0.006	0.034	0.072
$\kappa = 2$	0.008	0.032	0.070	0.008	0.032	0.068	0.010	0.038	0.082	0.008	0.036	0.076
$\phi_2 = 0.2$												
$\kappa = 1$	0.072	0.204	0.358	0.098	0.244	0.364	0.074	0.218	0.342	0.078	0.208	0.350
$\kappa = 2$	0.200	0.296	0.408	0.174	0.354	0.474	0.214	0.334	0.464	0.196	0.374	0.494
$\phi_2 = 0.25$												
$\kappa = 1$	0.188	0.410	0.582	0.238	0.482	0.612	0.192	0.420	0.562	0.200	0.424	0.592
$\kappa = 2$	0.408	0.564	0.680	0.422	0.640	0.744	0.434	0.596	0.708	0.440	0.650	0.768
$\phi_2 = 0.3$												
$\kappa = 1$	0.392	0.668	0.814	0.474	0.752	0.852	0.394	0.690	0.808	0.428	0.714	0.832
$\kappa = 2$	0.692	0.820	0.898	0.738	0.882	0.936	0.714	0.844	0.912	0.756	0.892	0.942
	$b = 0.3$			$b = 0.2$			$b = 0.3$			$b = 0.2$		
T=1024, N=128	$a = 0.01$	$a = 0.05$	$a = 0.1$	$a = 0.01$	$a = 0.05$	$a = 0.1$	$a = 0.01$	$a = 0.05$	$a = 0.1$	$a = 0.01$	$a = 0.05$	$a = 0.1$
$\phi_2 = 0.0$												
$\kappa = 1$	0.010	0.040	0.080	0.012	0.044	0.082	0.008	0.034	0.084	0.012	0.040	0.086
$\kappa = 2$	0.008	0.044	0.092	0.012	0.048	0.098	0.010	0.038	0.086	0.012	0.044	0.096
$\phi_2 = 0.2$												
$\kappa = 1$	0.272	0.512	0.618	0.226	0.474	0.584	0.256	0.492	0.638	0.222	0.466	0.590
$\kappa = 2$	0.480	0.732	0.800	0.500	0.674	0.776	0.502	0.696	0.788	0.518	0.658	0.776
$\phi_2 = 0.25$												
$\kappa = 1$	0.622	0.830	0.894	0.542	0.798	0.866	0.606	0.818	0.912	0.540	0.792	0.874
$\kappa = 2$	0.824	0.952	0.974	0.834	0.924	0.968	0.832	0.944	0.974	0.838	0.920	0.968
$\phi_2 = 0.3$												
$\kappa = 1$	0.914	0.992	0.996	0.884	0.974	0.992	0.904	0.986	0.996	0.876	0.972	0.992
$\kappa = 2$	0.986	1.000	1.000	0.986	1.000	1.000	0.986	1.000	1.000	0.988	1.000	1.000

TABLE 1. Rejection frequencies in 500 replications of the tvAR(2) model $X_{t,T} = 0.9\cos(1.5 - \cos(4\pi t/T))X_{t-1,T} - \phi_2 X_{t-2,T} + \varepsilon_t$ for different values of ϕ_2 and of the testing parameters. In Part a) rescaling is done using $f(u_i, \lambda_j; \hat{\vartheta})$ while in Part b) using $f(u_i, \lambda_j; \hat{\vartheta}^+)$.

	$b = 0.3$			$b = 0.2$		
T=512, N=64	$a = 0.01$	$a = 0.05$	$a = 0.1$	$a = 0.01$	$a = 0.05$	$a = 0.1$
$\kappa = 1$	0.726	0.842	0.904	0.720	0.868	0.902
$\kappa = 2$	0.766	0.874	0.936	0.740	0.882	0.934
	$b = 0.2$			$b = 0.1$		
T=1024, N=128	$a = 0.01$	$a = 0.05$	$a = 0.1$	$a = 0.01$	$a = 0.05$	$a = 0.1$
$\kappa = 1$	0.832	0.870	0.906	0.838	0.886	0.908
$\kappa = 2$	0.924	0.972	0.986	0.954	0.990	0.996

TABLE 2. Rejection frequencies in 500 replications of the model $X_{t,T} = 0.9X_{t-1,T} + \varepsilon_t$ for $1 \leq t \leq [T/2]$ and $X_{t,T} = -0.5X_{t-1,T} + \varepsilon_t$ for $[T/2] + 1 \leq t \leq T$.

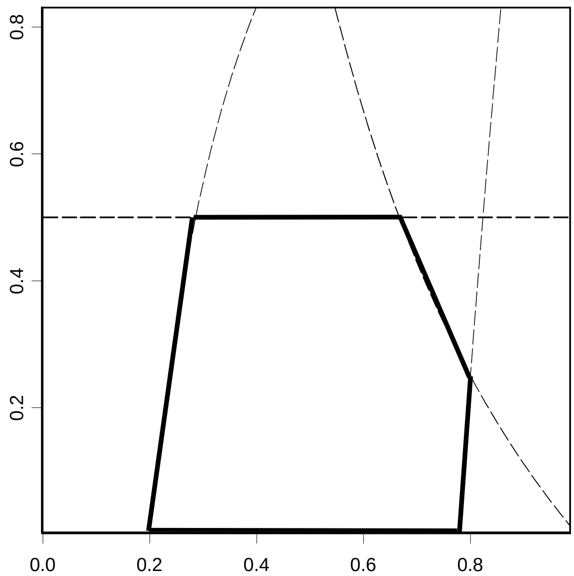
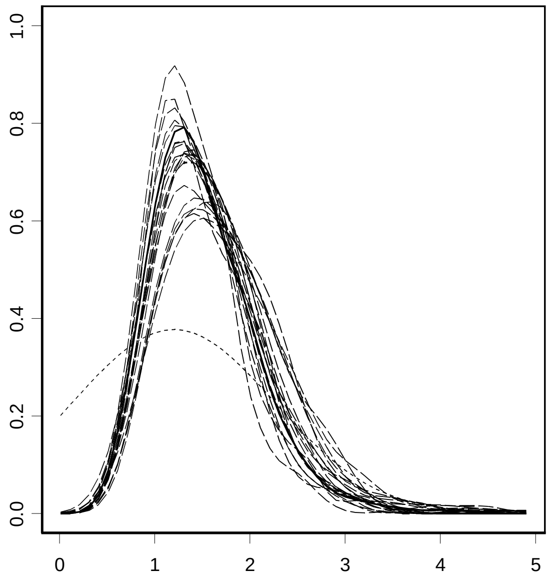
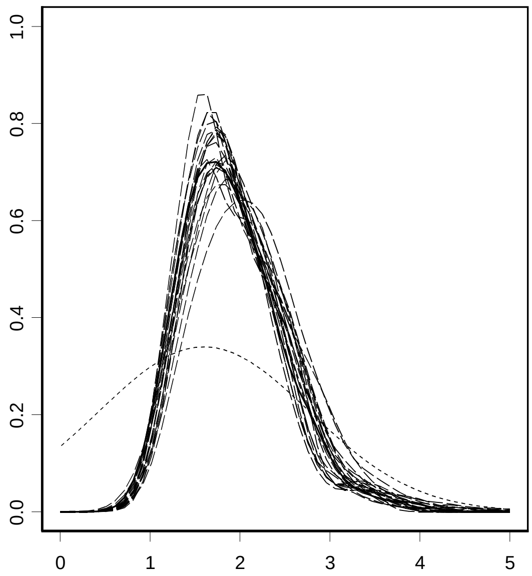


FIGURE 1. The area within the bold marked lines presents the possible range of values of the parameter λ, (y-axis) and δ, (x-axis), according to Assumption 2.2 (iii).



(a) $\kappa = 1$



(b) $\kappa = 2$

FIGURE 2. Estimated densities of the distribution of the test statistic Q_T under the null hypothesis of a first order tvAR process and its different approximations. The solid lines in (a) and (b) are the estimated exact densities, the dashed lines are the estimated densities of the bootstrap approximations while the dotted lines are the densities of the asymptotic Gaussian approximations.

University of Cyprus, Department of Mathematics and Statistics, P.O.Box 20537, CY-1678 Nicosia, Cyprus

University of Cyprus, Department of Mathematics and Statistics, P.O.Box 20537, CY-1678 Nicosia, Cyprus