

# Estimation of the Bispectrum for Locally Stationary Processes

Efstathios Paparoditis<sup>1</sup>, Philip Preuß<sup>2</sup>

<sup>1</sup>*Department of Mathematics and Statistics, University of Cyprus, P.O.Box 20537, CY 1678 Nicosia, CYPRUS*

<sup>2</sup>*Department of Mathematics, Ruhr University Bochum, 44780 Bochum, Germany*

email: stathisp@ucy.ac.cy

email: philip.preuss@ruhr-uni-bochum.de

December 6, 2013

## Abstract

We consider the problem of estimating the bispectrum of a locally stationary process. A non-parametric, lag-window type estimator is considered and its asymptotic properties are investigated. As a possible application, a test for linearity in the framework of locally stationary processes is discussed.

AMS subject classification: 62M10, 62G10

Keywords and phrases: bispectrum, non stationary processes, locally stationary process, linearity test

## 1 Introduction

Although the assumption of covariance-stationarity is extremely popular in the analysis of discrete time series (because it allows for an elegant asymptotic theory), it is usually not justified in practise. In the reality most processes change their second-order structure over time and one concept, which incorporates this feature, is that of locally stationary processes. A triangular array  $\{\mathbf{X}_T, T = 1, 2, \dots\}$  of stochastic processes  $\mathbf{X}_T = \{X_{t,T}, t = 1, 2, \dots, T\}$  is called locally stationary if it possesses a time varying  $MA(\infty)$ -representation

$$(1) \quad X_{t,T} = \sum_{l=0}^{\infty} \psi_{t,T,l} Z_{t-l}, \quad t = 1, \dots, T,$$

where the random variables  $Z_t$  are assumed to be independent and identically distributed with mean 0 and variance 1. To make the class  $\mathbf{X}_T$  of processes mathematically tractable, it is commonly assumed that  $X_{t,T}$  can be locally approximated by a stationary process, that is that there exist functions  $\psi_l :$

$[0, 1] \rightarrow \mathbb{R}$  such that the time varying coefficients  $\psi_{t,T,l}$  are close enough to some functions  $\psi_l(t/T)$ , i.e., that

$$(2) \quad \sum_{l=0}^{\infty} \sup_{t=1, \dots, T} |\psi_{t,T,l} - \psi_l(t/T)| = O(1/T).$$

The process

$$X_t(u) := \sum_{l=-\infty}^{\infty} \psi_l(u) Z_{t-l}$$

which is obtained by replacing  $\psi_{t,T,l}$  by  $\psi_l(t/T)$  is called the locally approximating stationary process, that is the stationary process that approximates  $X_{t,T}$  locally at the rescaled time point  $t/T$ . The above class of stochastic processes was introduced by Dahlhaus (1997) and became quite popular in recent years.

Consider now a locally stationary process as defined in (1) and assume that all moments of the innovations  $Z_t$  exist and that the locally approximating functions  $\psi_l : [0, 1] \rightarrow \mathbb{R}$  from (2) are twice continuously differentiable and satisfy the following conditions,

$$(3) \quad \sum_{l=-\infty}^{\infty} \sup_{u \in [0,1]} |\psi_l(u)| |l|^2 < \infty,$$

$$(4) \quad \sum_{l=-\infty}^{\infty} \sup_{u \in [0,1]} |\psi'_l(u)| |l| < \infty,$$

$$(5) \quad \sum_{l=-\infty}^{\infty} \sup_{u \in [0,1]} |\psi''_l(u)| < \infty.$$

Then a time varying spectral density of  $\mathbf{X}_T$  exists, is unique and it is given by

$$f(u, \lambda) := \frac{1}{2\pi} \sum_{l=-\infty}^{\infty} \mathbb{E}(X_t(u) X_{t+l}(u)) \exp(-i\lambda l), \quad u \in [0, 1], \quad \lambda \in \mathbb{R},$$

[see Dahlhaus (1997)]. Notice that under the assumptions made

$$\sup_{t, \lambda} \left| (2\pi)^{-1} |\Psi_{t,T}(e^{-i\lambda})|^2 - f(u, \lambda) \right| = O(T^{-1}),$$

where for  $z \in C$ ,  $\Psi_{t,T}(z) = \sum_l \psi_{t,T,l} z^l$ . Similarly, we define the time varying bispectrum of the locally stationary process  $\mathbf{X}$  through

$$f(u, \lambda_1, \lambda_2) := \frac{1}{(2\pi)^2} \sum_{l_1, l_2=-\infty}^{\infty} \mathbb{E}(X_t(u) X_{t+l_1}(u) X_{t+l_2}(u)) \exp(-i(\lambda_1 l_1 + \lambda_2 l_2)).$$

This definition is justified by the following, easily established property

$$\sup_{t, \lambda_1, \lambda_2} \left| (2\pi)^{-2} E(Z_1^3) \Psi_{t,T}(e^{-i\lambda_1}) \Psi_{t,T}(e^{-i\lambda_2}) \Psi_{t,T}(e^{-i(\lambda_1+\lambda_2)}) - f(u, \lambda_1, \lambda_2) \right| = O(T^{-1}).$$

In this note we consider the problem of estimating the bispectrum  $f(u, \lambda_1, \lambda_2)$  based on an observed time series  $X_{1,T}, X_{2,T}, \dots, X_{T,T}$  of the locally stationary process. For this, a lag-window estimator is considered and its asymptotic properties are derived. As an application, a test of linearity is discussed.

## 2 A Lag-Window Estimator and its Properties

In order to estimate the bispectrum  $f(u, \lambda_1, \lambda_2)$ , we choose a window length  $N$  such that  $N = o(T)$  holds, and define an estimator of the time-varying bispectrum through

$$(6) \quad \hat{f}(u, \lambda_1, \lambda_2) = \frac{1}{(2\pi)^2} \sum_{k_1, k_2 = -\infty}^{\infty} w(k_1/B, k_2/B) \hat{\gamma}_N(u, k_1, k_2) \exp(-i\lambda_{1,N}k_1 - i\lambda_{2,N}k_2),$$

where  $w : \mathbb{R}^2 \rightarrow \mathbb{R}_0^+$  is a continuously differentiable function with compact support,  $\lambda_N = 2\pi N^{-1} \lfloor (2\pi)^{-1} N \lambda \rfloor$  and

$$\hat{\gamma}_N(u, k_1, k_2) := \frac{1}{N} \sum_{p=0}^{N-1} X_{[uT]-N/2+1+p} X_{[uT]-N/2+1+p+k_1} X_{[uT]-N/2+1+p+k_2},$$

[we have set  $X_{t,T} = 0$  for  $t \notin \{1, \dots, T\}$ ]. We assume that  $N, B \rightarrow \infty$  and state the following assumptions on the weight function  $w$ .

**Assumption 2.1**  $w(\cdot, \cdot)$  has the compact support  $[0, 1]^2$  with  $w(0, 0)$  being equal to one, is twice continuously differentiable, and the first derivatives vanish at the origin.

The following theorem summarizes the asymptotic properties of the estimator (6).

**Theorem 2.2** Let Assumption 2.1 be true, assume that  $\inf_{u, \lambda} f(u, \lambda) \geq c > 0$  and that

$$N/B^2 \rightarrow \infty, \quad N^{5/2}/(BT^2) \rightarrow 0$$

as  $T \rightarrow \infty$ . . Then,

$$(i) \quad E(\hat{f}(u, \lambda_1, \lambda_2)) = f(u, \lambda_{1,N}, \lambda_{2,N}) + O(1/B^2 + N^2/T^2 + 1/T),$$

$$(ii) \quad \sqrt{N}B^{-1} \text{Var}(\hat{f}(u, \lambda_1, \lambda_2)) = \sigma_B f(u, \lambda_1) f(u, \lambda_2) f(u, \lambda_1 + \lambda_2) + o(1),$$

and

$$(iii) \quad \sqrt{N}B^{-1}(\hat{f}(u, \lambda_1, \lambda_2) - f(u, \lambda_{1,N}, \lambda_{2,N})) \xrightarrow{D} N(0, \sigma_B f(u, \lambda_1) f(u, \lambda_2) f(u, \lambda_1 + \lambda_2))$$

with  $\sigma_B := (2\pi)^{-1} \omega_B$  and  $\omega_B := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w^2(x, y) dy dx$ .

As it is seen from this theorem, the lag window estimator (6) is a mean square consistent estimator of the bispectrum  $f(u, \lambda_1, \lambda_2)$  which appropriate centered has an asymptotic Gaussian distribution.

### 3 A Test for Linearity

It is well known that, under the assumption of covariance stationarity, a linear model is very often not able to capture all stylized facts of a given time series dataset. If we look, for example, at financial time series, then days with either small or large movements are likely to be followed by days with similar fluctuation. This effect, called 'volatility clustering', can not be explained by a standard linear model like, for instance, an ARMA( $p, q$ ) one. One example to resolve this issue (which became particular famous during the last decades) is to employ GARCH( $p, q$ ) processes which are stationary under certain regularity conditions, but do not lie in the class of linear models anymore. It is obvious, however, that the effect of clustered volatility also occurs if we employ the simple locally stationary process

$$(7) \quad X_{t,T} = \sigma(t/T)Z_t \quad t = 1, \dots, T,$$

where  $\sigma(\cdot) : [0, 1] \rightarrow \mathbb{R}_+$  is a non parametric function and  $Z_t$  is some independent and identically distributed random variable. In fact, many authors have pointed out that most of the stylized facts which are observed for financial return data can be explained by using the model (7); see for example Starica and Granger (2005), Fryzlewicz et al. (2006). The local bispectrum discussed in the previous section can be used to develop a test for linearity.

To motivate our proposal, recall that every zero mean covariance-stationary time series  $X_t$  with a purely continuous spectral density can be represented as

$$(8) \quad X_t = \sum_{l=0}^{\infty} \psi_l Z_{t-l}$$

where  $Z_t$  is a white noise process (see Priestley (1983), Chapter 10). If the  $Z_t$  in (8) are independent and identically distributed, the process is called linear. If the underlying process is linear then the time varying bispectrum is equal to

$$\frac{\mathbb{E}(Z^3)}{(2\pi)^2} H(u, -\lambda_1 - \lambda_2) H(u, \lambda_1) H(u, \lambda_2)$$

with

$$H(u, \lambda) := \sum_{l=-\infty}^{\infty} \psi_l(u) \exp(-i\lambda l).$$

Because  $f(u, \lambda) = \frac{1}{2\pi} |H(u, \lambda)|^2$  we obtain for a linear process that

$$(9) \quad |f(u, \lambda_1, \lambda_2)|^2 = \frac{\mathbb{E}(Z^3)^2}{(2\pi)\sigma^6} f(u, \lambda_1) f(u, \lambda_2) f(u, \lambda_1 + \lambda_2).$$

and

$$(10) \quad |f(u, \lambda_1, \lambda_2)|^2 \equiv 0$$

if the process is Gaussian.

A global measure for linearity/Gaussianity is then obtained by dividing  $X_{1,T}, \dots, X_{T,T}$  into  $M$  segments of length  $N$ , setting  $u_j$  ( $j = 1, \dots, M$ ) as the midpoints of the corresponding interval in rescaled time and considering

$$\frac{1}{M} \sum_{j=1}^M \frac{|\hat{f}(u_j, \lambda_1, \lambda_2)|^2}{\hat{f}(u_j, \lambda_1)\hat{f}(u_j, \lambda_2)\hat{f}(u_j, \lambda_1 + \lambda_2)}$$

where

$$\hat{f}(u, \lambda) := \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} w(k/B) \hat{\gamma}_N(u, k) \exp(-i\lambda_N k),$$

$w : \mathbb{R} \rightarrow \mathbb{R}_0^+$  is a continuously differentiable function with compact support and

$$\hat{\gamma}_N(u, k) := \frac{1}{N} \sum_{q=0}^{N-1-k} X_{[uT]-N/2+1+q} X_{[uT]-N/2+1+q+k}.$$

Then for different  $\lambda_1, \lambda_2$ , [as it is formally proved in Theorem 3.2] this expression converges in probability to  $\int_0^1 |f(u, \lambda_1, \lambda_2)|^2 f^{-1}(u, \lambda_1) f^{-1}(u, \lambda_2) f^{-1}(u, \lambda_1 + \lambda_2) du$  which equals  $(2\pi)^{-1} \mathbb{E}(Z^3)^2$  if the process is linear and equals zero in the Gaussian case [cf. (9), (10)].

To derive the asymptotic distribution of our test statistic, the following set of additional conditions is required.

**Assumption 3.1** *Let  $N \sim T^\alpha$  and  $B \sim T^\beta$  for some  $\alpha, \beta > 0$ . Suppose that one of the following two conditions holds:*

*i) It is  $\mathbb{E}(Z_t^3) = 0$ ,  $\alpha \in (0, 1)$  and*

$$\max(1/2 - \alpha/2, 5/4\alpha - 3/4, 2\alpha/5) < \beta < \alpha/2 \quad \text{or}$$

*ii) It is  $\mathbb{E}(Z_t^3) \neq 0$ ,  $\alpha \in (1/2, 1)$  and*

$$\max(2\alpha - 3/2, 1/6) < \beta < \alpha/2.$$

The range of allowed values for  $(\alpha, \beta)$  is illustrated in Figure 1 and the following theorem establishes the limiting distribution of the test statistic under the null hypothesis of linearity.

**Theorem 3.2** *Let Assumption 2.1 and 3.1 be fulfilled and assume that  $\inf_{u,\lambda} f(u, \lambda) \geq c > 0$ . Then we have*

$$\hat{G}_T(\lambda_1, \lambda_2) := \sqrt{\frac{M}{\hat{\Sigma}_{B,N}}} \left( \frac{1}{M} \sum_{j=1}^M \frac{\sqrt{N}}{B} \left( \frac{|\hat{f}(u_j, \lambda_1, \lambda_2)|^2}{\sigma_B \hat{f}(u_j, \lambda_1) \hat{f}(u_j, \lambda_2) \hat{f}(u_j, \lambda_1 + \lambda_2)} - \hat{D}_T(u_j) \right) - \frac{B}{\sqrt{N}} \right) \xrightarrow{D} \mathcal{N}(0, 1)$$

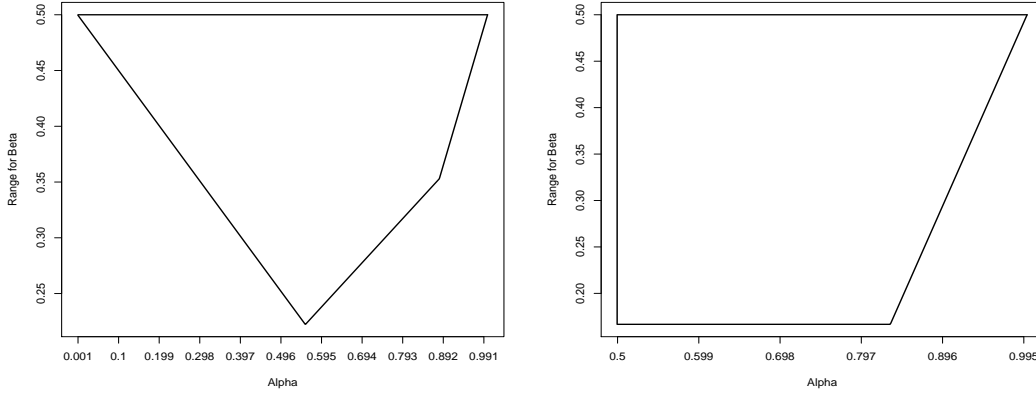


Figure 1: Range for  $(\alpha, \beta)$  if  $\mathbb{E}(Z_t^3) = 0$  (left panel) and  $\mathbb{E}(Z_t^3) \neq 0$  (right panel).

with

$$\hat{D}_N(u) := \frac{1}{2\pi\sigma_B} \frac{\left(\sum_{|k|<N} \frac{1}{N} \sum_{q=0}^{N-1} X_{[uT]-N/2+1+q}^2 X_{[uT]-N/2+1+q+k}\right)^2}{(\hat{\gamma}(u, 0))^3 2\pi \hat{f}(u, 0)}$$

$$\hat{\Sigma}_{B,N} := \frac{4}{M} \sum_{j=1}^M \hat{D}_N(u_j) + 2B^2/N$$

and  $\sigma_B$  being defined in Theorem 2.2. In addition we have

$$(11) \quad \lim_{T \rightarrow \infty} \text{Cov}(\hat{G}_T(\lambda_1^{(1)}, \lambda_2^{(1)}), \hat{G}_T(\lambda_1^{(2)}, \lambda_2^{(2)})) = 0$$

if  $(\lambda_1^{(1)}, \lambda_2^{(1)}) \neq (\lambda_1^{(2)}, \lambda_2^{(2)})$ .

Note that the proof of the above theorem reveals that  $\hat{D}_N(u)$  converges in probability to  $D := (2\pi\sigma_B)^{-1} E(Z^3)^2$  for every  $u$ , i.e. based on the above theorem, a formal level- $\alpha$  test for linearity can be obtained by calculating  $\hat{G}_T(\lambda_1, \lambda_2)$  for  $m$  different pairs

$$(\lambda_1^{(1)}, \lambda_2^{(1)}), \dots, (\lambda_1^{(m)}, \lambda_2^{(m)})$$

[cf. Berg et al. (2010)]. The null hypothesis is then rejected if the absolute value of the sum over all  $\hat{G}_T(\lambda_1^{(i)}, \lambda_2^{(i)})$  is larger than  $\sqrt{m}$  times the  $1 - \alpha/2$  quantile of the standard normal distribution. The assumption of Gaussianity can be verified in the same way if  $\hat{D}_T(u)$  in  $\hat{G}_T(\lambda_1, \lambda_2)$  is replaced by zero.

**Acknowledgements.** This work has been supported in part by the Collaborative Research Center “Statistical modeling of nonlinear dynamic processes” (SFB 823, Teilprojekt C1) of the German Research Foundation (DFG).

## 4 Appendix: technical details

**Proof of Theorem 2.2:** (i) By using (2) and a second order Taylor expansion similar to Dette et al. (2011) we obtain

$$\begin{aligned}\mathbb{E}(\hat{f}(u, \lambda_1, \lambda_2)) &= \frac{1}{(2\pi)^2} \sum_{k_1, k_2 = -\infty}^{\infty} w(k_1/B, k_2/B) \frac{1}{N} \sum_{p=0}^{N-1} \sum_{l, m, n = -\infty}^{\infty} \psi_l(u) \psi_m(u) \psi_n(u) \\ &\quad \times \mathbb{E}(Z_{[uT] - N/2 + 1 + p - l} Z_{[uT] - N/2 + 1 + p + k_1 - m} Z_{[uT] - N/2 + 1 + p + k_2 - n}) \\ &\quad \times \exp(-i\lambda_{1,N}k_1 - i\lambda_{2,N}k_2) + O(N^2/T^2 + 1/T).\end{aligned}$$

Since the innovations  $Z_t$  are independent and have mean zero, we obtain that the conditions  $k_1 = m - l$  and  $k_2 = n - l$  must hold for the corresponding sum not to vanish. This yields

$$\begin{aligned}\mathbb{E}(\hat{f}(u, \lambda_1, \lambda_2)) &= \frac{1}{(2\pi)^2} \sum_{l, m, n = -\infty}^{\infty} w((m-l)/B, (n-l)/B) \mathbb{E}(Z_1^3) \psi_l(u) \psi_m(u) \psi_n(u) \\ &\quad \times \exp(-i\lambda_{1,N}(m-l) - i\lambda_{2,N}(n-l)) + O(N^2/T^2 + 1/T) \\ &= f(u, \lambda_{1,N}, \lambda_{2,N}) + O(1/B^2 + N^2/T^2 + 1/T),\end{aligned}$$

where the last equality follows with a second order Taylor expansion of  $w$  around zero in combination with Assumption 2.1 and (3).

(ii) It is

$$\begin{aligned}N/B^2 \text{Var}(\hat{f}(u, \lambda_1, \lambda_2)) &= N/B^2 \text{cum}(\hat{f}(u, \lambda_1, \lambda_2), \hat{f}(u, \lambda_1, \lambda_2)^*) \\ &= \frac{N}{B^2(2\pi)^4} \sum_{k_1, k_2, k_3, k_4 = -B}^B w(k_1/B, k_2/B) w(k_3/B, k_4/B) \frac{1}{N^2} \sum_{p_1, p_2 = 0}^{N-1} \sum_{l_0, l_1, l_2, m_0, m_1, m_2 = -\infty}^{\infty} \\ &\quad \times \psi_{l_0}(u) \psi_{l_1}(u) \psi_{l_2}(u) \psi_{m_0}(u) \psi_{m_1}(u) \psi_{m_2}(u) \\ &\quad \times \text{cum}(Z_{[uT] - N/2 + 1 + p_1 - l_0} Z_{[uT] - N/2 + 1 + p_1 + k_1 - l_1} Z_{[uT] - N/2 + 1 + p_1 + k_2 - l_2}, \\ &\quad \quad Z_{[uT] - N/2 + 1 + p_2 - m_0} Z_{[uT] - N/2 + 1 + p_2 + k_3 - m_1} Z_{[uT] - N/2 + 1 + p_2 + k_4 - m_2}) \\ &\quad \times \exp(-i\lambda_{1,N}(k_1 - k_3) - i\lambda_{2,N}(k_2 - k_4)) + O(N^2/T^2 + 1/T).\end{aligned}$$

We denote the six innovations  $Z$  above with  $Y_1, \dots, Y_6$  and by using the product theorem for cumulants [cf. Theorem 2.3.2 in Brillinger (1981)] we obtain

$$\text{cum}(Y_1 Y_2 Y_3, Y_4 Y_5 Y_6) = \sum_{\nu} \text{cum}(Y_a; a \in \nu_1) \cdots \text{cum}(Y_a; a \in \nu_l),$$

where the sum runs over all indecomposable partitions  $\nu = (\nu_1, \dots, \nu_l)$  of

$$\begin{array}{c} Y_1 \ Y_2 \ Y_3 \\ Y_4 \ Y_5 \ Y_6. \end{array}$$

The only partition which is not vanishing in the limit is  $\nu^* = ((1, 4), (2, 5), (3, 6))$ . In this case we obtain (due to the independence of the  $Z_t$ ) the conditions

$$\begin{aligned} p_1 &= p_2 + l_0 - m_0 \\ k_3 &= k_1 + l_0 - m_0 + m_1 - l_1 \\ k_4 &= k_2 + l_0 - m_0 + m_2 - l_2 \end{aligned}$$

and the corresponding summand in the covariance is therefore equal to

$$\begin{aligned} & \frac{1}{B^2(2\pi)^4} \sum_{k_1, k_2, k_3, k_4 = -B}^B w(k_1/B, k_2/B) w((k_1 + l_0 - m_0 + m_1 - l_1)/B) w((k_2 + l_0 - m_0 + m_2 - l_2)/B) \\ & \sum_{l_0, l_1, l_2, m_0, m_1, m_2 = -\infty}^{\infty} \psi_{l_0}(u) \psi_{l_1}(u) \psi_{l_2}(u) \psi_{m_0}(u) \psi_{m_1}(u) \psi_{m_2}(u) \\ & \exp(-i\lambda_{1,N}(m_0 - l_0 + l_1 - m_1) - i\lambda_{2,N}(m_0 - l_0 + l_2 - m_2)) \\ & = \frac{1}{2\pi B^2} \sum_{k_1, k_2 = -B}^B w(k_1/B, k_2/B)^2 f(u, \lambda_{1,N}) f(u, \lambda_{2,N}) f(u, \lambda_{1,N} + \lambda_{2,N}) + O(1/B) \end{aligned}$$

which yields the claim by approximating the integral by the corresponding Riemann sum.

(iii) By proceeding as in the proof of Theorem 1 in Dette et al. (2011) we obtain

$$N^{l/2}/B^l \text{cum}_l(\hat{f}(u, \lambda_1, \lambda_2)) = o(1) \quad \text{for } l \geq 3$$

which directly yields the claim with (i), (ii) and the grow assumptions on  $N, B, T$ .  $\square$

**Proof of Theorem 3.2:** We start by showing that it suffices to prove

$$(4.1) \quad \sqrt{\frac{M}{\Sigma_{B,N}}} \left( \frac{1}{M} \sum_{j=1}^M \frac{\sqrt{N}}{B} \left( \frac{|\hat{f}(u_j, \lambda_1, \lambda_2)|^2}{\sigma_B f(u_j, \lambda_1) f(u_j, \lambda_2) f(u_j, \lambda_1 + \lambda_2)} - D \right) - \frac{B}{\sqrt{N}} \right) \xrightarrow{D} \mathcal{N}(0, 1),$$

where

$$D := \frac{E(Z^3)^2}{2\pi\sigma_B\sigma^6} \quad \text{and} \quad \Sigma_{B,N} := 4D + 2B^2/N.$$

This is ensured by the following two claims:

$$(4.2) \quad \sqrt{M/\Sigma_{B,N}} \frac{\sqrt{N}}{B} \left( \frac{1}{M} \sum_{j=1}^M \hat{D}_N(u_j) - D \right) \xrightarrow{P} 0,$$

$$(4.3) \quad \begin{aligned} & \sqrt{\frac{M}{\Sigma_{B,N}}} \frac{1}{M} \sum_{j=1}^M \frac{\sqrt{N}}{B} \frac{|\hat{f}(u_j, \lambda_1, \lambda_2)|^2}{\sigma_B \hat{f}(u_j, \lambda_1) \hat{f}(u_j, \lambda_2) \hat{f}(u_j, \lambda_1 + \lambda_2)} \\ & = \sqrt{\frac{M}{\Sigma_{B,N}}} \frac{1}{M} \sum_{j=1}^M \frac{\sqrt{N}}{B} \frac{|\hat{f}(u_j, \lambda_1, \lambda_2)|^2}{\sigma_B f(u_j, \lambda_1) f(u_j, \lambda_2) f(u_j, \lambda_1 + \lambda_2)} + o_P(1). \end{aligned}$$



Proof of (4.2): It is straightforward to show that  $\frac{1}{M} \sum_{j=1}^M \hat{D}_N(u_j)$  converges to

$$\int_0^1 \frac{1}{2\pi\sigma_B} \frac{\left(\sum_{h=-\infty}^{\infty} \text{Cov}(X_t^2(u), X_{t+h}(u))\right)^3}{\text{Var}(X_t(u))^3 2\pi f(u, 0)} du$$

in probability and a simple calculation yields that this quantity already equals  $D$ . By using (2) and a second order Taylor expansion similar to Dette et al. (2011) we obtain

$$\begin{aligned} \mathbb{E}(\hat{f}(u, \lambda)) &= \frac{1}{2\pi} \sum_{k=-B}^B w(k/B) \frac{1}{N} \sum_{p=0}^{N-1} \sum_{l,m=-\infty}^{\infty} \psi_l(u) \psi_m(u) \mathbb{E}(Z_{\lfloor uT \rfloor - N/2 + 1 + p - m} Z_{\lfloor uT \rfloor - N/2 + 1 + p + k - l}) \exp(-i\lambda_N k) \\ &\quad + O(N^2/T^2 + 1/T) \\ &= \frac{1}{2\pi} \sum_{\substack{l,m=-\infty \\ |l-m| \leq B}}^{\infty} w((l-m)/B) \psi_l(u) \psi_m(u) \exp(-i\lambda_N(l-m)) + O(N^2/T^2), \end{aligned}$$

where the last equality is due to the independence of the innovations. By using a second order Taylor expansion of  $w$  around zero in combination with Assumption 2.1 and (3) it follows that

$$\mathbb{E}(\hat{f}(u, \lambda)) = f(u, \lambda) + O(N^2/T^2 + 1/B^2 + 1/T)$$

and a straightforward calculation yields  $\text{Var}(\hat{f}(u, \lambda)) = O(B/N)$ . So we have

$$(4.4) \quad \hat{f}(u, \lambda) - f(u, \lambda) = O_P(\sqrt{B/N} + N^2/T^2 + 1/T + 1/B^2) = O_P(N^2/T^2 + \sqrt{B/N})$$

and similarly we obtain

$$(4.5) \quad \hat{\gamma}(u, k) - \gamma(u, k) = O_P(1/\sqrt{N} + N^2/T^2)$$

for every fixed  $k \in \mathbb{N}$ . (4.4), (4.5) and a second order Taylor expansion now imply

$$(4.6) \quad \begin{aligned} \frac{1}{M} \sum_{j=1}^M \hat{D}_N(u_j) &= \frac{1}{2\pi\sigma_B M} \sum_{j=1}^M \frac{\left(\sum_{|k| < N} \frac{1}{N} \sum_{q=0}^{N-1} X_{\lfloor t_j \rfloor - N/2 + 1 + q}^2 X_{\lfloor t_j \rfloor - N/2 + 1 + q + k}\right)^2}{\text{Var}(X_t(u_j))^3 2\pi f(u_j, 0)} \\ &\quad + O_P\left(\frac{1}{M} \sum_{j=1}^M C_1(u_j) (\hat{f}(u_j, 0) - f(u_j, 0))\right) + O_P\left(\frac{1}{M} \sum_{j=1}^M C_2(u_j) (\hat{\gamma}(u_j, 0)^2 - \text{Var}(X_t(u_j))^2)\right) \\ &\quad + O_P(B/N + N^4/T^4 + B^{1/2} N^{3/2}/T^2) \end{aligned}$$

with

$$C_1(u) := -\frac{3(\sum_h \text{Cov}(X_t(u)^2, X_{t+h}(u)))^2}{\text{Var}(X_t(u))^4 2\pi f(u, 0)}, \quad C_2(u) := -\frac{(\sum_h \text{Cov}(X_t(u)^2, X_{t+h}(u)))^2}{\text{Var}(X_t(u))^3 2\pi f(u, 0)^2}.$$

If  $\mathbb{E}(Z^3) = 0$  we have  $C_1(u) = C_2(u) = 0$  for all  $u$ , so we now assume that  $\mathbb{E}(Z^3) \neq 0$  holds. Analogously to the calculation of  $\mathbb{E}(\hat{f}(u, \lambda))$  we obtain

$$\mathbb{E}\left(\frac{1}{M} \sum_{j=1}^M C_1(u_j) \hat{f}(u_j, 0)\right) = \frac{1}{M} \sum_{j=1}^M C_1(u_j) f(u_j, 0) + O(N^2/T^2 + 1/B^2)$$

and we next examine the variance of  $1/M \sum_{j=1}^M C_1(u_j) \hat{f}(u_j, 0)$ . It is

$$\begin{aligned} & \text{Var}\left(\frac{1}{M} \sum_{j=1}^M C_1(u_j) \hat{f}(u_j, 0)\right) \\ &= \frac{1}{(2\pi M)^2} \sum_{j_1, j_2=1}^M C_1(u_{j_1}) C_1(u_{j_2}) \sum_{k_1, k_2=-B}^B w(k_1/B) w(k_2/B) \frac{1}{N^2} \sum_{p_1, p_2=0}^{N-1} \sum_{l_i, m_i=-\infty}^{\infty} \psi_{t, T, l_1} \psi_{t, T, m_1} \psi_{t, T, l_2} \psi_{t, T, m_2} \\ & \quad \text{Cov}(Z_{t_{j_1}-N/2+1+p_1-l_1} Z_{t_{j_1}-N/2+1+p_1+k_1-m_1}, Z_{t_{j_2}-N/2+1+p_2-l_2} Z_{t_{j_2}-N/2+1+p_2+k_2-m_2}). \end{aligned}$$

Because of

$$\begin{aligned} & \text{Cov}(Z_{t_{j_1}-N/2+1+p_1-l_1} Z_{t_{j_1}-N/2+1+p_1+k_1-m_1}, Z_{t_{j_2}-N/2+1+p_2-l_2} Z_{t_{j_2}-N/2+1+p_2+k_2-m_2}) \\ &= \text{Cov}(Z_{t_{j_1}-N/2+1+p_1-l_1}, Z_{t_{j_2}-N/2+1+p_2+k_2-m_2}) \text{Cov}(Z_{t_{j_1}-N/2+1+p_1+k_1-m_1}, Z_{t_{j_2}-N/2+1+p_2-l_2}) \\ & \quad + \text{Cov}(Z_{t_{j_1}-N/2+1+p_1-l_1}, Z_{t_{j_2}-N/2+1+p_2-l_2}) \text{Cov}(Z_{t_{j_1}-N/2+1+p_1+k_1-m_1}, Z_{t_{j_2}-N/2+1+p_2+k_2-m_2}) \\ & \quad + \text{cum}(Z_{t_{j_1}-N/2+1+p_1-l_1} Z_{t_{j_1}-N/2+1+p_1+k_1-m_1}, Z_{t_{j_2}-N/2+1+p_2-l_2} Z_{t_{j_2}-N/2+1+p_2+k_2-m_2}) \end{aligned}$$

[cf. Brillinger (1981)] the variance splits into three summands and we restrict ourselves to the first one, because the remaining expressions are treated analogously. In this case we obtain the conditions  $k_1 = m_1 - l_2 + p_2 - p_1 + t_{j_2} - t_{j_1}$ ,  $k_2 = l_1 - m_2 + p_1 - p_2 + t_{j_2} - t_{j_1}$  due to the independence of the  $Z_t$ , thus the first considered term equals

$$\begin{aligned} & \frac{1}{(2\pi M)^2} \sum_{j_1=1}^M \sum_{l_i, m_i=-\infty}^{\infty} \frac{1}{N^2} \sum_{p_1, p_2=0}^{N-1} C_1(u_{j_1}) C_1(u_{j_2}) \sum_{\substack{j_2=1 \\ -B \leq m_1 - l_2 + p_2 - p_1 + t_{j_2} - t_{j_1} \leq B \\ -B \leq l_1 - m_2 + p_1 - p_2 + t_{j_2} - t_{j_1} \leq B}}^M \\ & \quad w((m_1 - l_2 + p_2 - p_1 + t_{j_2} - t_{j_1})/B) w((l_1 - m_2 + p_1 - p_2 + t_{j_2} - t_{j_1})/B) \psi_{t, T, l_1} \psi_{t, T, m_1} \psi_{t, T, l_2} \psi_{t, T, m_2}. \end{aligned}$$

The conditions on  $j_2$  and  $\beta < \alpha$  imply that, if  $j_1, l_i, m_i, p_i$  are chosen and  $T$  is sufficiently large, there is at most one possible choices for  $j_1$  such that the corresponding term does not vanish. Since

$$\frac{1}{B} \sum_{p_1=0}^{N-1} w((m_1 - l_2 + p_2 - p_1 + t_{j_2} - t_{j_1})/B) w((l_1 - m_2 + p_1 - p_2 + t_{j_2} - t_{j_1})/B) \leq C < \infty$$

uniformly in  $p_2, m_i, l_i, t_{j_i}$ , we then obtain

$$\text{Var}\left(\frac{1}{M} \sum_{j=1}^M C_1(u_j) \hat{f}(u_j, 0)\right) = O(B/T),$$

thus the second summand in (4.6) is of order  $O_P(\sqrt{B/T} + N^2/T^2 + 1/B^2)$ . So, roughly speaking, by taking the sample-mean of  $(C_1(u_j)\hat{f}(u_j, 0))_{j=1, \dots, M}$ , we get an 'additional'  $1/\sqrt{M}$  which becomes  $\sqrt{B/T}$  after multiplication with the rate  $\sqrt{B/N}$  of  $\hat{f}(u, \lambda)$ . In the same way it can be shown that the third term in (4.6) is a  $O_P(1/\sqrt{T} + N^2/T^2)$ , thus altogether we get

$$\begin{aligned} \frac{1}{M} \sum_{j=1}^M \hat{D}_T(u_j) &= \frac{1}{2\pi\sigma_B M} \sum_{j=1}^M \frac{\left( \sum_{|k| < N} \frac{1}{N} \sum_{q=0}^{N-1} X_{[t_j]-N/2+1+q}^2 X_{[t_j]-N/2+1+q+k} \right)^2}{\text{Var}(X_t(u_j))^3 2\pi f(u_j, 0)} \\ &+ O_P(\sqrt{B/T} + N^2/T^2 + 1/B^2) \times 1\{\mathbb{E}(Z^3) \neq 0\} + O_P(B/N + N^4/T^4 + B^{1/2}N^{3/2}/T^2). \end{aligned}$$

Using the same arguments as above it follows that the first summand in this expression equals  $D$  plus an error term of order  $O_P(1/\sqrt{T} + N^2/T^2)$ . This in combination with Assumption 2.1 yields (4.2).

Proof of (4.3): We start with the case where  $\mathbb{E}(Z^3) = 0$ . Because of Theorem 2.2 we have  $f(u_j, \lambda_1, \lambda_2)|^2 = O_P(B^2/N)$ , which combined with a Taylor expansion yields

$$\begin{aligned} &\frac{|\hat{f}(u_j, \lambda_1, \lambda_2)|^2}{\hat{f}(u_j, \lambda_1)\hat{f}(u_j, \lambda_2)\hat{f}(u_j, \lambda_1 + \lambda_2)} \\ &= \frac{|\hat{f}(u_j, \lambda_1, \lambda_2)|^2}{f(u_j, \lambda_1)f(u_j, \lambda_2)f(u_j, \lambda_1 + \lambda_2)} + O_P(B^2/N)O_P\left((\hat{f}(u_j, \lambda_1) - f(u_j, \lambda_1)) + (\hat{f}(u_j, \lambda_2) - f(u_j, \lambda_2))\right. \\ &\quad \left.+ (\hat{f}(u_j, \lambda_1 + \lambda_2) - f(u_j, \lambda_1 + \lambda_2))\right). \end{aligned}$$

Analogously to the treatment of  $\frac{1}{M} \sum_{j=1}^M C_1(u_j)\hat{f}(u_j, 0)$  it can be shown that

$$1/M \sum_{j=1}^M (\hat{f}(u_j, \lambda_1) - f(u_j, \lambda_1)) = O_P(\sqrt{B/T} + N^2/T^2 + 1/B^2)$$

and since the last two summands in the  $O_P(\cdot)$ -term are treated in the same way the claim follows for  $\mathbb{E}(Z^3) = 0$ . If  $\mathbb{E}(Z^3) \neq 0$  we have

$$|\hat{f}(u_j, \lambda_1, \lambda_2)|^2 - |f(u_j, \lambda_1, \lambda_2)|^2 = O_P(B/\sqrt{N})$$

[this is shown analogously to the proof of Theorem 2.2] and the claim then follows with a Taylor expansion analogously to the case  $\mathbb{E}(Z^3) = 0$ .

Proof of (4.1): We obtain as in the proof of Theorem 2.2 (i) that it is

$$\begin{aligned}
& \mathbb{E}(|\hat{f}(u, \lambda_1, \lambda_2)|^2) \\
&= \frac{1}{(2\pi)^4} \sum_{k_1, k_2, k_3, k_4 = -B}^B w(k_1/B, k_2/B) w(k_3/B, k_4/B) \frac{1}{N^2} \sum_{p_1, p_2 = 0}^{N-1} \sum_{l_0, l_1, l_2, m_0, m_1, m_2 = -\infty}^{\infty} \\
&\quad \times \psi_{l_0}(u) \psi_{l_1}(u) \psi_{l_2}(u) \psi_{m_0}(u) \psi_{m_1}(u) \psi_{m_2}(u) \\
&\quad \times \mathbb{E}(Z_{[uT]-N/2+1+p_1-l_0} Z_{[uT]-N/2+1+p_1+k_1-l_1} Z_{[uT]-N/2+1+p_1+k_2-l_2} \\
&\quad \quad Z_{[uT]-N/2+1+p_2-m_0} Z_{[uT]-N/2+1+p_2+k_3-m_1} Z_{[uT]-N/2+1+p_2+k_4-m_2}) \\
(4.7) \quad &\quad \times \exp(-i\lambda_{1,N}(k_1 - k_3) - i\lambda_{2,N}(k_2 - k_4)) + O(N^2/T^2 + 1/T).
\end{aligned}$$

We denote the six innovations  $Z$  above with  $Y_1, \dots, Y_6$  and by using the product theorem for cumulants [cf. Theorem 2.3.2 in Brillinger (1981)] we obtain

$$\mathbb{E}(Y_1 \cdots Y_6) = \sum_{\nu} \text{cum}(Y_a; a \in \nu_1) \cdots \text{cum}(Y_a; a \in \nu_l),$$

where the sum runs over all indecomposable partitions  $\nu = (\nu_1, \dots, \nu_l)$  of  $1\ 2 \cdots 6$ . The only two partitions which do not vanish in the limit are  $\nu_1^* = ((1, 2, 3), (4, 5, 6))$  and  $\nu_2^* = ((1, 4), (2, 5), (3, 6))$  and we denote the corresponding summands in (4.7) with  $E_{\nu_1^*}$  and  $E_{\nu_2^*}$ , respectively. Concerning  $\nu_1^*$  we obtain with the independence of the innovations that the relations  $k_1 = l_1 - l_0$ ,  $k_2 = l_2 - l_0$ ,  $k_3 = m_1 - m_0$  and  $k_4 = m_2 - m_0$  have to hold, which, with a second order Taylor expansion of  $w$  around zero in combination with Assumption 2.1 and (3), yields

$$E_{\nu_1^*} = \frac{\mathbb{E}(Z^3)^2}{2\pi\sigma^6} f(u, \lambda_{1,N}) f(u, \lambda_{2,N}) f(u, \lambda_{1,N} + \lambda_{2,N}) + O(1/B^2).$$

As in the proof of Theorem 2.2 (ii), we additionally obtain

$$E_{\nu_2^*} = B^2/N\sigma_B f(u, \lambda_{1,N}) f(u, \lambda_{2,N}) f(u, \lambda_{1,N} + \lambda_{2,N}) + O(1/B^2)$$

and therefore  $\mathbb{E}(\hat{G}_{T,2}(\lambda_1, \lambda_2)) = o(1)$ . By employing the exact same arguments one obtains  $\text{Var}(\hat{G}_{T,2}(\lambda_1, \lambda_2)) = 1 + o(1)$ ,  $\text{cum}_l(\hat{G}_{T,2}(\lambda_1, \lambda_2)) = o(1)$  for  $l \geq 3$  and (11).  $\square$

## References

- Berg, A., Paparoditis, E., and Politis, D. N. (2010). A bootstrap test for time series linearity. *Journal of Statistical Planning and Inference*, 140:3841–3857.
- Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory*. McGraw Hill, New York.
- Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Annals of Statistics*, 25(1):1–37.

- Dette, H., Preuß, P., and Vetter, M. (2011). A measure of stationarity in locally stationary processes with applications to testing. *Journal of the American Statistical Association*, 106(495):1113–1124.
- Fryzlewicz, P., Sapatinas, T., and Subba Rao, S. (2006). A Haar-Fisz technique for locally stationary volatility estimation. *Biometrika*, 93:687–704.
- Priestley, M. (1983). *Spectral analysis and time series, Vol 1 and II*. Academic Press.
- Starica, C. and Granger, C. (2005). Nonstationarities in stock returns. *The Review of Economics and Statistics*, 87:503–522.