# Discriminating between long-range dependence and non-stationarity

Philip Preuß, Mathias Vetter
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: philip.preuss@ruhr-uni-bochum.de
email: mathias.vetter@ruhr-uni-bochum.de

September 24, 2012

**Abstract**

This paper is devoted to the discrimination between a stationary long-range dependent model and a non stationary process. We develop a nonparametric test for stationarity in the framework of locally stationary long memory processes which is based on a Kolmogorov-Smirnov type distance between the time varying spectral density and its best approximation through a stationary spectral density. We show that the test statistic converges to the same limit as in the short memory case if the (possibly time varying) long memory parameter is smaller than 1/4 and justify why the limiting distribution is different if the long memory parameter exceeds this boundary. Concerning the latter case the novel FARI($\infty$) bootstrap is introduced which provides a bootstrap-based test for stationarity that only requires the long memory parameter to be smaller than 1/2 which is the usual restriction in the framework of long-range dependent time series. We investigate the finite sample properties of our approach in a comprehensive simulation study and apply the new test to a data set containing log returns of the S&P 500.

AMS subject classification: 62M10, 62M15, 62G10
Keywords and phrases: bootstrap, empirical spectral measure, goodness-of-fit test, integrated periodogram, locally stationary process, long memory, non stationary process, spectral density.

## 1 Introduction

For many decades one of the leading paradigms in time series analysis is the assumption of stationarity which means that the second-order charateristics of the considered time series are constant over time. One of the prime examples which fits into the framework of stationary processes is the well-known ARMA$(p, q)$ model. Such processes are widely used in applications due to their simplicity and flexibility, and they belong to the class of so called short memory models containing a summable autocovariance function $\gamma$.

However, many time series in reality exhibit an effect which is known as long-range dependence (or long memory) and which means that $\gamma$ decays to zero slowly. Usually one has $\gamma(k) \sim Ck^{2d-1}$ as $k \to \infty$ for some $d \in (0, 1/2)$, so in particular the autocovariance function is not absolutely summable. The coefficent $d$ is called long memory parameter, and the most common way to model these kinds of strong dependencies is to employ FARIMA$(p, d, q)$ processes which were introduced in Granger and Joyeux (1980) and Hosking (1981). These long memory extensions of ARMA$(p, q)$ processes are stationary under certain regularity conditions as well. There exists a large literature on long-range dependence in applications, as it occurs e.g. in the modeling of asset volatility, computer network traffic or various other phenomena; see for example Park and Willinger (2000), Henry and Zaffaroni (2002) and Doukhan et al. (2002) for an overview. The assumption of stationarity, however, is always imposed.

More recently, several authors have pointed out that a slow decrease of $\gamma(k)$ might also occur if the true underlying process does actually not possess long memory but is non stationary instead; see Mikosch and Starica (2004), among others. In addition, Starica and Granger (2005) compared the performance of a non stationary model with that of a FARIMA$(1, d, 1)$ and a GARCH$(1, 1)$ process in the framework of volatility forecasting and found out that their non stationary model is leading to superior results. Fryzlewicz et al. (2006) proved that most of the stylized facts which are observed for financial return data can be explained by fitting the simple (but usually non stationary) model

$$X_{t,T} = \sigma(t/T)Z_t, \quad t = 1, \ldots, T, \tag{1.1}$$

to the data, where $T$ here and throughout the paper denotes the sample size, $\sigma(\cdot) : [0, 1] \to I\!R_+$ is a non parametric function and $(Z_t)_t$ is some i.i.d. white noise process. Thus many phenomena in reality can be explained by either fitting a stationary long memory process or a non stationary (short memory) model to the data. A natural question then is how to discriminate between these two approaches.

Although the importance of statistical tests concerning this matter was pointed out by many authors (see e.g. Perron and Qu (2010) or Chen et al. (2010)), there does not exist much research on this topic. Berkes et al. (2006) and Dehling et al. (2011) developed CUSUM and Wilcoxon type tests which discriminate between long-range dependence and changes in mean. While the authors of the first article are testing the null hypothesis that there is no long-range dependence but one change in mean at some unknown point in time (i.e. the alternative corresponds to the case where the process possesses long memory), the latter paper considers the null hypothesis that there is no change in mean but possibly long-range dependence (i.e. the alternative corresponds to the case where there is a change in mean). A similar approach can be found in Sibbertsen and Kruse (2009). However, there exist many other deviations from stationarity besides changes in mean and it is of particular importance to detect variations in the dependency structure of a given time series as well.

This paper is devoted to the construction of a test for stationarity in the framework of locally stationary long memory processes. The concept of local stationarity became quite famous in recent years, because in contrast to other proposals to model non-stationarity it allows for a meaningful asymptotic theory. Locally stationary processes were introduced by Dahlhaus (1997) and there exist numerous articles which are concerned with estimation techniques or segmentation methods in this framework; see Neumann and von Sachs (1997), Adak (1998), Chang and Morettin (1999), Sakiyama and Taniguchi (2004), Dahlhaus and Polonik (2006), Van Bellegem and von Sachs (2008) or Kreiß and Paparoditis (2011), among others. Articles allowing for long memory effects are rare, however, as only Beran (2009), Palma and Olea (2010) and Roueff and von Sachs (2011) considered parametric and semiparametric estimation.

Similarly, there exist several tests for stationarity in the context of locally stationary models [see for example von Sachs and Neumann (2000), Paparoditis (2009), Paparoditis (2010), Dwivedi and Subba Rao (2010), Dette et al. (2011) and Preuß et al. (2012)], but in all articles long-range dependence is excluded, i.e. these methods cannot be employed for discriminating between long memory and non-stationarity. Our aim is to fill this gap, and for this reason we consider a Kolmogorov-Smirnov type distance which was already discussed in Dahlhaus (2009) and Preuß et al. (2012) to measure deviations from stationarity in the short memory case. Precisely, set

$$E := \sup_{(v,\omega)\in[0,1]^2} |E(v,\omega)|, \tag{1.2}$$

where

$$E(v,\omega) := \frac{1}{2\pi}\Big(\int_0^v \int_0^{\pi\omega} f(u,\lambda)d\lambda du - v\int_0^{\pi\omega}\int_0^1 f(u,\lambda)dud\lambda\Big), \quad (v,\omega)\in[0,1]^2,$$

and $f(u,\lambda)$ denotes the time-varying spectral density. Under the null hypothesis of stationary $f(u,\lambda)$ does not depend on $u$ and therefore $E$ equals zero. For this reason it is natural to consider an empirical version of the measure in (1.2) in order to construct a test for stationarity.

The literature on empirical spectral processes in the long memory framework is surprisingly small, even when restricted to the simpler stationary case. To the best of our knowledge, only Kokoszka and Mikosch (1997) have discussed weak convergence of the integrated periodogram to a Gaussian process in the stationary case. Our first goal is therefore to derive the asymptotics of an empirical version $\hat{E}_T(v,\omega)$ of the measure proposed above. In particular, we are able to prove weak convergence of the process $\hat{E}_T(v,\omega)-E(v,\omega)$ at the parametric rate $T^{-1/2}$, but only one if the long memory parameter satisfies $d<1/4$. This is a natural restriction in this framework (see e.g. Fox and Taqqu (1987) for a similar result on quadratic forms) since the covariances of the finite-dimensional limits contain integrals over the square of the spectral density. These do not exist if the boundary at $1/4$ is exceeded.

As a consequence, we obtain a central limit theorem for $\sup_{v,\omega}|\hat{E}_T(v,\omega)|$ under the null hypothesis and if $d<1/4$, but with a rather complicated dependence structure due to the unknown spectral density. Our second main contribution is therefore the invention of the novel FARI($\infty$) bootstrap for which we are able to prove consistency in the situation above. Interestingly, as it automatically adopts to a switch in the rate of convergence this procedure works indeed for the entire case of $d<1/2$ which is the usual assumption in the framework of long-range dependent time series; see for example Berkes et al. (2006).

The paper is organized as follows. In Section 2 we introduce the necessary notation, whereas we describe the testing procedure in Section 3. The FARI($\infty$) bootstrap required to obtain asymptotic quantiles of the test statistic is discussed in Section 4, and we investigate the finite sample behaviour of our approach in Section 5. Finally, we defer all proofs to an appendix in Section 6.

## 2 Locally stationary long memory processes

Locally stationary processes are usually defined via a sequence of stochastic processes $\{X_{t,T}\}_{t=1,\dots,T}$ which possess a time-varying MA($\infty$) representation

$$X_{t,T} = \sum_{l=0}^{\infty} \psi_{t,T,l} Z_{t-l}, \quad t=1,\dots,T, \tag{2.1}$$

with independent and identically distributed $Z_t$ where $\mathbb{E}(|Z_t|^k)<\infty$ for all $k\in\mathbb{N}$; see Dahlhaus and Polonik (2009). For the coefficents $\psi_{t,T,l}$ we assume that

$$\sup_{t,T}\sum_{l=0}^{\infty}\psi_{t,T,l}^2 < \infty \tag{2.2}$$

is fulfilled which ensures that the process in (2.1) is well defined; see Brockwell and Davis (1991). If the $\psi_{t,T,l}$ are independent of $t$ and $T$ the process $X_{t,T}$ is stationary. However, the coefficents $\psi_{t,T,l}$ depend on $t$ and $T$ in general. To ensure that in this case the process $X_{t,T}$ behaves approximately like a stationary process on a small time interval, it is typically assumed that

$$\sup_{t=1,\ldots,T} \sum_{l=0}^{\infty} \big|\psi_{t,T,l} - \psi_l(t/T)\big| = O(1/T) \tag{2.3}$$

holds for twice continuously differentiable functions $\psi_l : [0,1] \to I\!R$, $l \in Z\!\!Z$. Different smoothness conditions on the functions $\psi_l(\cdot)$ are imposed in the literature, and in essentially all articles in the framework of local stationarity it is assumed that in addition to (2.2) the condition

$$\sup_{t,T} \sum_{l=0}^{\infty} |\psi_{t,T,l}||l|^{\delta} < \infty \tag{2.4}$$

is satisfied for some $\delta > 0$. This implies $\sup_{t,T} \sum_{h=0}^{\infty} |\mathrm{Cov}(X_{t,T}, X_{t+h,T})| < \infty$, and therefore long memory models are excluded. For this reason we replace (2.4) by a growth condition which is flexible enough to include long-range dependent time series as well.

**Assumption 2.1** *Suppose we have a sequence of stochastic processes $\{X_{t,T}\}_{t=1,\ldots,T}$ which have an MA($\infty$) representation as in (2.1) with independent and standard normal distributed $Z_t$ such that (2.2) is fulfilled. Furthermore, we assume that (2.3) holds with twice continuously differentiable functions $\psi_l : [0,1] \to I\!R$ which satisfy the following conditions:*

1) *There exist twice differentiable functions $a, d : [0,1] \to I\!R_+$ such that $D := \sup_{u\in[0,1]} |d(u)| < 1/2$ and*

$$\psi_l(u) = a(u)I(l)^{d(u)-1} + O(I(l)^{D-2}) \tag{2.5}$$

*holds uniformly in $u$ as $l \to \infty$, where $I(x) := |x| \cdot 1_{\{x\neq 0\}} + 1_{\{x=0\}}$.*

2) *The time varying spectral density*

$$f(u,\lambda) := \frac{1}{2\pi}\Big|\sum_{l=0}^{\infty} \psi_l(u) \exp(-i\lambda l)\Big|^2 \tag{2.6}$$

*is twice continuously differentiable on $(0,1) \times (0,\pi)$. Furthermore, $f(u,\lambda)$ and all its partial derivatives up to order two are continuous on $[0,1] \times (0,\pi]$.*

3) *There exists a constant $C \in I\!R$ which is independent of $u$ and $\lambda$ such that the first and second derivative of the approximating functions $\psi_l(\cdot)$ satisfy*

$$\sup_{u\in(0,1)} |\psi_l'(u)| \leq C \log|l|/|l|^{1-D}, \tag{2.7}$$
$$\sup_{u\in(0,1)} |\psi_l''(u)| \leq C \log^2|l|/|l|^{1-D}$$

*for $l \neq 0$ and are bounded otherwise. Furthermore, we assume*

$$\sup_{u\in(0,1)} |\partial/\partial u \; f(u,\lambda)| \leq C \log(\lambda)/\lambda^{2D},$$
$$\sup_{u\in(0,1)} |\partial^2/\partial u^2 \; f(u,\lambda)| \leq C \log^2(\lambda)/\lambda^{2D}.$$

*4) We have*

$$\sup_{t,T} |\psi_{t,T,l}| \leq C|l|^{D-1}. \tag{2.8}$$

To simplify the notation we use $C \in I\!R$ as a universal constant throughout this paper. Note that it is common sense to consider only zero mean processes in this framework since observed data can be easily transformed into data with mean zero. Furthermore, innovations $Z_t$ with a time varying variance $\sigma^2(t/T)$ can be included by choosing other coefficents $\psi_{t,T,l}$. The assumption of Gaussianity is standard (see Palma and Olea (2010) or Dette et al. (2011)) and only imposed to simplify technical arguments since the proofs are already quite involved in this situation. It is straightfoward but cumbersome to extend the results to a more general class of linear processes.

To obtain an impression for local stationarity, note that the process

$$X_t(u) = \sum_{l=0}^{\infty} \psi_l(u) Z_{t-l} \tag{2.9}$$

is stationary for every $u \in [0,1]$, and that $X_t(t/T)$ serves as an approximation of $X_{t,T}$ in the sense of (2.3). It is easy to see that (2.5) implies

$$|\mathrm{Cov}(X_t(u), X_{t+k}(u))| \sim y_1(u)/k^{1-2d(u)} \quad \text{as } k \to \infty$$

and

$$f(u,\lambda) \sim y_2(u)/\lambda^{2d(u)} \quad \text{as } \lambda \to 0 \tag{2.10}$$

for some functions $y_i(\cdot)$; see the proof of Theorem 3.1 in Palma (2007) for details. This shows that the autocovariance function $\gamma(u,k) = \mathrm{Cov}(X_0(u), X_k(u))$ is not absolutely summable and that the time varying spectral density $f(u,\lambda)$ has a pole in $\lambda = 0$ for every $u \in [0,1]$. If the considered process is stationary then $u \mapsto d(u)$ is independent of $u$ which yields that $D$ equals the long memory parameter $d$ of a stationary time series. Let us present two examples which fit into the above framework of locally stationary long memory processes. To this end we define the backshift operator $B$ through $B^k X_t := X_{t-k}$, $k \in I\!N$, and we set

$$(1-B)^{d(u)} = \sum_{j=0}^{\infty} \binom{d(u)}{j} (-1)^j B^j,$$

just as for the binomial series. We justify first that stationary FARIMA$(p,d,q)$ processes are included in our theoretical framework and then motivate a time-varying extension of them.

**Example 2.1** (FARIMA$(p,d,q)$). We consider the FARIMA$(p,d,q)$ equation

$$a(B)(1-B)^d X_t = b(B) Z_t \tag{2.11}$$

with $a(z) := 1 - \sum_{j=1}^{p} a_j z^j$ and $b(z) := 1 + \sum_{j=1}^{q} b_j z^j$. Theorem 3.4 in Palma (2007) states that if the polynomials $a(\cdot)$ and $b(\cdot)$ have no common zeros and the zeros of $a(\cdot)$ furthermore lie outside the unit disc $\{z \in \mathbb{C} : |z| \leq 1\}$, then for $d \in (-1, 1/2)$ the equation (2.11) possesses a stationary solution. Moreover, the coefficents $\psi_{t,T,l} = \psi_l(t/T) = \psi_l$ in the MA$(\infty)$ representation of the process are given by

$$\sum_{l=0}^{\infty} \psi_l z^l = (1-z)^{-d} b(z)/a(z),$$

thus the time-homogeneous spectral density $f(u,\lambda) = f(\lambda)$ becomes

$$f(\lambda) = \frac{\sigma^2}{2\pi}|1-\exp(-i\lambda)|^{-2d}\frac{|1+\sum_{j=1}^{q} b_j \exp(-i\lambda j)|^2}{|1-\sum_{j=1}^{p} a_j \exp(-i\lambda j)|^2}. \tag{2.12}$$

In addition, using Lemma 3.2 in Kokoszka and Taqqu (1995) and equation (1.18) in chapter 3 of Zygmund (1959), it can be shown that

$$\psi_l = \frac{b(1)}{a(1)\Gamma(d)}\frac{1}{l^{1-d}} + O\left(\frac{1}{l^{2-d}}\right) \quad \text{as } l \to \infty, \tag{2.13}$$

which in combination with (2.12) yields that part 1) and 2) in Assumption 2.1 are fulfilled. Part 3) of Assumption 2.1 is automatically fulfilled since the process is stationary and hence there is no time dependence. Finally, part 4) follows with $\psi_{t,T,l} = \psi_l$ and (2.13).

**Example 2.2** (tvFARIMA$(p,d,q)$). The time varying extension of (2.11) is given by

$$a(t/T,B)(1-B)^{d(t/T)}X_{t,T} = b(t/T,B)Z_t, \quad t = 1,\dots,T, \tag{2.14}$$

where $a(u,z) := 1-\sum_{j=1}^{p} a_j(u)z^j$, $b(u,z) := 1+\sum_{j=1}^{q} b_j(u)z^j$ for some functions $a_j(\cdot)$, $b_j(\cdot)$ on $[0,1]$, and $d : [0,1] \to (0,D]$ is twice continuously differentiable with $D < 1/2$. (2.14) is called a time varying FARIMA (tvFARIMA) equation. It can be shown that under certain regularity conditions on the functions $a_j(\cdot), b_j(\cdot)$, the equation (2.14) possesses a solution which is a locally stationary long memory process in the sense of Assumption 2.1; see Jensen and Whitcher (2000) for more details. For example, if we are in the framework of a time varying fractional noise (i.e. $p = q = 0$), then a Taylor expansion in $x$ yields

$$(1-x)^{-d(u)} = \sum_{j=0}^{\infty} \eta_j(u)x^j \quad \text{with} \quad \eta_j(u) := \frac{\Gamma(j+d(u))}{\Gamma(d(u))\Gamma(j+1)} \quad \text{for } j \geq 0; \tag{2.15}$$

see Section 1.3.1 in Pipiras and Taqqu (2011). This implies part 2) of Assumption 2.1, and

$$\frac{\Gamma(l+d(u))}{\Gamma(d(u))\Gamma(l+1)} = \frac{1}{\Gamma(d(u))l^{1-d(u)}} + O\left(\frac{1}{l^{2-D}}\right) \quad \text{as } l \to \infty \tag{2.16}$$

as above proves that parts 1) and 3) are satisfied as well. Part 4) holds since it is $\psi_{t,T,l} = \psi_l(t/T)$.

# 3 The testing procedure

Let us now come to the development of a test for stationarity in the case of long memory models. We are thus interested in testing the null hypothesis

$$H_0 : f(u,\lambda) \text{ is independent of } u \tag{3.1}$$

against the alternative that there exists an $\omega \in [0,\pi]$ such that $u \mapsto f(u,\omega)$ is not independent of $u$. Our test will be based on empirical versions of the quantities $E$ and $E(v,\omega)$ specified in (1.2), and we see that $E$ vanishes under the null hypothesis while it is positive under the alternative due to the continuity of the spectral density.

In order to obtain an estimator for $E$ we have to define an empirical version of $E(v,\omega)$ at first, and for this reason we require an estimator for $f(u,\lambda)$. We assume without loss of generality that the sample size $T$ can be decomposed as $T = NM$ where $N$ and $M$ are integers with $N$ even. We then define the local periodogram at the rescaled time point $u \in [0,1]$ by

$$I_N(u,\lambda) := \frac{1}{2\pi N}\Big|\sum_{s=0}^{N-1} X_{\lfloor uT\rfloor - N/2+1+s,T}\exp(-i\lambda s)\Big|^2$$

[see Dahlhaus (1997)], where we have set $X_{j,T} = 0$, if $j \notin \{1,\dots,T\}$. This is the usual periodogram computed from the observations $X_{\lfloor uT\rfloor-N/2+1,T},\dots,X_{\lfloor uT\rfloor+N/2,T}$. It can be shown that the quantity $I_N(u,\lambda)$ is an asymptotically unbiased estimator for the spectral density if $N\to\infty$ and $N = o(T)$. However, $I_N(u,\lambda)$ is not consistent just as the usual periodogram.
An empirical version of $E(v,\omega)$ is now constructed by replacing the integral by its Riemann sum and substituting the time varying spectral density $f(u,\lambda)$ by its (asymptotically) unbiased estimator $I_N(u,\lambda)$. In other words, we define an estimator for $E(v,\omega)$ by

$$\hat{E}_T(v,\omega) := \frac{1}{T}\sum_{j=1}^{\lfloor vM\rfloor}\sum_{k=1}^{\lfloor \omega\frac{N}{2}\rfloor} I_N(u_j,\lambda_k) - \frac{\lfloor vM\rfloor}{M}\frac{1}{T}\sum_{j=1}^{M}\sum_{k=1}^{\lfloor \omega\frac{N}{2}\rfloor} I_N(u_j,\lambda_k), \tag{3.2}$$

where $u_j := t_j/T := (N(j-1)+N/2)/T$ and $\lambda_k := 2\pi k/N$ with $j = 1,\dots,M$ and $k = 1,\dots,N/2$. Note that in this procedure the $T$ observations are divided into $M$ intervals with length $N$ and that the $u_j$ correspond to the midpoints of these intervals in rescaled time. The $\lambda_k$ are the usual Fourier frequencies. We then set

$$E_T(v,\omega) := \frac{1}{T}\sum_{j=1}^{\lfloor vM\rfloor}\sum_{k=1}^{\lfloor \omega\frac{N}{2}\rfloor} f(u_j,\lambda_k) - \frac{\lfloor vM\rfloor}{M}\frac{1}{T}\sum_{j=1}^{M}\sum_{k=1}^{\lfloor \omega\frac{N}{2}\rfloor} f(u_j,\lambda_k),$$

which is the Riemann sum approximation of $E(v,\omega)$, and consider the empirical spectral process

$$\hat{G}_T(v,\omega) := \sqrt{T}\left(\hat{E}_T(v,\omega) - E_T(v,\omega)\right), \quad v,\omega\in[0,1].$$

The following theorem specifies the asymptotic properties of the process $(\hat{G}_T(v,\omega))_{v,\omega}$ in the case $D < 1/4$. Note that the results hold both under the null hypothesis and the alternative, and throughout this paper the symbol $\Rightarrow$ denotes weak convergence in $[0,1]^2$.

**Theorem 3.1** *Suppose that Assumption 2.1 with $D < 1/4$ is satisfied and let*

$$N\to\infty, \quad N/T\to 0, \quad T^{1/2}\log(N)/N^{1-2D}\to 0. \tag{3.3}$$

*Then as $T\to\infty$ we have*

$$\left(\hat{G}_T(v,\omega) + \sqrt{T}C_T\left(v,\omega,(\psi_l(\cdot))_{l\in\mathbb{Z}}\right)\right)_{(v,\omega)\in[0,1]^2} \Rightarrow (G(v,\omega))_{(v,\omega)\in[0,1]^2},$$

*where $(G(v,\omega))_{(v,\omega)\in[0,1]^2}$ is a Gaussian process with mean zero and covariance structure*

$$\mathrm{Cov}(G(v_1,\omega_1),G(v_2,\omega_2)) = \frac{1}{2\pi}\int_0^1\int_0^{\pi\min(\omega_1,\omega_2)} (1_{[0,v_1]}(u)-v_1)(1_{[0,v_2]}(u)-v_2)f^2(u,\lambda)d\lambda du.$$

*$C_T\left(v,\omega,(\psi_l(\cdot))_{l\in\mathbb{Z}}\right)$ denotes a bias term which equals zero if the functions $\psi_l(u)$ are independent of $u$ for all $l\in\mathbb{Z}$ and which is some $O(N^2/T^2 + \log(N)/N^{1-2D})$, uniformly in $v,\omega$, otherwise.*

Even under the alternative the bias term above is negligible for $D < 1/6$, at least for a suitable choice of $N$. This is why it does not appear in the related result in Preuß et al. (2012). More interesting for us is the behaviour under (3.1), however. In this case we have $C_T\left(v, \omega, (\psi_l(\cdot))_{l\in\mathbb{Z}}\right) = E_T(v,\omega) = 0$ for all $v, \omega, T$. Thus Theorem 3.1 implies

$$(\sqrt{T}\hat{E}_T(v,\omega))_{(v,\omega)\in[0,1]^2} \Rightarrow (G(v,\omega))_{(v,\omega)\in[0,1]^2},$$

under the null hypothesis which yields

$$\sqrt{T}\sup_{(v,\omega)\in[0,1]^2}|\hat{E}_T(v,\omega)| \xrightarrow{D} \sup_{(v,\omega)\in[0,1]^2}|G(v,\omega)|.$$

An asymptotic level $\alpha$ test is then given by rejecting (3.1) whenever $\sqrt{T}\sup_{(v,\omega)\in[0,1]^2}|\hat{E}_T(v,\omega)|$ exceeds the $(1-\alpha)$ quantile of the distribution of the random variable $\sup_{(v,\omega)\in[0,1]^2}|G(v,\omega)|$. To obtain consistency of the test, note that $E_T(v,\omega) \geq C$ for some $v,\omega \in [0,1]$ and $T$ large enough, if we are under the alternative. Since Theorem 3.1 implies $|\hat{E}_T(v,\omega) - E_T(v,\omega)| \to 0$ in probability for this specific $(v,\omega)$, it follows that $\sqrt{T}\sup_{(v,\omega)\in[0,1]^2}|\hat{E}_T(v,\omega)|$ blows up to infinity (in probability).
The restriction $D < 1/4$ in Theorem 3.1 is necessary since $f^2(u,\lambda)$ in the asymptotic variance is not integrable anymore if $D \geq 1/4$ due to (2.10). In fact, in the latter case the rate of convergence is different to $T^{-1/2}$ and the calculation of the corresponding variance becomes extremely messy. To circumvent this step, we introduce the FARI($\infty$) bootstrap in the next section and show that it can be employed to approximate the distribution of $\hat{G}_T(v,\omega)$ if $D < 1/2$. This implies a test for stationarity which does not require $D$ to be smaller than $1/4$ but only to be less than $1/2$. This is the usual restriction in this framework since for example FARIMA$(p,d,q)$ models are not stationary anymore if $D \geq 1/2$.
But even in the situation of Theorem 3.1 it is important to use a bootstrap approximation to obtain empirical quantiles, since already under the null hypothesis the limiting distribution depends in a complicated way on the unknown spectral density.

# 4 Bootstrapping the test statistic

In this section we introduce a bootstrap procedure which approximates the distribution of $\hat{G}_T(v,\omega)$ in the case $D < 1/2$. We call our procedure the FARI($\infty$) bootstrap as it extends the AR($\infty$) bootstrap of Kreiß (1988) to the long memory situation. While the AR($\infty$) bootstrap works by choosing a $p = p(T) \in \mathbb{N}$ and then fitting an AR($p$) model to the data, the FARI($\infty$) bootstrap fits an FARIMA$(p,d,0)$ model to the data where in both cases $p = p(T)$ grows to infinity as $T$ gets larger. We will describe this method in more detail later and state now the main technical assumptions which will be required.

**Assumption 4.1** *For the stationary process $X_t$ with strictly positive spectral density $\lambda \mapsto \int_0^1 f(u,\lambda)du$, there exists a $0 < \underline{d} < 1/2$ such that the process $Y_t = (1-B)^{\underline{d}}X_t$ possesses an AR($\infty$)-representation, i.e.*

$$Y_t = \sum_{j=1}^{\infty} a_j Y_{t-j} + Z_t^{AR} \tag{4.1}$$

*where $(Z_t^{AR})_{t\in\mathbb{Z}}$ denotes a Gaussian white noise process with variance $\sigma^2 > 0$, $1 - \sum_{j=1}^{\infty} a_j z^j \neq 0$ for $|z| \leq 1$ and*

$$\sum_{j=1}^{\infty} |a_j| j^7 < \infty. \tag{4.2}$$

The aim of the bootstrap procedure is to reproduce the behaviour of the previous test statistic in case the process $X_t$ is observed. Note that under the null hypothesis $X_t$ basically equals $X_{t,T}$ and $\underline{d}$ is the corresponding long memory parameter.
We start by choosing some $p = p(T) \in I\!N$, estimating $\underline{d}$ through some $\hat{\underline{d}}$ and then fitting an AR($p$) model to the process $Y_t$ from (4.1), i.e. estimating

$$(a_{1,p}, \ldots, a_{p,p}) = \underset{b_{1,p},\ldots,b_{p,p}}{\operatorname{argmin}} \mathbb{E}\Big(Y_t - \sum_{j=1}^{p} b_{j,p} Y_{t-j}\Big)^2.$$

We then consider the approximating process $Y_t^{AR}(p)$ which is defined through

$$Y_t^{AR}(p) = \sum_{j=1}^{p} a_{j,p} Y_{t-j}^{AR}(p) + Z_t^{AR}(p), \tag{4.3}$$

where $Z_t^{AR}(p)$ is a Gaussian white noise process with mean zero and variance $\sigma_p^2 = \mathbb{E}(Y_t - \sum_{j=1}^{p} a_{j,p} Y_{t-j})^2$. The idea is that for $p = p(T) \to \infty$ the process $Y_t^{AR}(p)$ is close to the process $Y_t$ and therefore $(1-B)^{-\underline{d}} Y_t^{AR}(p)$ is close to the stationary process $X_t$ whose spectral density is given through $\lambda \mapsto \int_0^1 f(u,\lambda)du$ as well.
So if we observe the data $X_{1,T}, \ldots, X_{T,T}$, the FARI($\infty$) bootstrap precisely works as follows:

1) Choose $p = p(T) \in I\!N$ and calculate $\hat{\theta}_{T,p} = (\hat{\underline{d}}, \hat{\sigma}_p^2, \hat{a}_{1,p}, \ldots, \hat{a}_{p,p})$ as the minimizer of

$$\frac{1}{T} \sum_{k=1}^{T/2} \Big( \log f_{\theta_p}(\lambda_{k,T}) + \frac{I_T(\lambda_{k,T})}{f_{\theta_p}(\lambda_{k,T})} \Big)$$

where $\lambda_{k,T} = 2\pi k/T$ for $k = 1, \ldots, T/2$, $I_T(\lambda) = \frac{1}{2\pi T} |\sum_{t=1}^{T} X_{t,T} \exp(-i\lambda t)|^2$ is the usual periodogram for stationary processes and

$$f_{\theta_p}(\lambda) = \frac{|1 - \exp(-i\lambda)|^{-2\underline{d}}}{2\pi} \times \frac{\sigma_p^2}{|1 - \sum_{j=1}^{p} a_{j,p} \exp(-i\lambda j)|^2}$$

is the spectral density of a stationary FARIMA($p, \underline{d}, 0$) model which we want to fit. Note that the estimator $\hat{\theta}_{T,p}$ is the classical Whittle estimator of a stationary process; see Whittle (1951).

2) Calculate $Y_{t,T} = (1-B)^{\hat{\underline{d}}} X_{t,T}$ for $t = 1, \ldots, T$ and simulate a pseudo-series $Y_{1,T}^*, \ldots, Y_{T,T}^*$ according to

$$Y_{t,T}^* = Y_{t,T};\ t = 1, \ldots, p, \quad Y_{t,T}^* = \sum_{j=1}^{p} \hat{a}_{j,p} Y_{t-j,T}^* + \hat{\sigma}_p Z_t^*;\ p < t \leq T,$$

where the $Z_t^*$ are independent standard normal distributed random variables with variance $\hat{\sigma}_p^2$.

3) Create the pseudo-series $X_{1,T}^*, \ldots, X_{T,T}^*$ by calculating $X_{i,T}^* = (1-B)^{-\hat{\underline{d}}} Y_{i,T}^*$ and compute $\hat{G}_T^*(v,\omega)$ in the same way as $\hat{G}_T(v,\omega)$ but with the original observations $X_{1,T}, \ldots, X_{T,T}$ replaced by the bootstrap replicates $X_{1,T}^*, \ldots, X_{T,T}^*$ and with $E_T(v,\omega)$ replaced by zero.

Let us mention some implications: Assumption 4.1 together with Lemma 2.3 of Kreiß et al. (2011) yields that there exists a $p_0 \in I\!N$ such that for all $p \geq p_0$ the approximating process $Y_t^{AR}(p)$ defined in (4.3) possesses an MA($\infty$) representation

$$Y_t^{AR}(p) = \sum_{l=0}^{\infty} c_{l,p} Z_{t-l}^{AR}(p)$$

where the additional condition

$$\sum_{l=0}^{\infty} |c_{l,p}| l^7 \leq C < \infty \tag{4.4}$$

follows from (4.2) and Lemma 2.4 of Kreiß et al. (2011). Furthermore, since we use the Whittle estimator, the fitted AR($p$) process $Y_{t,T}^*$ has an $MA(\infty)$ representation

$$Y_{t,T}^* = \sum_{l=0}^{\infty} \hat{c}_{l,p} Z_{t-l}^* \tag{4.5}$$

for every $p$, if at least two observations are different which is typically the case; see for example the discussion following Lemma 2.4 in Kreiß et al. (2011).
Our goal now is to prove consistency of the FARI($\infty$) bootstrap which is concerned with the series $X_{t,T}^*$. Some technical assumptions on rates regarding $p$ and $\hat{\theta}_{T,p}$ are necessary which are standard in the framework of an $AR(\infty)$ bootstrap; see for example Berg et al. (2010) or Kreiß et al. (2011).

**Assumption 4.2** *i) We have $p = p(T) \in [p_{min}(T), p_{max}(T)]$ with $p_{min}(T) \to \infty$, where also*

$$p_{max}^9(T) \log(T)^3 / T \leq C \quad \text{and} \quad \sqrt{T} p_{min}(T)^{-9} / \sqrt{\log(T)} \to 0.$$

*ii) The condition*

$$||\hat{\theta}_{T,p} - \theta_p||_\infty = O_P(\sqrt{\log(T)/T}) \tag{4.6}$$

*holds uniformly in $p$, where $\theta_p = (\underline{d}, \sigma_p^2, a_{1,p}, \ldots, a_{p,p})$ denotes the vector of the true parameters.*

We want to investigate the properties of an $MA(\infty)$ representation of the bootstrap replicates $X_{t,T}^*$ now. If $\hat{\underline{d}} > 0$, a Taylor expansion yields

$$(1-z)^{-\hat{\underline{d}}} = \sum_{l=0}^{\infty} \hat{\eta}_l z^l \quad \text{with } \hat{\eta}_l := \frac{\Gamma(l + \hat{\underline{d}})}{\Gamma(\hat{\underline{d}})\Gamma(l+1)}$$

for $l \in I\!N$; see (2.15) with $d(u)$ replaced by $\hat{\underline{d}}$. Otherwise, for $\hat{\underline{d}} = 0$ we have $\hat{\eta}_l = 1_{\{l=0\}}$. Using this expansion and (4.5) we obtain

$$X_{t,T}^* = (1-B)^{-\hat{\underline{d}}} Y_{t,T}^* = \sum_{l=0}^{\infty} \hat{\psi}_{l,p} Z_{t-l}^*, \tag{4.7}$$

where the parameters $\hat{\psi}_{l,p}$ are given through the relation

$$\hat{\psi}_{l,p} = \sum_{k=0}^{l} \hat{c}_{k,p}\hat{\eta}_{l-k}; \tag{4.8}$$

see for example the proof of Lemma 3.2 in Kokoszka and Taqqu (1995). Under the null hypothesis (3.1) the approximating functions $\psi_l(u)$ of the true process $X_{t,T}$ are constant, i.e. $\psi_l(u) = \psi_l$. If we show consistency of the FARI($\infty$) bootstrap later, we will naturally use similar arguments as in the proof of Theorem 3.1. For this reason we require the coefficents $\hat{\psi}_{l,p} - \psi_l$ to satisfy conditions which are similar to the conditions on the true coefficents as stated in Assumption 2.1. Note that the coefficents $\hat{\psi}_{l,p} - \psi_l$ do not depend on the rescaled time $u$. Therefore all conditions but (2.5) in Assumption 2.1 are automatically fulfilled and the following lemma ensures that we obtain a condition similar to (2.5) as well.

**Lemma 4.3** *Suppose the null hypothesis* (3.1) *holds and let the Assumptions 2.1, 4.1 and 4.2 be satisfied. Then we have*

$$|\hat{\psi}_{l,p} - \psi_l| l^{1-\max(\hat{\underline{d}},\underline{d})} = O_P(p^4\sqrt{\log(T)}/\sqrt{T}), \quad \textit{uniformly in } p, l, \hat{\underline{d}}, \underline{d}.$$

Empirical quantiles of $\sup_{(v,\omega)\in[0,1]^2} |\hat{G}_T(v,\omega)|$ are now obtained by calculating

$$\hat{F}^*_{T,i} := \sup_{(v,\omega)\in[0,1]^2} |\hat{G}^*_{T,i}(v,\omega)| \quad \text{for } i = 1,\dots,B,$$

where $\hat{G}^*_{T,1}(v,\omega),\dots,\hat{G}^*_{T,B}(v,\omega)$ are the $B$ bootstrap replicates of $\hat{G}_T(v,\omega)$. We then reject the null hypothesis, whenever

$$\sqrt{T} \sup_{(v,\omega)\in[0,1]^2} |\hat{E}_T(v,\omega)| > (\hat{F}^*_T)_{T,\lfloor(1-\alpha)B\rfloor}, \tag{4.9}$$

where $(\hat{F}^*_T)_{T,1},\dots,(\hat{F}^*_T)_{T,B}$ denotes the order statistic of $\hat{F}^*_{T,1},\dots,\hat{F}^*_{T,B}$. In order to explain why this bootstrap procedure works, we have to introduce approximations of $\hat{G}_T(v,\omega)$ and $\hat{G}^*_T(v,\omega)$. First, if we replace $X_{t,T}$ in the definition of $\hat{G}_T(v,\omega)$ by $X_t(t/T)$ from (2.9), we denote the resulting process with $\hat{G}_{T,a}(v,\omega)$. Similarly, we set

$$X^*_{t,T,a} = \sum_{l=0}^{\infty} \psi_l Z^*_{t-l}, \tag{4.10}$$

where the $Z^*_t$ are the innovations from part 2) above. We then define $\hat{G}^*_{T,a}(v,\omega)$ in the same way as $\hat{G}^*_T(v,\omega)$, but with the bootstrap series $X^*_{t,T}$ replaced by $X^*_{t,T,a}$.

**Theorem 4.4** *Suppose the null hypothesis* (3.1) *holds and let the Assumptions 2.1, 4.1 and 4.2 be fulfilled. Choose $N \to \infty$ such that $N/T \to 0$. Then*

*a)* $$\sup_{(v,\omega)\in[0,1]^2} |\hat{G}_{T,a}(v,\omega)| \overset{D}{=} \sup_{(v,\omega)\in[0,1]^2} |\hat{G}^*_{T,a}(v,\omega)|,$$

*b)* $$\mathbb{E}\Big(\sup_{(v,\omega)\in[0,1]^2} |\hat{G}^*_{T,a}(v,\omega)|^2\Big)^{-1/2}\Big(\sup_{(v,\omega)\in[0,1]^2} |\hat{G}_T(v,\omega)| - \sup_{(v,\omega)\in[0,1]^2} |\hat{G}_{T,a}(v,\omega)|\Big) = o_P(1),$$

*c)* $$\mathbb{E}\Big(\sup_{(v,\omega)\in[0,1]^2} |\hat{G}^*_{T,a}(v,\omega)|^2\Big)^{-1/2}\Big(\sup_{(v,\omega)\in[0,1]^2} |\hat{G}^*_T(v,\omega)| - \sup_{(v,\omega)\in[0,1]^2} |\hat{G}^*_{T,a}(v,\omega)|\Big) = o_P(1),$$

*d)* $$\mathbb{E} \sup_{(v,\omega)\in[0,1]^2} |\hat{G}^*_{T,a}(v,\omega)|^2 \le C(N^{\max(4\underline{d}-1,0)} + \log(N)1_{\{\underline{d}=1/4\}}).$$

*Part d) holds also under the alternative.*

We need the standardisation in parts b) and c) above in order to incorporate all cases corresponding to $D < 1/2$. Assertion d) proves that the factors can be skipped, if we are in the framework of Theorem 3.1.
Theorem 4.4 and the arguments from Paparoditis (2010) indicate that the test constructed in (4.9) has asymptotic level $\alpha$. It is consistent, since each bootstrap statistic $\sup_{(v,\omega)\in[0,1]^2} |\hat{G}_T^*(v,\omega)|/\sqrt{T}$ converges to zero in probability from part d) above, while $\sup_{(v,\omega)\in[0,1]^2} |\hat{E}_T(v,\omega)|$ becomes under the alternative larger than some positive constant due to Theorem 6.1 a), b) and (2.10).

# 5 Finite sample properties

Our aim now is to demonstrate how the test for stationarity performs in finite sample situations. Since the proposed decision rule (4.9) depends on the choice of $N$ in the estimation of the Kolmogorov-Smirnov type distance and furthermore on the selection of the AR parameter $p$ in the bootstrap procedure, we start by discussing how we choose both parameters. We then investigate the size and power of our test where all reported results are based on 200 bootstrap replications and 1000 simulation runs. Finally we apply our test to a data set containing S&P 500 returns.

## 5.1 Choice of the parameters $N$ and $p$

Although the proposed method does not show much sensitivity with respect to different choices of the AR parameter, we select $p$ throughout this section as the minimizer of the AIC criterion dating back to Akaike (1973), which is defined by

$$\hat{p} = \mathrm{argmin}_p \frac{2\pi}{T} \sum_{k=1}^{T/2} \Big( \log f_{\hat{\theta}(p)}(\lambda_{k,T}) + \frac{I_T(\lambda_{k,T})}{f_{\hat{\theta}(p)}(\lambda_{k,T})} \Big) + p/T$$

in the context of stationary processes due to Whittle (1951). Here, $f_{\hat{\theta}(p)}$ is the spectral density of the fitted stationary FARIMA$(p,d,0)$ process and $I_T$ is the usual stationary periodogram; see step 1) in the description of the FARI$(\infty)$ bootstrap. Therefore we focus in the following discussion on a sensitivity analysis of the test (4.9) with respect to different choices of $N$. We will see that the particular choice of that tuning parameter has typically very little influence on the outcome of the test under the null hypothesis while it can change the power substantially under certain alternatives.

## 5.2 Size of the test

In order to study the approximation of the nominal level, we consider the FARIMA$(1,d,1)$ model

$$(1-\phi B)(1-B)^d X_t = (1+\theta B) Z_t \tag{5.1}$$

for independent and standard Gaussian $Z_t$ and present the results for different values of $\phi, \theta$ and $d$. To be more precise, we simulate

$$(1-\phi B)(1-B)^d X_t = Z_t \tag{5.2}$$

and

$$(1-B)^d X_t = (1+\theta B) Z_t \tag{5.3}$$

for $d \in \{0.2, 0.4\}$ and $\phi, \theta \in \{-0.9, -0.5, 0, 0.5, 0.9\}$. The corresponding results for $d = 0.2$ are depicted in Tables 1 and 2 for the models (5.2) and (5.3), respectively. In the latter case we observe a precise approximation of the nominal level even for $T = 128$ and it can be seen that the results are basically not affected by the choice of $N$ in these cases. For the model (5.2) we obtain very good results for $\phi \in \{-0.5, 0, 0.5\}$ while the nominal level is overestimated for $|\phi| = 0.9$ and smaller $T$. However, the approximation becomes much more precise if $T$ grows and is also robust with respect to different choices of the window length $N$.

**Please insert Tables 1–4 about here**

The results for the case $d = 0.4$ are presented in Table 3 and Table 4 and we can draw exactly the same picture from it as for $d = 0.2$. In fact, apart from the process (5.2) with $\phi = 0.9$, the performance under the null hypothesis does not change at all with different $d$.

### 5.3 Power of the test

To study the power of our test we consider the following three time varying FARIMA$((1, d, 1))$ models

$$(1-B)^d X_{t,T} = \sqrt{\sin(\pi t/T)} Z_t \tag{5.4}$$

$$(1-B)^d X_{t,T} = Z_t + 1.1 \cos\left(1.5 - \cos(4\pi t/T)\right) Z_{t-1} \tag{5.5}$$

$$\left(1 + 0.9\sqrt{t/T} B\right)(1-B)^d X_{t,T} = Z_t \tag{5.6}$$

for independent and standard Gaussian $Z_t$ and different values of $d$. In addition we simulate the time varying fractional noise processes

$$(1-B)^{d(t/T)} X_{t,T} = Z_t \tag{5.7}$$

with either $d_1(u) = 0.4u^2$ or $d_2(u) = 0.1 \times 1(u \leq 0.5) + 0.4 \times 1(u > 0.5)$. Here, in contrast to the models (5.4)–(5.6), the long memory parameter $d(u)$ varies over time.

The results for the alternatives (5.4)–(5.6) are depicted in Table 5, and it is remarkable that the choice of $N$ seems to affect the results more than under the null hypothesis. This is less important for model (5.4), for which the observed rejection frequencies are large even for small sample sizes, whereas the effect can have an extreme impact on the power for the other ones; see first and foremost model (5.5) for $d = 0.2$. We display the results for the alternatives from (5.7) in Table 6 and it can be seen that for these kinds of models the power seems to grow slower in $T$ than for the alternatives (5.4)–(5.6).

**Please insert Tables 5–6 about here**

Again, the sensitivity of the results with respect to the choice of $N$ is rather large, where the best overall performance is obtained if we choose $N$ large.

## 5.4 Data example

In this section we apply the test (4.9) to 4097 observations of the S&P 500 which were recorded between April 10th 1996 and July 13th 2012. We consider the log returns $X_t = \log(Y_{t+1}/Y_t)$ $(t = 1, \ldots, 4096)$ which are plotted in the right panel of Figure 1. We observe that days with either small or large movements are likely to be followed by days with similar fluctuation. This effect is called 'volatility clustering' and serves as the usual motivation to employ GARCH$(p, q)$ processes in the modelling of stock returns.

**Please insert Figures 1–2 about here**

In Figure 2 the ACF (autocorrelation function) is plotted for the log returns $X_t$ (left panel), the absolute values $|X_t|$ (middle panel) and squared returns $X_t^2$ (right panel). It can be seen that the autocorrelation function $\gamma(k)$ of the log returns is rather small if $k \neq 0$. However, if we take the absolute values $|X_t|$ or the squared returns $X_t^2$ then $\gamma(k)$ decays to zero very slow as $k \to \infty$. The latter observation is the main reason to use a long memory model if the volatility of a financial asset is analyzed.
It was shown in Mikosch and Starica (2004) and Fryzlewicz et al. (2006) that all these effects can also occur if model (1.1) is used. Starica and Granger (2005), among others, demonstrated that a simple and natural model like (1.1) is leading to a superior volatility forecast compared to a GARCH or a long range dependent FARIMA model. So it might be beneficial to consider not only complicated (e.g. long-range dependent) stationary processes in the analysis of a financial time series but to take into account models which are not stationary anymore.

**Please insert Figures 3–4 about here**

We applied our test (4.9) with $T = 64$ and $N = 8$ to a rolling window of the 4096 log returns, i.e. we employed our approach using the data $X_i, \ldots X_{i+63}$ for $i = 1, \ldots, 4033$. Thus we obtain 4033 p-values whose histogram is displayed in in the left panel of Figure 3. It can be seen that the assumption of stationarity is usually not justified since for example 789 of the 4033 p-values are equal to zero and 1789 are smaller than 0.2. This effect becomes even more evident if we use a rolling window of $T = 256$ data. In this case we obtain 3841 p-values whose histograms are presented in Figure 4 for different window lengths $N$. If we take $N = 32$ then 2413 of the 3841 p-values are equal to zero and 3300 are smaller than 0.2. So the more data we look at the bigger is the urgency to employ also non stationary processes in the statistical analysis. Moreover, we observe that the histograms in Figure 4 look similar and therefore the results are basically not affected by the choice of $N$.
One interesting observation is that during the period we took into account the data seem to become more non stationary in time which can be oberserved from the two histograms in the middle and the right panel of Figure 3. In the middle panel we display the histogram of the p-values if our test (with $T = 64$ and $N = 8$) is applied to $X_i, \ldots X_{i+63}$ with $i = 1, \ldots, 1000$ while the same is shown in the right panel if our approach is applied to $X_i, \ldots X_{i+63}$ with $i = 3034, \ldots, 4033$. If we look at both histograms it can be seen that there is a significant shift towards lower p-values.
**Acknowledgements** This work has been supported in part by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823, Teilprojekt A1, C1) of the German Research Foundation (DFG).

# 6 Appendix: Proofs

In this section we present the proofs of all results above. We define

$$\phi_{v,\omega,T}(u,\lambda) := \Big(I_{[0,\frac{\lfloor vM\rfloor}{M}]}(u) - \lfloor vM\rfloor/M\Big) I_{[0,\frac{2\pi\lfloor \omega \frac{N}{2}\rfloor}{N}]}(\lambda) \quad \text{for } u,\lambda \geq 0,\ v,\omega \in [0,1],$$

$$\rho_{2,T,D}(y_1,y_2) := \Big(\frac{1}{T}\sum_{j=1}^{M}\sum_{k=1}^{N/2}(\phi_{v_1,\omega_1,T}(u_j,\lambda_k) - \phi_{v_2,\omega_2,T}(u_j,\lambda_k))^2\frac{1}{\lambda_k^{4D}}\Big)^{1/2} \text{ for } y_i = (v_i,\omega_i) \in [0,1]^2 \tag{6.1}$$

and set

$$\phi_{v,\omega}(u,\lambda) := \lim_{T\to\infty}\phi_{v,\omega,T}(u,\lambda) = \big(I_{[0,v]}(u) - v\big)\, I_{[0,\pi\omega]}(\lambda), \quad v,\omega \in [0,1]. \tag{6.2}$$

Note that $M$ and $N$ depend on $T$ and observe the relations

$$\hat{E}_T(v,\omega) = \frac{1}{T}\sum_{j=1}^{M}\sum_{k=1}^{N/2}\phi_{v,\omega,T}(u_j,\lambda_k)I_N(u_j,\lambda_k) \quad \text{and} \quad E_T(v,\omega) = \frac{1}{T}\sum_{j=1}^{M}\sum_{k=1}^{N/2}\phi_{v,\omega,T}(u_j,\lambda_k)f(u_j,\lambda_k)$$

which will be employed in the proofs of the following two main theorems. All results below are assumed to hold uniformly in $v,\omega$ unless otherwise stated.

**Theorem 6.1** *Suppose Assumption 2.1 holds and assume that $v,\omega,v_i,\omega_i \in [0,1]$ for $i \in \mathbb{N}$. Then we have*

*a)* $\mathbb{E}(\hat{E}_T(v,\omega)) = E_T(v,\omega) + C_T\left(v,\omega,(\psi_l(\cdot))_{l\in\mathbb{Z}}\right) + O\left(\log(N)/N^{1-2D}\right) + O\left(1/T\right),$
*where $C_T\left(v,\omega,(\psi_l(\cdot))_{l\in\mathbb{Z}}\right)$ is the bias term specified in Theorem 3.1.*

*b)*
$$\mathrm{Cov}(\hat{E}_T(v_1,\omega_1),\hat{E}_T(v_1,\omega_1)) = \frac{1}{T^2}\sum_{j=1}^{M}\sum_{k=1}^{\lfloor\min(\omega_1,\omega_2)N/2\rfloor}\big(1_{[0,v_1]}(u_j) - v_1\big)\big(1_{[0,v_2]}(u_j) - v_2\big) f^2(u_j,\lambda_k) + O(\log(N)^2/(TN^{1-4D})) + O(N/T^2).$$

*c)* $\mathrm{cum}(\hat{E}_T(v_1,\omega_1),\ldots,\hat{E}_T(v_l,\omega_l)) = o(T^{-l/2})$ *for $D < 1/4$ and $l \geq 3$.*

*d)* $\mathbb{E}|\hat{G}_T(v_1,\omega_1) - \hat{G}_T(v_2,\omega_2)|^k \leq (2k)!C^k\rho_{2,T,D}\left((v_1,\omega_1),(v_2,\omega_2)\right)^k$ *for all even $k \in \mathbb{N}$.*

Thorem 6.1 is the main tool for proving the results from Section 3. Regarding the bootstrap, suppose that the null hypothesis (3.1) holds. The next theorem ensures that the random variable $\hat{G}_T^*(v,\omega)/\sqrt{T}$ can be approximated by the random variable $\hat{G}_{T,a}^*(v,\omega)/\sqrt{T}$.

**Theorem 6.2** *Suppose the null hypothesis* (3.1) *holds, Assumptions 2.1, 4.1 and 4.2 are satisfied, and let $v,\omega,v_i,\omega_i \in [0,1]$ for $i = 1,2$. Let $\alpha > 0$ be fixed and denote with $A_T(\alpha)$ the set where $|\hat{\underline{d}} - \underline{d}| \leq \alpha/4$ and*

$$|\hat{\psi}_{l,p} - \psi_l| l^{1-\max(\hat{\underline{d}},\underline{d})} \leq C\frac{p^4\log(T)^{3/2}}{\sqrt{T}} \quad \forall l \in \mathbb{N} \tag{6.3}$$

*is fulfilled. Then we have*

$$a)\ \mathbb{E}\Big((\hat{G}_T^*(v,\omega)-\hat{G}_{T,a}^*(v,\omega))1_{A_T(\alpha)}\Big)/\sqrt{T}=0.$$

$$b)\ \mathrm{Var}\Big((\hat{G}_T^*(v,\omega)-\hat{G}_{T,a}^*(v,\omega))1_{A_T(\alpha)}\Big)/T=O\Big(p^8\log(T)^3\log(N)^2N^{\max(4\underline{d}-1,0)+\alpha}T^{-2}\Big).$$

$$c)\ \mathbb{E}\Big(|(\hat{G}_T^*(v_1,\omega_1)-\hat{G}_{T,a}^*(v_1,\omega_1))-(\hat{G}_T^*(v_2,\omega_2)-\hat{G}_{T,a}^*(v_2,\omega_2))|^k1_{A_T(\alpha)}\Big)$$
$$\leq (2k)!C^k\tilde{\rho}^k((v_1,\omega_1),(v_2,\omega_2))(p^8\log(T)^3N^{\max(4\underline{d}-1,0)+\alpha}T^{-1})^{k/2}$$

*for all $k\in\mathbb{N}$ even, where $\tilde{\rho}((v_1,\omega_1),(v_2,\omega_2)):=1_{\{v_1\neq v_2\ or\ \omega_1\neq\omega_2\}}$.*

We begin with the proof of Theorem 6.1 and 6.2 for which we require some technical lemmata.

**Lemma 6.3** *Suppose Assumption 2.1 is satisfied. Then for all $\lambda\in(0,\pi)$ and $N\in\mathbb{N}$*

$$\Big|\sum_{\substack{l,m=0\\|l-m|>N}}^{\infty}\psi_l(u)\psi_m(u)\exp(-i\lambda(l-m))\Big|\leq\frac{C}{\lambda N^{1-2D}}.$$

**Proof:** Without loss of generality we only consider the case $m>l$. We have

$$\sum_{\substack{l,m=0\\m-l>N}}^{\infty}\psi_l(u)\psi_m(u)\exp(-i\lambda(l-m))=\sum_{l=0}^{\infty}\psi_l(u)\sum_{m=l+N+1}^{\infty}\psi_m(u)\exp(-i\lambda(l-m)),$$

and the absolute value of the right term can be bounded through

$$\sum_{l=0}^{\infty}\Big|\psi_l(u)\exp(-i\lambda l)\Big|\left(\Big|\sum_{m=l+N+1}^{\infty}\frac{a(u)}{m^{1-d(u)}}\exp(i\lambda m)\Big|+\sum_{m=l+N+1}^{\infty}\Big|\psi_m(u)-\frac{a(u)}{m^{1-d(u)}}\Big|\right)\tag{6.4}$$

where $a(u)$ is the function from (2.5). Equation (2.9) in chapter 5 of Zygmund (1959) says that

$$\Big|\sum_{m=l+N+1}^{\infty}\frac{1}{m^{1-d(u)}}\exp(-i\lambda m)\Big|\leq\frac{C}{\lambda}\frac{1}{(l+N)^{1-D}}$$

holds for a constant $C\in\mathbb{R}$ which is independent of $l$, $N$, $u$ and $\lambda$. In addition, (2.5) implies

$$\sup_u|\psi_l(u)|\leq C|l|^{D-1}\quad\forall l\geq1.\tag{6.5}$$

If we combine the last two statements with (2.5) we can bound (6.4) up to a constant through

$$\sum_{l=1}^{\infty}\frac{1}{l^{1-D}}\left(\frac{1}{\lambda}\frac{1}{(l+N)^{1-D}}+\frac{1}{(l+N)^{1-D}}\right)\leq\frac{C}{\lambda}\frac{1}{N^{1-2D}}.$$

□

**Lemma 6.4** *a) For all $n\geq1$ and $k_1,k_2\in\mathbb{N}$ there exists a constant $C(k_1,k_2)>0$ such that:*

$$\sum_{\substack{l,m=1\\|l-m|\geq n}}^{\infty}\frac{\log^{k_1}|l|\log^{k_2}|m|}{|lm|^{1-D}}\frac{1}{|l-m|}\leq C(k_1,k_2)\left(\frac{\log^{k_1+k_2+1}(n)}{n^{1-2D}}+1_{\{n=1\}}\right).$$

*b) For $n \geq 1$ we have*

$$\sum_{\substack{l,m=1 \\ 0<|l-m|<n}}^{\infty} \frac{1}{|lm|^{1-D}} \leq Cn^{2D}.$$

*c) We write $(+)^{\geq}_{\neq}$ if $|m_1 - l_2| \geq n$ and $m_1 - l_2 + m_2 - l_1 \neq 0$ are fulfilled. Then we have for $n \geq 2$*

$$\sum_{\substack{m_1,m_2,l_1,l_2=1 \\ (+)^{\geq}_{\neq}}}^{\infty} \frac{1}{|m_1 m_2 l_1 l_2|^{1-D}} \frac{1}{|m_1 - l_2 + m_2 - l_1|} \leq \frac{C\log(n)}{n^{1-4D}}.$$

*d) We write $(+)_{\neq}$ if $|m_1 - l_2| \leq n$, $|m_2 - l_1| \leq n$ and $m_1 - l_2 + m_1 - l_1 \neq 0$ hold. Then for $n \geq 2$*

$$\frac{1}{n} \sum_{\substack{m_1,m_2,l_1,l_2=1 \\ (+)_{\neq}}}^{\infty} \frac{1}{|m_1 m_2 l_1 l_2|^{1-D}} \frac{|m_2 - l_1|}{|m_1 - l_2 + m_2 - l_1|} \leq \frac{C\log(n)}{n^{1-4D}}.$$

*e) For $l \geq 3$ we write $(+)_{\neq,l}$ if $|n_1 - m_l| \leq n$ and $|n_{i+1} - m_i| \leq n$ are satisfied for $i \in \{1,\ldots,l-1\}$ and furthermore $m_1 - n_1 + m_2 - n_2 + \ldots + m_l - n_l \neq 0$ holds. Then there exists $C_l > 0$ such that for all $n \geq 2$*

$$\sum_{\substack{m_i,n_i=1 \\ (+)_{\neq,l}}}^{\infty} \frac{1}{|m_1 n_1 m_2 n_2 \cdots m_l n_l|^{1-D}} \frac{1}{|m_1 - n_1 + m_2 - n_2 + \ldots + m_l - n_l|} \leq C_l \log(n) n^{2Dl-4D}.$$

**Proof:** Before we begin with the proof, note that a simple change of variables yields

$$\int_a^b \frac{1}{x^{1-D}} \frac{\log^k x}{(c \pm x)^e} dx = \frac{1}{c^{e-D}} \int_{a/c}^{b/c} \frac{\log^k(cz)}{z^{1-D}} \frac{1}{(1 \pm z)^e} dz \tag{6.6}$$

for $a, b, c, e \in I\!R$ with $a \leq b$, $c > 0$, $k \in I\!N$ if any of the integrals exist. The proof now basically works by considering approximating integrals instead of the sums, using (6.6) and afterwards employing that

$$\int_a^b \frac{|\log^k(z)|}{z^{1-D}} \frac{1}{1-z} dz \leq C(k) + C|\log(1-b)| \tag{6.7}$$

holds for $k \in I\!N_0$, $0 < a < b < 1$ and constants $C(k) \in I\!R$ which are independent of $a$ and $b$. Note that the absolute value of the right hand side of (6.6) is bounded by

$$\frac{1}{c^{e-D}} \int_0^{\infty} \frac{|\log^k(cz)|}{z^{1-D}} \frac{1}{(1 \pm z)^e} dz$$

which is in any case finite if $0 < D < 1$, $0 < e < 1$ and $1 - D + e > 1$. If $e = 1$ and $b/c$ is close to the possible pole 1, (6.7) implies that the integral on the right hand side of (6.6) is only bounded by a constant times some additional log term which incorporates in some way how close the boundary is to 1. This rule of thumb will be helpful in understanding the treatment of the approximating integrals in the following. Since all proofs work in that particular way of replacing the sum through integrals and applying (6.6) and (6.7) afterwards we will

present the details for part a) only.
Proof of a): For $n = 1$ the claim holds obviously, thus let $n \geq 2$. Then

$$\sum_{\substack{l,m=1 \\ |l-m|\geq n}}^{\infty} \frac{\log^{k_1}|l| \log^{k_2}|m|}{|lm|^{1-D}} \frac{1}{|m-l|}$$

can be bounded by the sum of four terms corresponding to the cases $\{l \geq n/2\}$, $\{l \leq -n/2\}$, $\{m \geq n/2\}$ and $\{m \leq -n/2\}$. For symmetry reasons we present the details for the first term only, and it will be divided into two summands corresponding to $m > 0$ and $m < 0$, respectively. Also, we assume w.l.o.g. that $l > m$, i.e. $l - m \geq n$. We treat the case $m > 0$ first which equals

$$\sum_{l\geq n/2, 0\leq m\leq l-n}^{\infty} \frac{\log^{k_1}|l| \log^{k_2}|m|}{|lm|^{1-D}} \frac{1}{|m-l|} = \sum_{l=n+1}^{\infty} \frac{\log^{k_1} l}{l^{1-D}} \sum_{m=2}^{l-n} \frac{\log^{k_2} m}{m^{1-D}} \frac{1}{l-m}.$$

If we treat the expression in the second summand as a function in $m$, it can be seen that this function only has a finite number of points where the first derivate equals zero. Thus it is piecewise monotonic, which allows us to bound the sum over $m$ by its approximating integral, i.e. by

$$\int_1^{l-n+1} \frac{\log^{k_2} x}{x^{1-D}} \frac{1}{l-x} dx = \frac{1}{l^{1-D}} \int_{1/l}^{1-\frac{n-1}{l}} \frac{\log^{k_2}(lz)}{z^{1-D}} \frac{1}{1-z} dz \leq \frac{\log^{k_2}(l)}{l^{1-D}} \int_{1/l}^{1-\frac{n-1}{l}} \frac{1}{z^{1-D}} \frac{1}{1-z} dz.$$

With (6.7) it follows that the entire expression can be (up to a further constant) bounded by

$$\sum_{l=n+1}^{\infty} \frac{\log^{k_1+k_2} l}{l^{2-2D}} \left(1 + \left|\log\left(\frac{n-1}{l}\right)\right|\right) \leq 3 \sum_{l=n+1}^{\infty} \frac{\log^{k_1+k_2+1} l}{l^{2-2D}} = O\left(\frac{\log^{k_1+k_2+1} n}{n^{1-2D}}\right).$$

This yields the claim for $m > 0$ and we now consider the case $m < 0$. A straightforward calculation yields that

$$\sum_{l\geq n/2, m\leq \min(0,l-n)}^{\infty} \frac{\log^{k_1}|l| \log^{k_2}|m|}{|lm|^{1-D}} \frac{1}{|m-l|} \leq \sum_{l=n/2}^{n-1} \frac{\log^{k_1} l}{l^{1-D}} \left(\frac{\log^{k_2} n}{n} + \sum_{m=n-l+1}^{\infty} \frac{\log^{k_2} m}{m^{1-D}} \frac{1}{l+m}\right)$$
$$+ \sum_{l=n}^{\infty} \frac{\log^{k_1} l}{l^{1-D}} \sum_{m=2}^{\infty} \frac{\log^{k_2} m}{m^{1-D}} \frac{1}{l+m},$$

and by replacing the sum over $m$ through its approximating integral we can bound this expression by

$$\frac{\log^{k_2} n}{n} \sum_{l=n/2}^{n-1} \frac{\log^{k_1} l}{l^{1-D}} + \sum_{l=n/2}^{n-1} \frac{\log^{k_1} l}{l^{1-D}} \int_{n-l}^{\infty} \frac{\log^{k_2} x}{x^{1-D}} \frac{1}{l+x} dx + \sum_{l=n}^{\infty} \frac{\log^{k_1} l}{l^{1-D}} \int_1^{\infty} \frac{\log^{k_2} x}{x^{1-D}} \frac{1}{l+x} dx.$$

By using (6.6) we can bound both integrals through a constant times $\log^{k_2}(l)/l^{1-D}$ which then yields the claim by calculating the resulting sums. $\square$

Analogously to the above proof we can show the next lemma, which, although it looks similar to Lemma 6.4 (and is proven in the same way), is different since the index of summation $m$ is fixed.

**Lemma 6.5** *For all $m \in \mathbb{Z}$ and $n \geq 1$ we have*

$$a) \quad \sum_{\substack{l=1 \\ 0<|l-m|<n}}^{\infty} \frac{1}{|l|^{1-D}} \frac{1}{|l-m|} \leq C \left( \frac{\log|m|}{|m|^{1-D}} 1_{\{m \neq 0\}} + 1_{\{m=0\}} \right) \leq C$$

$$b) \quad \sum_{\substack{l=-\infty \\ n/2 \leq |l-m|<n \\ l \neq 0}}^{\infty} \frac{1}{|l|^{1-d}} \frac{1}{n-|l-m|} \leq C \left( \max\left( \frac{\log|n-m|}{|n-m|^{1-d}}, \frac{\log|n+m|}{|n+m|^{1-d}} \right) 1_{\{m \neq n\}} + 1_{\{m=n\}} \right) \leq C.$$

## 6.1 Proof of Theorem 6.1

**Proof of a):** We have

$$\mathbb{E}\big(\frac{1}{T}\sum_{j=1}^{M}\sum_{k=1}^{N/2}\phi_{v,\omega,T}(u_j,\lambda_k)I_N(u_j,\lambda_k)\big) = \frac{1}{T}\sum_{j=1}^{M}\sum_{k=1}^{N/2}\phi_{v,\omega,T}(u_j,\lambda_k)\frac{1}{2\pi N}\sum_{p,q=0}^{N-1}\sum_{l,m=0}^{\infty}$$
$$\psi_{t_j-N/2+1+p,T,l}\psi_{t_j-N/2+1+q,T,m}\mathbb{E}(Z_{t_j-N/2+1+p-m}Z_{t_j-N/2+1+q-l})\exp(-i\lambda_k(p-q)).$$

Set $e_{j,N} := t_j - N/2 + 1$. By using the independence of the innovations $Z_i$ we obtain that the above term equals

$$\frac{1}{T}\sum_{j=1}^{M}\sum_{k=1}^{N/2}\phi_{v,\omega,T}(u_j,\lambda_k)\frac{1}{2\pi N}\sum_{\substack{l,m=0 \\ |l-m|<N}}^{\infty}\sum_{\substack{q=0 \\ 0\leq q+m-l\leq N-1}}^{N-1}\psi_{e_{j,N}+q+m-l,T,l}\psi_{e_{j,N}+q,T,m}\exp(-i\lambda_k(m-l)). \tag{6.8}$$

Write the product of the $\psi$-terms above as

$$\psi_l\big(\frac{e_{j,N}+q+m-l}{T}\big)\psi_m\big(\frac{e_{j,N}+q}{T}\big) + \psi_{e_{j,N}+q+m-l,T,l}\big(\psi_{e_{j,N}+q,T,m} - \psi_m\big(\frac{e_{j,N}+q}{T}\big)\big)$$
$$+ \psi_m\left(\frac{e_{j,N}+q}{T}\right)\left(\psi_{e_{j,N}+q+m-l,T,l} - \psi_l\left(\frac{e_{j,N}+q+m-l}{T}\right)\right)\Bigg\}, \tag{6.9}$$

so (6.8) splits into a sum of three terms. We will now demonstrate that the second summand is of order $O(1/T)$ and analogously for the third one. The absolute value of the second summand can be bounded by

$$\frac{1}{M}\sum_{j=1}^{M}\sum_{\substack{l,m=0 \\ |l-m|<N}}^{\infty}\frac{1}{2\pi N}\sum_{\substack{q=0 \\ 0\leq q+m-l\leq N-1}}^{N-1}|\psi_{e_{j,N}+q+m-l,T,l}|\left|\psi_{e_{j,N}+q,T,m} - \psi_m\left(\frac{e_{j,N}+q}{T}\right)\right|$$
$$\times\left|\frac{1}{N}\sum_{k=1}^{N/2}\phi_{v,\omega,T}(u_j,\lambda_k)\exp(-i\lambda_k(m-l))\right|. \tag{6.10}$$

With (2.3) it follows that in (6.10) the cases $l = 0$ and $l = m$ are of order $O(1/T)$, thus we only consider the case where $0 < |l-m| < N$ and $l \neq 0$. We employ (A.2) of Eichler (2008) which says that there exists a constant $C \in \mathbb{R}$ such that for all $\{r \in \mathbb{Z} : r \bmod N/2 \neq 0\}$ we have

$$\Big|\frac{1}{N}\sum_{k=1}^{N/2}\phi_{v,\omega,T}(u,\lambda_k)\exp(-i\lambda_k r)\Big| \leq \frac{C}{|r \bmod N/2|} \tag{6.11}$$

uniformly in $v,\omega$. Using (2.8), (6.11) and a symmetry argument we can bound (6.10) up to a constant by

$$2\sum_{\substack{l,m=1\\0<|l-m|<N/2}}^{\infty}\frac{1}{|l|^{1-D}}\sup_{q,t_j}\Big|\psi_{t_j-N/2+1+q,T,m}-\psi_m\Big(\frac{t_j-N/2+1+q}{T}\Big)\Big|\frac{1}{|l-m|}$$
$$+\sum_{\substack{l,m=1\\|l-m|=N/2\,\vee\,l=m}}^{\infty}\frac{1}{|l|^{1-D}}\sup_{q,t_j}\Big|\psi_{t_j-N/2+1+q,T,m}-\psi_m\Big(\frac{t_j-N/2+1+q}{T}\Big)\Big|$$

which is of order $O(1/T)$ due to Lemma 6.5 and (2.3). In the following we will bound expressions like the above one w.l.o.g. by a constant times the first summand, i.e. from now on we will only consider the case $0<|l-m|<N/2$ if we derive the order of error terms. We do this since the remaining terms will be either of the same or of smaller order and are treated analogously.
Following the above argumentation we obtain that (6.9) equals

$$\begin{aligned}&\frac{1}{T}\sum_{j=1}^{M}\sum_{k=1}^{N/2}\phi_{v,\omega,T}(u_j,\lambda_k)\frac{1}{2\pi N}\sum_{p,q=0}^{N-1}\sum_{l,m=0}^{\infty}\psi_l\Big(\frac{e_{j,N}+p}{T}\Big)\psi_m\Big(\frac{e_{j,N}+q}{T}\Big)\\&\times\mathbb{E}(Z_{e_{j,N}+p-m}Z_{e_{j,N}+q-l})\exp(-i\lambda_k(p-q))+O(1/T)\\&=\frac{1}{T}\sum_{j=1}^{M}\sum_{k=1}^{N/2}\phi_{v,\omega,T}(u_j,\lambda_k)\frac{1}{2\pi N}\sum_{p,q=0}^{N-1}\sum_{l,m=0}^{\infty}\mathbb{E}(Z_{e_{j,N}+p-m}Z_{e_{j,N}+q-l})\psi_l(u_j)\psi_m(u_j)\exp(-i\lambda_k(p-q))\\&+C_T\left(v,\omega,(\psi_l(\cdot))_{l\in\mathbb{Z}}\right)+O(1/T)\end{aligned}\tag{6.12}$$

with

$$\begin{aligned}C_T\left(v,\omega,(\psi_l(\cdot))_{l\in\mathbb{Z}}\right):=&\frac{1}{T}\sum_{j=1}^{M}\sum_{k=1}^{N/2}\phi_{v,\omega,T}(u_j,\lambda_k)\frac{1}{2\pi N}\sum_{p,q=0}^{N-1}\sum_{l,m=0}^{\infty}\mathbb{E}(Z_{e_{j,N}+p-m}Z_{e_{j,N}+q-l})\exp(-i\lambda_k(p-q))\\&\times\left\{\left(\psi_l\left(\frac{e_{j,N}+p}{T}\right)-\psi_l(u_j)\right)\psi_m(u_j)+\left(\psi_m\left(\frac{e_{j,N}+q}{T}\right)-\psi_m(u_j)\right)\psi_l(u_j)\right.\\&\left.+\left(\psi_l\left(\frac{e_{j,N}+p}{T}\right)-\psi_l(u_j)\right)\left(\psi_m\left(\frac{e_{j,N}+q}{T}\right)-\psi_m(u_j)\right)\right\}.\end{aligned}\tag{6.13}$$

Let us begin with the first summand of (6.12). This term can be rewritten as

$$\begin{aligned}&\frac{1}{T}\sum_{j=1}^{M}\sum_{k=1}^{N/2}\phi_{v,\omega,T}(u_j,\lambda_k)\frac{1}{2\pi N}\sum_{\substack{l,m=0\\|l-m|<N}}^{\infty}\sum_{\substack{q=0\\0\le q+m-l\le N-1}}^{N-1}\psi_l(u_j)\psi_m(u_j)\exp(-i\lambda_k(m-l))\\&=\frac{1}{2\pi T}\sum_{j=1}^{M}\sum_{k=1}^{N/2}\phi_{v,\omega,T}(u_j,\lambda_k)\sum_{\substack{l,m=0\\|l-m|\le N-1}}^{\infty}\psi_l(u_j)\psi_m(u_j)\exp(-i\lambda_k(m-l))\\&-\frac{1}{2\pi TN}\sum_{j=1}^{M}\sum_{k=1}^{N/2}\phi_{v,\omega,T}(u_j,\lambda_k)\sum_{\substack{l,m=0\\|l-m|\le N-1}}^{\infty}|l-m|\psi_l(u_j)\psi_m(u_j)\exp(-i\lambda_k(m-l))=A_T-B_T\end{aligned}\tag{6.14}$$

where $A_T$ and $B_T$ are defined implicitly. (6.5) and (6.11) prove that $B_T$ is up to a constant bounded by

$$\frac{1}{N} \sum_{\substack{l,m=1 \\ 0<|l-m|<N/2}}^{\infty} \frac{1}{l^{1-D}} \frac{1}{m^{1-D}}$$

which is of order $O(\log(N)/N^{1-2D})$ due to Lemma 6.4 b). Note that the cases with either $l = 0$, $m = 0$ or $N/2 \leq |l-m| < N$ are of the same or of smaller order. Consider $A_T$ next. Our aim is to skip the condition $|l-m| \leq N-1$. By employing Lemma 6.3 we obtain

$$\left| \frac{1}{T} \sum_{j=1}^{M} \sum_{k=1}^{N/2} \phi_{v,\omega,T}(u_j, \lambda_k) \frac{1}{2\pi} \sum_{\substack{l,m=0 \\ |l-m|\geq N}}^{\infty} \psi_l(u_j) \psi_m(u_j) \exp(-i\lambda_k(m-l)) \right| = O\left( \frac{\log(N)}{N^{1-2d}} \right),$$

and therefore $A_T$ is the same as

$$\frac{1}{T} \sum_{j=1}^{M} \sum_{k=1}^{N/2} \phi_{v,\omega,T}(u_j, \lambda_k) f(u_j, \lambda_k) + O\left( \frac{\log(N)}{N^{1-2d}} \right) = E_T(v,\omega) + O\left( \frac{\log(N)}{N^{1-2d}} \right).$$

Finally we show that $C_T(v, \omega, (\psi_l(\cdot))_{l\in\mathbb{Z}}) = O(N^2/T^2) + O(\log(N)/N^{1-2D})$ holds uniformly in $v, \omega \in [0,1]$. Without loss of generality we only consider the first summand in (6.13) which equals

$$\frac{1}{T} \sum_{j=1}^{M} \sum_{k=1}^{N/2} \phi_{v,\omega,T}(u_j, \lambda_k) \frac{1}{2\pi N} \sum_{\substack{p=0 \\ 0\leq p+l-m\leq N-1}}^{N-1} \sum_{\substack{l,m=0 \\ |l-m|<N}}^{\infty} \psi_l'(u_j) \psi_m(u_j) \frac{p-N/2+1}{T} \exp(-i\lambda_k(m-l)) + O(N^2/T^2)$$

due to a second order Taylor expansion. We proceed here as for $A_T$ and $B_T$ in (6.14) above, and a similar argument as for $B_T$ proves that we can skip the condition $0 \leq p+l-m \leq N-1$ at the cost of an error of order $O(\log(N)/N^{1-2D})$. Therefore the above expression equals

$$\frac{1}{T} \sum_{j=1}^{M} \sum_{k=1}^{N/2} \phi_{v,\omega,T}(u_j, \lambda_k) \frac{1}{2\pi N} \sum_{p=0}^{N-1} \sum_{\substack{l,m=0 \\ |l-m|<N}}^{\infty} \psi_l'(u_j) \psi_m(u_j) \frac{p-N/2+1}{T} \exp(-i\lambda_k(m-l)) + O\left( \frac{\log(N)}{N^{1-2d}} \right).$$

Using $\sum_{p=0}^{N-1} (p-N/2+1)/T = N/(2T)$ we see that this term is the same as

$$\frac{1}{T} \sum_{j=1}^{M} \sum_{k=1}^{N/2} \phi_{v,\omega,T}(u_j, \lambda_k) \frac{1}{4\pi T} \sum_{\substack{l,m=0 \\ |l-m|<N}}^{\infty} \psi_l'(u_j) \psi_m(u_j) \exp(-i\lambda_k(m-l)) + O\left( \frac{\log(N)}{N^{1-2d}} \right),$$

and its first part is some $O(1/T)$ because of (2.7), (6.5), (6.11) and Lemma 6.4 a) with $k_1 = 1$ and $k_2 = 0$. $\square$

**Proof of b):** We set

$$
\begin{aligned}
V_T^{true} &= \mathrm{Cov}\Big(\frac{1}{T}\sum_{j_1=1}^{M}\sum_{k_1=1}^{N/2}\phi_{v_1,\omega_1,T}(u_{j_1},\lambda_{k_1})I_N(u_{j_1},\lambda_{k_1}), \frac{1}{T}\sum_{j_2=1}^{M}\sum_{k_2=1}^{N/2}\phi_{v_2,\omega_2,T}(u_{j_2},\lambda_{k_2})I_N(u_{j_2},\lambda_{k_2})\Big) \\
&= \frac{1}{T^2}\sum_{j_1,j_2=1}^{M}\sum_{k_1,k_2=1}^{N/2}\phi_{v_1,\omega_1,T}(u_{j_1},\lambda_{k_1})\phi_{v_2,\omega_2,T}(u_{j_2},\lambda_{k_2}) \\
&\times \frac{1}{(2\pi N)^2}\sum_{p_1,p_2,q_1,q_2=0}^{N-1}\sum_{m_1,m_2,l_1,l_2=0}^{\infty}\psi_{e_{j_1,N}+p_1,T,m_1}\psi_{e_{j_1,N}+q_1,T,l_1}\psi_{e_{j_2,N}+p_2,T,m_2}\psi_{e_{j_2,N}+q_2,T,l_2} \\
&\times\ \mathrm{cum}(Z_{e_{j_1,N}+p_1-m_1}Z_{e_{j_1,N}+q_1-l_1}, Z_{e_{j_2,N}+p_2-m_2}Z_{e_{j_2,N}+q_2-l_2})\exp(-i\lambda_{k_1}(p_1-q_1))\exp(-i\lambda_{k_2}(p_2-q_2)).
\end{aligned}
$$

with $e_{j_i,N} = t_{j_i} - N/2 + 1$. We start by considering the approximating version $V_T^{appr}$ which is the same as above, but where all $\psi$-terms have been replaced, so e.g. $\psi_{e_{j_1,N}+p_1,T,m_1}$ by $\psi_{m_1}(u_{j_1})$ and similarly for the others. Using the well-known formula

$$
\begin{aligned}
&\mathrm{cum}(Z_{e_{j_1,N}+p_1-m_1}Z_{e_{j_1,N}+q_1-l_1}, Z_{e_{j_2,N}+p_2-m_2}Z_{e_{j_2,N}+q_2-l_2}) \\
&= \mathrm{cum}(Z_{e_{j_1,N}+p_1-m_1}Z_{e_{j_2,N}+q_2-l_2})\mathrm{cum}(Z_{e_{j_2,N}+p_2-m_2}Z_{e_{j_1,N}+q_1-l_1}) \\
&+ \mathrm{cum}(Z_{e_{j_1,N}+p_1-m_1}Z_{e_{j_2,N}+p_2-m_2})\mathrm{cum}(Z_{e_{j_2,N}+q_1-l_1}Z_{e_{j_1,N}+q_2-l_2}).
\end{aligned} \tag{6.15}
$$

the computation of $V_T^{appr}$ splits into two similar terms which we denote with $V_{T,1}$ and $V_{T,2}$. We start by considering the first one. Because of the independence of the innovations $Z_i$ we obtain

$$
\begin{aligned}
V_{T,1} &= \frac{1}{T^2}\sum_{j_1,j_2=1}^{M}\sum_{k_1,k_2=1}^{N/2}\phi_{v_1,\omega_1,T}(u_{j_1},\lambda_{k_1})\phi_{v_2,\omega_2,T}(u_{j_2},\lambda_{k_2}) \\
&\times \frac{1}{(2\pi N)^2}\sum_{m_1,m_2,l_1,l_2=0}^{\infty}\sum_{\substack{q_1,q_2=0\\ 0\le q_2+m_1-l_2+t_{j_2}-t_{j_1}\le N-1\\ 0\le q_1+m_2-l_1+t_{j_1}-t_{j_2}\le N-1}}^{N-1}\psi_{m_1}(u_{j_1})\psi_{l_1}(u_{j_1})\psi_{m_2}(u_{j_2})\psi_{l_2}(u_{j_2}) \\
&\times \exp(-i\lambda_{k_1}(q_2-q_1+t_{j_2}-t_{j_1}+m_1-l_2))\exp(i\lambda_{k_2}(q_2-q_1+t_{j_2}-t_{j_1}+l_1-m_2)).
\end{aligned}
$$

We divide the sum over the $m_i$ and $l_i$ into two sums, namely one sum where furthermore $|m_1-l_2|<N$ and $|m_2-l_1|<N$ are satisfied [denoted by (+)] and one sum where either $|m_1-l_2|\ge N$ or $|m_2-l_1|\ge N$. Then

$$
\begin{aligned}
V_{T,1} &= \frac{1}{T^2}\sum_{j_1,j_2=1}^{M}\sum_{k_1,k_2=1}^{N/2}\phi_{v_1,\omega_1,T}(u_{j_1},\lambda_{k_1})\phi_{v_2,\omega_2,T}(u_{j_2},\lambda_{k_2}) \\
&\times \frac{1}{(2\pi N)^2}\sum_{\substack{m_1,m_2,l_1,l_2=0\\ (+)}}^{\infty}\sum_{\substack{q_1,q_2=0\\ 0\le q_2+m_1-l_2+t_{j_2}-t_{j_1}\le N-1\\ 0\le q_1+m_2-l_1+t_{j_1}-t_{j_2}\le N-1}}^{N-1}\psi_{m_1}(u_{j_1})\psi_{l_1}(u_{j_1})\psi_{m_2}(u_{j_2})\psi_{l_2}(u_{j_2}) \\
&\times \exp(-i\lambda_{k_1}(q_2-q_1+t_{j_2}-t_{j_1}+m_1-l_2))\exp(i\lambda_{k_2}(q_2-q_1+t_{j_2}-t_{j_1}+l_1-m_2)) + V_{T,1}^{+},
\end{aligned} \tag{6.16}
$$

where $V_{T,1}^{+}$ corresponds to the case of either $|m_1-l_2|\ge N$ or $|m_2-l_1|\ge N$. The first claim will be

$$
V_{T,1}^{+} = O\left(\log(N)^2/(TN^{1-4D})\right), \tag{6.17}
$$

i.e. the second summand in (6.16) vanishes asymptotically. Furthermore, if we consider the first summand in (6.16) and assume that $j_1$ has been chosen, $j_2$ must be equal to $j_1, j_1-1$ or $j_1+1$, as all other combination of $j_1$ and $j_2$ vanish, because of the condition $0 \le q_2+m_1-l_2+t_{j_2}-t_{j_1} \le N-1$ and the fact that the summation is only performed with respect to indices satisfying $|m_1-l_2|<N$. If $j_2$ equals $j_1+1$ we call the corresponding term $V_{T,1}^{j_2=j_1+}$ and denote with $V_{T,1}^{j_2=j_1-}$ the case $j_2=j_1-1$. Jointly with (6.17) we will (only) show that

$$V_{T,1}^{j_2=j_1+} = O\left(\log(N)^2/(TN^{1-4D})\right), \tag{6.18}$$

which means that we can finally restrict ourselves to the case $j_1=j_2$ in the first summand of (6.16).

Proof of (6.17) and (6.18): Note that we can bound the absolute value of $V_{T,1}^{+}$ by the sum of four terms $V_{N,T,i}^{+}$ $[i=1,\dots,4]$ which are the absolute values of the terms corresponding to the following four cases:

$$\begin{aligned}
&1)\ q_2-q_1+t_{j_2}-t_{j_1}+m_1-l_2 \neq 0 \text{ and } q_2-q_1+t_{j_2}-t_{j_1}+l_1-m_2 = 0\\
&2)\ q_2-q_1+t_{j_2}-t_{j_1}+m_1-l_2 = 0 \text{ and } q_2-q_1+t_{j_2}-t_{j_1}+l_1-m_2 \neq 0\\
&3)\ q_2-q_1+t_{j_2}-t_{j_1}+m_1-l_2 \neq 0 \text{ and } q_2-q_1+t_{j_2}-t_{j_1}+l_1-m_2 \neq 0\\
&4)\ q_2-q_1+t_{j_2}-t_{j_1}+m_1-l_2 = 0 \text{ and } q_2-q_1+t_{j_2}-t_{j_1}+l_1-m_2 = 0
\end{aligned} \tag{6.19}$$

Analogously, the absolute value of $V_{T,1}^{j_2=j_1+}$ can be bounded by four terms $V_{N,T,i}^{j_2=j_1+}$ $[i=1,\dots,4]$ where $t_{j_2}-t_{j_1}$ in the above cases is replaced by $N$. We will present the details for the terms $V_{N,T,3}^{+}$ and $V_{N,T,3}^{j_2=j_1+}$ only since they are the dominating ones due to the least restrictive conditions. We start with the treatment of $V_{N,T,3}^{+}$. Setting $\Delta t = t_{j_2}-t_{j_1}$ we obtain from symmetry arguments

$$\begin{aligned}
|V_{N,T,3}^{+}| &= \Big| \frac{1}{T^2} \sum_{j_1,j_2=1}^{M} \sum_{k_1,k_2=1}^{N/2} \phi_{v_1,\omega_1,T}(u_{j_1},\lambda_{k_1}) \phi_{v_2,\omega_2,T}(u_{j_2},\lambda_{k_2}) \\
&\quad \times \frac{1}{(2\pi N)^2} \sum_{\substack{m_1,m_2,l_1,l_2=0\\ \max(|m_1-l_2|,|m_2-l_1|)\ge N}}^{\infty} \sum_{\substack{q_1,q_2=0\\ 0\le q_2+m_1-l_2+\Delta t\le N-1\\ 0\le q_1+m_2-l_1-\Delta t\le N-1\\ (6.19)}}^{N-1} \psi_{m_1}(u_{j_1})\psi_{l_1}(u_{j_1})\psi_{m_2}(u_{j_2})\psi_{l_2}(u_{j_2}) \\
&\quad \times \exp(-i\lambda_{k_1}(q_2-q_1+\Delta t+m_1-l_2)) \exp(-i\lambda_{k_2}(q_1-q_2-\Delta t+m_2-l_1)) \Big| \\
&\le \frac{2}{(2\pi T)^2} \sum_{j_1=1}^{M} \sum_{\substack{m_1,m_2,l_1,l_2=0\\ |m_1-l_2|\ge N}}^{\infty} \sum_{q_2=0}^{N-1} \sum_{\substack{j_2=1\\ 0\le q_2+m_1-l_2+\Delta t\le N-1}}^{M} |\psi_{m_1}(u_{j_1})\psi_{l_1}(u_{j_1})\psi_{m_2}(u_{j_2})\psi_{l_2}(u_{j_2})| \\
&\quad \times \sum_{\substack{q_1=0\\ 0\le q_1+m_2-l_1-\Delta t\le N-1\\ (6.19)}}^{N-1} \Big| \frac{1}{N} \sum_{k_1=1}^{N/2} \phi_{v_1,\omega_1,T}(u_{j_1},\lambda_{k_1}) \exp(-i\lambda_{k_1}(q_2-q_1+\Delta t+m_1-l_2)) \Big| \\
&\quad \times \Big| \frac{1}{N} \sum_{k_2=1}^{N/2} \phi_{v_2,\omega_2,T}(u_{j_2},\lambda_{k_2}) \exp(-i\lambda_{k_2}(q_1-q_2-\Delta t+m_2-l_1)) \Big|.
\end{aligned} \tag{6.20}$$

The conditions $0 \le q_2+m_1-l_2+\Delta t \le N-1$ and $0 \le q_1+m_2-l_1-\Delta t \le N-1$ can only be satisfied if $|m_1-l_2+\Delta t|<N$ and $|m_2-l_1-\Delta t|<N$ hold. By combining this with (6.5) and (6.11) it can be seen that

the above term is up to a constant bounded by

$$
\begin{aligned}
&\frac{1}{T^2}\sum_{j_1=1}^{M}\sum_{\substack{m_1,m_2,l_1,l_2=0\\(+)^{\geq}_{\neq}}}^{\infty}\sum_{q_2=0}^{N-1}\sum_{\substack{j_2=1\\0\leq q_2+m_1-l_2+\Delta t\leq N-1\\|m_1-l_2+\Delta t|<N\\|m_2-l_1-\Delta t|<N}}^{M}\frac{1}{|m_1m_2l_1l_2|^{1-D}}\\
&\times\sum_{\substack{q_1\in A_N\\|q_2-q_1+\Delta t+m_1-l_2|<N/2\\|q_1-q_2-\Delta t+m_2-l_1|<N/2}}\frac{1}{|q_2-q_1+\Delta t+m_1-l_2|}\frac{1}{|q_1-q_2-\Delta t+m_2-l_1|}
\end{aligned}
\tag{6.21}
$$

where $(+)^{\geq}_{\neq}$ was defined in Lemma 6.4 c) and $A_N=\{0,1,2,\ldots,N-1\}\backslash\{z_1,z_2\}$ with $z_1=q_2+\Delta t+m_1-l_2$, $z_2=q_2+\Delta t+l_1-m_2$. We used once more that the cases with $m_i=0$, $l_i=0$, $m_1-l_2+m_2-l_1=0$, $|q_2-q_1+\Delta t+m_1-l_2|\geq N/2$ or $|q_1-q_2-\Delta t+m_2-l_1|\geq N/2$ are of the same or smaller order and that $z_1$ and $z_2$ correspond to the values of $q_1$ for which the argument in one of the exp-function is zero which cannot occur because of (6.19). By considering the approximating integral we can bound the latter sum up to a constant by

$$
\int_A\frac{1}{|q_2-q_1+\Delta t+m_1-l_2|}\frac{1}{|q_1-q_2-\Delta t+m_2-l_1|}dq_1
$$

with $A=[0,N-1]\backslash\{[z_1-1,z_1+1]\cup[z_2-1,z_2+1]\}$. A simple integration via a decomposition into partial fractions yields that (6.21) is thus (up to a constant) bounded by

$$
\begin{aligned}
&\frac{1}{T^2}\sum_{j_1=1}^{M}\sum_{\substack{m_1,m_2,l_1,l_2=0\\(+)^{\geq}_{\neq}}}^{\infty}\sum_{q_2=0}^{N-1}\sum_{\substack{j_2=1\\0\leq q_2+m_1-l_2+\Delta t\leq N-1\\|m_1-l_2+\Delta t|<N\\|m_2-l_1-\Delta t|<N}}^{M}\frac{1}{|m_1m_2l_1l_2|^{1-D}}\\
&\times\left.\frac{\log|q_2-q_1+\Delta t+m_1-l_2|+\log|q_1-q_2-\Delta t+m_2-l_1|}{|m_1-l_2+m_2-l_1|}\right|_{\partial A}
\end{aligned}
$$

where $\big|_{\partial A}$ means that the antiderivative with respect to $q_1$ is computed at all values of the boundary of $A$ and always combined via a sum. We observe that the construction of $A$ together with the conditions on $q_i, m_i, l_i$ and $j_2$ imply that the arguments in the log-function are between 1 and $2N$. Furthermore, for chosen $q_2$, $m_1$, $l_2$ and $j_1$, there is at most one possible choice for $j_2$ for which the corresponding summand does not vanish. Thus the above term can be up to a constant bounded by

$$
\frac{1}{TN}\sum_{\substack{m_1,m_2,l_1,l_2=0\\(+)^{\geq}_{\neq}}}^{\infty}\frac{1}{|m_1m_2l_1l_2|^{1-D}}\frac{1}{|m_1-l_2+m_2-l_1|}\sum_{q_2=0}^{N-1}\log(N)
$$

which is of order $O\left(\log(N)^2/(TN^{1-4D})\right)$ due to Lemma 6.4 c). In the same way we can bound the term $V_{N,T,3}^{j_2=j_1+}$ (up to a constant) by

$$
\frac{1}{TN}\sum_{\substack{m_1,m_2,l_1,l_2=0\\(+)_{\neq}}}^{\infty}\frac{1}{|m_1m_2l_1l_2|^{1-D}}\frac{1}{|m_1-l_2+m_2-l_1|}\sum_{\substack{q_2=0\\0\leq q_2+m_1-l_2+N\leq N-1}}^{N-1}\log(N)
$$

where the differences between the quantities above are the different summation conditions on the $m_i$ and $l_i$ and the constraint $0 \leq q_2 + m_1 - l_2 + N \leq N-1$ on $q_2$. Note that there are only $|m_1 - l_2|$ possible choices for $q_2$ if $m_1$ and $l_2$ are chosen. Therefore $V_{N,T,3}^{j_2=j_1+}$ is bounded by

$$\frac{\log(N)}{TN} \sum_{\substack{m_1,m_2,l_1,l_2=0 \\ (+)_{\neq}}}^{\infty} \frac{1}{|m_1 m_2 l_1 l_2|^{1-D}} \frac{|m_1 - l_2|}{|m_1 - l_2 + m_2 - l_1|} = O\left(\frac{\log(N)^2}{TN^{1-4D}}\right)$$

where the last equality follows with Lemma 6.4 d). We have thus shown (6.17) and (6.18) and can restrict ourselves to the case $j_1 = j_2$ in the first term of (6.16), i.e. $V_{T,1}$ equals

$$\frac{1}{T^2} \sum_{j=1}^{M} \sum_{k_1,k_2=1}^{N/2} \phi_{v_1,\omega_1,T}(u_j, \lambda_{k_1}) \phi_{v_2,\omega_2,T}(u_j, \lambda_{k_2}) \sum_{\substack{m_1,m_2,l_1,l_2=0 \\ (+)}}^{\infty} \sum_{\substack{q_1,q_2=0 \\ 0\leq q_2+m_1-l_2\leq N-1 \\ 0\leq q_1+m_2-l_1\leq N-1}}^{N-1} \psi_{m_1}(u_j)\psi_{l_1}(u_j)\psi_{m_2}(u_j)\psi_{l_2}(u_j)$$
$$\times \frac{1}{(2\pi N)^2} \exp(-i\lambda_{k_1}(q_2 - q_1 + m_1 - l_2)) \exp(i\lambda_{k_2}(q_2 - q_1 + l_1 - m_2)) + O\left(\frac{\log(N)^2}{TN^{1-4D}}\right).$$

Note first that we make an error of order $O(\log(N)^2/(TN^{1-4D}))$ if we skip the conditions on the choice of $q_1$ and $q_2$. This follows in a similar way as above, using (6.5), (6.11) and Lemma 6.4 d) once more. Therefore

$$V_{T,1} = \frac{1}{T^2} \sum_{j=1}^{M} \sum_{k_1,k_2=1}^{N/2} \phi_{v_1,\omega_1,T}(u_j, \lambda_{k_1}) \phi_{v_2,\omega_2,T}(u_j, \lambda_{k_2}) \frac{1}{(2\pi N)^2} \sum_{\substack{m_1,m_2,l_1,l_2=0 \\ (+)}}^{\infty} \sum_{q_1,q_2=0}^{N-1} \psi_{m_1}(u_j)\psi_{l_1}(u_j)\psi_{m_2}(u_j)\psi_{l_2}(u_j)$$
$$\times \exp(-i\lambda_{k_1}(q_2 - q_1 + m_1 - l_2)) \exp(i\lambda_{k_2}(q_2 - q_1 + l_1 - m_2)) + O\left(\log(N)^2/(TN^{1-4D})\right).$$

By employing the well known identity

$$\frac{1}{N} \sum_{q=0}^{N-1} \exp(-i(\lambda_{k_1} - \lambda_{k_2})q) = \begin{cases} 1, & k_1 - k_2 = lN \text{ with } l \in \mathbb{Z} \\ 0, & \text{else} \end{cases}, \tag{6.22}$$

it can be seen that all terms with $k_1 \neq k_2$ are equal to zero and we therefore get

$$V_{T,1} = \frac{1}{T^2} \sum_{j=1}^{M} \sum_{k=1}^{N/2} \phi_{v_1,\omega_1,T}(u_j, \lambda_k) \phi_{v_2,\omega_2,T}(u_j, \lambda_k) \frac{1}{(2\pi)^2} \sum_{\substack{m_1,m_2,l_1,l_2=0 \\ (+)}}^{\infty} \psi_{m_1}(u_j)\psi_{l_1}(u_j)\psi_{m_2}(u_j)\psi_{l_2}(u_j)$$
$$\times \exp(-i\lambda_k(m_1 - l_2 + m_2 - l_1)) + O\left(\log(N)^2/(TN^{1-4D})\right).$$

The same error arises due to (6.5), (6.11) and Lemma 6.4 c), if we finally skip the condition (+). Note that we can proceed completely analogously for the term $V_{T,2}$ with the difference that instead of the right hand side in (6.22) we obtain the corresponding term with $\lambda_{k_1} - \lambda_{k_2}$ replaced by $\lambda_{k_1} + \lambda_{k_2}$. Because of (6.22) we then only have to consider the case $k_1 = k_2 = N/2$ and therefore the whole term is of order $O\left(\log(N)^2/(TN^{1-4D})\right)$. Using the definition of the spectral density the claim follows for $V_T^{appr}$.
What remains is to show $V_T^{true} = V_T^{appr} + O(N/T^2)$. However, the only property of the coefficents $\psi_l(\cdot)$ used in the treatment of $V_{T,1}$ is (6.5). Since (2.8) provides the same property as (6.5) for the original coefficents,

we obtain that $V_T^{true}$ equals the final quantity above but with the approximating functions $\psi_l(u_j)$ recplaced by some $\psi_{t_j+cN,T,l}$, $c \in (-1,1)$. Condition (2.3) then yields that we make an error of order $O(1/T^2)$ if we replace $\psi_{t_j+cN,T,l}$ by $\psi_l(t_j+cN/T)$ and a Taylor expansion combined with (2.7) gives the result. $\square$

**Proof of c):** Assume w.l.o.g. that $(v,\omega) := (v_1,\omega_1) = (v_2,\omega_2) = \ldots = (v_l,\omega_l)$. Using the same replacement of coefficients as in the previous proof we obtain from (2.3), a Taylor expansion and (2.7) the relation

$$\begin{aligned}\mathrm{cum}_l(\hat{E}_T(v,\omega)) = \frac{1}{T^l}\sum_{j_1,\ldots,j_l=1}^{M}\sum_{k_1,\ldots,k_l=1}^{N/2}\phi_{v,\omega,T}(u_{j_1},\lambda_{k_1})\cdots\phi_{v,\omega,T}(u_{j_l},\lambda_{k_l})\frac{1}{(2\pi N)^l}\sum_{p_1,\ldots,q_l=0}^{N-1}\sum_{m_1,\ldots,n_l=0}^{\infty}\\ \times\,\mathrm{cum}(Z_{t_{j_1}-N/2+1+p_1-m_1}Z_{t_{j_1}-N/2+1+q_1-n_1},\ldots,Z_{t_{j_l}-N/2+1+p_l-m_l}Z_{t_{j_l}-N/2+1+q_l-n_l})\\ \times\,\psi_{m_1}(u_{j_1})\cdots\psi_{n_l}(u_{j_l})\exp(-i\lambda_{k_1}(p_1-q_1))\cdots\exp(-i\lambda_{k_l}(p_l-q_l))(1+o(1))\end{aligned}$$

for $l \geq 3$. We define $Y_{i,1} := Z_{t_{j_i}-N/2+1+p_i-m_i}$ and $Y_{i,2} := Z_{t_{j_i}-N/2+1+q_i-n_i}$ for $i \in \{1,\ldots,l\}$. Following chapter 2.3 of Brillinger (1981) we obtain $\mathrm{cum}_l(\hat{E}_T(v,\omega)) = \sum_\nu V_T(\nu)(1+o(1))$, where the sum runs over all indecomposable partitions $\nu = \nu_1 \cup \ldots \cup \nu_l$ with $|\nu_i| = 2$ $(1 \leq i \leq l)$ of the matrix

$$\begin{matrix} Y_{1,1} & Y_{1,2} \\ \vdots & \vdots \\ Y_{l,1} & Y_{l,2} \end{matrix}$$

and

$$\begin{aligned}V_T(\nu) := \frac{1}{T^l}\sum_{j_1,\ldots,j_l=1}^{M}\sum_{k_1,\ldots,k_l=1}^{N/2}\phi_{v,\omega,T}(u_{j_1},\lambda_{k_1})\cdots\phi_{v,\omega,T}(u_{j_l},\lambda_{k_l})\frac{1}{(2\pi N)^l}\sum_{p_1,\ldots,q_l=0}^{N-1}\sum_{m_1,\ldots,n_l=0}^{\infty}\psi_{m_1}(u_{j_1})\cdots\psi_{n_l}(u_{j_l})\\ \times\,\mathrm{cum}(Y_{i,k};(i,k)\in\nu_1)\cdots\mathrm{cum}(Y_{i,k};(i,k)\in\nu_l)\exp(-i\lambda_{k_1}(p_1-q_1))\cdots\exp(-i\lambda_{k_l}(p_l-q_l)).\end{aligned}$$

We now fix one indecomposable partition $\tilde{\nu}$ and assume without loss of generality that

$$\tilde{\nu} = \bigcup_{i=1}^{l-1}(Y_{i,1},Y_{i+1,2}) \cup (Y_{l,1},Y_{1,2}). \tag{6.23}$$

Because of $\mathrm{cum}(Z_i,Z_j) \neq 0$ for $i \neq j$ we obtain the equations $q_1 = p_l + n_1 - m_l + t_{j_l} - t_{j_1}$ and $q_{i+1} = p_i + n_{i+1} - m_i + t_{j_i} - t_{j_{i+1}}$ for $i \in \{1,\ldots,l-1\}$. Therefore only $l$ variables of the $2l$ variables $p_1,q_1,p_2,\ldots,q_l$ are free to choose and must satisfy the conditions

$$0 \leq p_l + n_1 - m_l + t_{j_l} - t_{j_1} \leq N-1 \quad\text{and}\quad 0 \leq p_i + n_{i+1} - m_i + t_{j_i} - t_{j_{i+1}} \leq N-1 \text{ for } i \in \{1,\ldots,l-1\}. \tag{6.24}$$

Thus we obtain

$$\begin{aligned}V_T(\tilde{\nu}) = \frac{1}{T^l}\sum_{j_1,\ldots,j_l=1}^{M}\sum_{k_1,\ldots,k_l=1}^{N/2}\phi_{v,\omega,T}(u_{j_1},\lambda_{k_1})\cdots\phi_{v,\omega,T}(u_{j_l},\lambda_{k_l})\frac{1}{(2\pi N)^l}\sum_{p_1,\ldots,p_l=0}^{N-1}\sum_{\substack{m_1,\ldots,n_l=0\\(6.24)}}^{\infty}\psi_{m_1}(u_{j_1})\cdots\psi_{n_l}(u_{j_l})\\ \times\exp(-i\lambda_{k_1}(p_1-p_l+m_l-n_1+t_{j_1}-t_{j_l}))\prod_{i=2}^{l}\exp(-i\lambda_{k_i}(p_i-p_{i-1}+m_{i-1}-n_i+t_{j_i}-t_{j_{i-1}})).\end{aligned}$$

Note that (6.24) can only be satisfied if $|n_1 - m_l + t_{j_l} - t_{j_1}| < N$ and $|n_{i+1} - m_i + t_{j_i} - t_{j_{i+1}}| < N$ hold for $i \in \{1, 2, \ldots, l-1\}$. Using this fact in combination with (6.5) and (6.11) the term above is (up to a constant) bounded by

$$\frac{1}{T^l} \sum_{j_1=1}^{M} \sum_{\substack{m_1,n_1,\ldots,m_l,n_l=0 \\ m_i,n_i \neq 0}}^{\infty} \sum_{\substack{j_2,\ldots,j_M=1 \\ |n_{i+1}-m_i+t_{j_i}-t_{j_{i+1}}|<N \\ |n_1-m_l+t_{j_l}-t_{j_1}|<N}}^{M} \frac{1}{|m_1|^{1-d}} \cdots \frac{1}{|n_l|^{1-d}} \sum_{\substack{p_1,p_2,\ldots,p_l=0 \\ |p_i-p_{i-1}+m_{i-1}-n_i+t_{j_i}-t_{j_{i-1}}|<N/2}}^{N-1}$$
$$\times \frac{1}{|p_1 - p_l + m_l - n_1 + t_{j_1} - t_{j_l}|} \prod_{i=2}^{l} \frac{1}{|p_i - p_{i-1} + m_{i-1} - n_i + t_{j_i} - t_{j_{i-1}}|} \prod_{i=1}^{l} 1(p_i \notin \{z_{i1}, z_{i2}\})$$

where $z_{i1}$, $z_{i2}$ are the $p_i$ for which the denominator vanishes, i.e. $z_{i1} = p_{i-1} + n_i - m_{i-1} + t_{j_{i-1}} - t_{j_i}$ and $z_{i2} = p_{i+1} + m_i - n_{i+1} + t_{j_{i+1}} - t_{j_i}$ for $i = \{1, \ldots, l\}$, where we identified 0 with $l$ and $l+1$ with 1. Note that the cases with $p_i = z_{ij}$ for a $j \in \{1, 2\}$ or $|p_i - p_{i-1} + m_{i-1} - n_i + t_{j_i} - t_{j_{i-1}}| \geq N/2$ are again of smaller or equal order. Recall the treatment of (6.21). If we set $A_i = [0, N-1] \backslash ([z_{i1} - 1, z_{i1} + 1] \cup [z_{i2} - 1, z_{i2} + 1])$ for $i = \{1, \ldots, l\}$, the final line of the previous display can be bounded by

$$\sum_{p_l=0}^{N-1} \int_{A_1 \times \ldots \times A_{l-1}} \frac{1}{|p_1 - p_l + m_l - n_1 + t_{j_1} - t_{j_l}|} \prod_{i=2}^{l} \frac{1}{|p_i - p_{i-1} + m_{i-1} - n_i + t_{j_i} - t_{j_{i-1}}|} d(p_1, \ldots, p_{l-1})$$
$$\leq \sum_{p_l=0}^{N-1} \int_{A_2 \times \ldots \times A_{l-1}} \frac{\log |p_1 - p_l + m_l - n_1 + t_{j_1} - t_{j_l}| + \log |p_2 - p_1 + m_1 - n_2 + t_{j_2} - t_{j_1}|}{|p_2 - p_l + t_{j_2} - t_{j_l} + m_l - n_1 + m_1 - n_2|} \Bigg|_{\partial A_1}$$
$$\times \prod_{i=3}^{l} \frac{1}{|p_i - p_{i-1} + m_{i-1} - n_i + t_{j_i} - t_{j_{i-1}}|} d(p_2, \ldots, p_{l-1}).$$

where we considered partial fractions again and with the same notation as before. The conditions on $p_i, m_i, n_i$ and $j_i$ imply that the arguments of the log-functions are between 1 and $2N$, so also smaller than $2lN$. Thus the above term bounded by

$$\sum_{p_l=0}^{N-1} \int_{A_2 \times \ldots \times A_{l-1}} \frac{\log(2lN)}{|p_2 - p_l + t_{j_2} - t_{j_l} + m_l - n_1 + m_1 - n_2|} \prod_{i=3}^{l} \frac{1}{|p_i - p_{i-1} + m_{i-1} - n_i + t_{j_i} - t_{j_{i-1}}|} d(p_2, \ldots, p_{l-1}).$$

Using this argumentation also in the integration over $p_2, \ldots, p_{l-1}$, we can bound $V_T(\tilde{\nu})$ (up to a constant) by

$$\frac{1}{T^l} \sum_{j_1=1}^{M} \sum_{\substack{m_1,n_1,\ldots,m_l,n_l=0 \\ m_1-n_1+\ldots+m_l-n_l \neq 0 \\ m_i,n_i \neq 0}}^{\infty} \sum_{\substack{j_2,\ldots,j_M=1 \\ |n_{i+1}-m_i+t_{j_i}-t_{j_{i+1}}|<N \\ |n_1-m_l+t_{j_l}-t_{j_1}|<N}}^{M} \frac{1}{|m_1|^{1-d}} \cdots \frac{1}{|n_l|^{1-d}} \frac{1}{|m_1 - n_1 + \ldots + m_l - n_l|} \sum_{p_1=0}^{N-1} \log(2lN)^{l-1}$$

where all the differences of $p_i$- and $t_{j_i}$-terms vanish in a telescoping sum. Note that for $T$ large enough, the conditions

$$|n_{i+1} - m_i + t_{j_i} - t_{j_{i+1}}| < N \quad \text{and} \quad |n_1 - m_l + t_{j_l} - t_{j_1}| < N \tag{6.25}$$

can only be satisfied if $|n_{i+1} - m_i| \leq 2T$ for $i \in \{1, \ldots, l\}$ where we identified $l+1$ with 1. Therefore the above term is smaller or equal to

$$\frac{1}{T^l} \sum_{j_1=1}^{M} \sum_{\substack{m_1,n_1,\ldots,m_l,n_l=0 \\ (+)_{\neq,l}}}^{\infty} \sum_{\substack{j_2,\ldots,j_M=1 \\ (6.25)}}^{M} \frac{1}{|m_1|^{1-d}} \cdots \frac{1}{|n_l|^{1-d}} \frac{1}{|m_1 - n_1 + \ldots + m_l - n_l|} \sum_{p_1=0}^{N-1} \log(2lN)^{l-1}$$

where $(+)_{\neq,l}$ was defined in Lemma 6.4 e) and we now have $n = 2T$. As in the proof of part b), it can be seen that if $j_1, m_i, n_i$ are chosen, there are only finitely many possible choices for $j_2, \ldots, j_l$ because of the conditions (6.25). By using this and Lemma 3.13 e), we finally obtain

$$V_T(\tilde{\nu}) = O(T^{1-l} \log(N)^{l-1} \log(T) T^{2Dl-4D}) = O\left(T^{(1-4D)-l(1/2-2D)-l/2} \log(T)^l\right)$$

which is of order $o(1/T^{l/2})$ for $l \geq 3$ and $D < 1/4$. $\square$

**Proof of d):** Analogously to the proof of Theorem 5.1 in Preuß et al. (2012) we show the claim by proving

$$|\text{cum}_l(\hat{G}_T(v_1, \omega_1) - \hat{G}_T(v_2, \omega_2))| \leq (2l)! C^l \rho_{2,T,D}\left((v_1, \omega_1), (v_2, \omega_2)\right)^l \quad \forall l \in I\!N.$$

We assume without loss of generality that $l$ is even since the case for odd $l$ follows in the same way. In order to simplify technical arguments we furthermore define $\phi_{v,\omega,T}(u, \lambda) := \phi_{v,\omega,T}(u, -\lambda)$ for $u \in [0, 1]$ and $\lambda \in [-\pi, 0]$. Due to the symmetry of $I_N(u, \lambda)$ in $\lambda$ we then obtain that the $l$-th cumulant of $\hat{G}_T(v_1, \omega_1) - \hat{G}_T(v_2, \omega_2))$ is given by

$$\frac{1}{2^l T^{l/2}} \sum_{j_1,\ldots,j_l=1}^{M} \sum_{k_1,\ldots,k_l=-\lfloor (N-1)/2 \rfloor}^{N/2} (\phi_{v_1,\omega_1,T}(u_{j_1}, \lambda_{k_1}) - \phi_{v_2,\omega_2,T}(u_{j_1}, \lambda_{k_1})) \cdots (\phi_{v_1,\omega_1,T}(u_{j_l}, \lambda_{k_l}) - \phi_{v_2,\omega_2,T}(u_{j_l}, \lambda_{k_l}))$$
$$\times \frac{1}{(2\pi N)^l} \sum_{p_1,\ldots,q_l=0}^{N-1} \sum_{m_1,\ldots,n_l=0}^{\infty} \psi_{m_1}(u_{j_1}) \cdots \psi_{n_l}(u_{j_l}) \exp(-i\lambda_{k_1}(p_1 - q_1)) \cdots \exp(-i\lambda_{k_l}(p_l - q_l))$$
$$\times \text{cum}(Z_{t_{j_1}-N/2+1+p_1-m_1} Z_{t_{j_1}-N/2+1+q_1-n_1}, \ldots, Z_{t_{j_l}-N/2+1+p_l-m_l} Z_{t_{j_l}-N/2+1+q_l-n_l})(1 + o(1))$$

Set $\phi_{1,2,T}(u, \lambda) := \phi_{v_1,\omega_1,T}(u, \lambda) - \phi_{v_2,\omega_2,T}(u, \lambda)$. We restrict ourselves again to the indecomposable partition $\tilde{\nu}$ defined in (6.23) and call the corresponding summand $V_{2,T}(\tilde{\nu})$. Then as in the proof of Theorem 5.1 in Preuß et al. (2012) we see that

$$0 \leq p_i + m_i - n_i + t_{j_i} - t_{j_{i+1}} \leq N - 1 \qquad \text{for } i \in \{1, 3, 5, \ldots, l-3, l-1\} \tag{6.26}$$

must be satisfied and that $V_{2,T}(\tilde{\nu})$ is bounded by $\sqrt{J_{1,T}J_{2,T}}$ with

$$J_{1,T} = \frac{1}{2^l T^{l/2}} \sum_{j_1,\ldots,j_l=1}^{M} \sum_{k_1,k_3,\ldots,k_{l-1}=-\lfloor (N-1)/2 \rfloor}^{N/2} \phi_{1,2,T}^2(u_{j_1},\lambda_{k_1})\phi_{1,2,T}^2(u_{j_3},\lambda_{k_3})^2 \cdots \phi_{1,2,T}^2(u_{j_{l-1}},\lambda_{k_{l-1}})^2$$
$$\frac{1}{(2\pi N)^l} \sum_{p_1,p_3,\ldots,p_{l-1}=0}^{N-1} \sum_{\tilde{p}_1,\tilde{p}_3,\ldots,\tilde{p}_{l-1}=0}^{N-1} \sum_{\substack{m_1,n_1,m_3,n_3,\ldots,m_{l-1},n_{l-1}=0 \\ (6.26)}}^{\infty} \sum_{\substack{\tilde{m}_1,\tilde{n}_1,\tilde{m}_3,\tilde{n}_3,\ldots,\tilde{m}_{l-1},\tilde{n}_{l-1}=0 \\ \widetilde{(6.26)}}}^{\infty}$$
$$\exp(-i\lambda_{k_1}(p_1-\tilde{p}_1))\exp(-i\lambda_{k_3}(p_3-\tilde{p}_3))\cdots\exp(-i\lambda_{k_{l-1}}(p_{l-1}-\tilde{p}_{l-1}))$$
$$\psi_{m_1}(u_{j_2})\psi_{n_1}(u_{j_1})\cdots\psi_{m_{l-1}}(u_{j_l})\psi_{n_{l-1}}(u_{j_{l-1}})\psi_{\tilde{m}_1}(u_{j_2})\psi_{\tilde{n}_1}(u_{j_1})\cdots\psi_{\tilde{m}_{l-1}}(u_{j_l})\psi_{\tilde{n}_{l-1}}(u_{j_{l-1}})$$
$$\sum_{k_2,k_4,\ldots,k_l=-\lfloor (N-1)/2 \rfloor}^{N/2} \exp(-i\lambda_{k_2}(\tilde{p}_1-p_1+n_1-m_1+\tilde{m}_1-\tilde{n}_1))\exp(-i\lambda_{k_4}(\tilde{p}_3-p_3+n_3-m_3+\tilde{m}_3-\tilde{n}_3))$$
$$\cdots\exp(-i\lambda_{k_l}(\tilde{p}_{l-1}-p_{l-1}+n_{l-1}-m_{l-1}+\tilde{m}_{l-1}-\tilde{n}_{l-1})) \tag{6.27}$$

and $J_{2,T}$ being defined for even $p_i, m_i, n_i$. Here, the condition $\widetilde{(6.26)}$ says that (6.26) holds but with the $p_i, m_i, n_i$ replaced by $\tilde{p}_i, \tilde{m}_i, \tilde{n}_i$. The identity (6.22) implies that in (6.27) the restrictions

$$\tilde{p}_i = p_i + m_i - n_i + \tilde{n}_i - \tilde{m}_i \quad \text{and} \quad 0 \le p_i + m_i - n_i + \tilde{n}_i - \tilde{m}_i \le N-1 \text{ for odd } i \tag{6.28}$$

must be fulfilled and that $J_{1,T}$ therefore equals

$$\frac{1}{(4\pi)^l (TN)^{l/2}} \sum_{j_1,\ldots,j_l=1}^{M} \sum_{k_1,k_3,\ldots,k_{l-1}=-\lfloor (N-1)/2 \rfloor}^{N/2} \phi_{1,2,T}(u_{j_1},\lambda_1)^2\phi_{1,2,T}(u_{j_3},\lambda_3)^2\cdots\phi_{1,2,T}(u_{j_{l-1}},\lambda_{l-1})^2$$
$$\sum_{p_1,\ldots,p_{l-1}=0}^{N-1} \sum_{\substack{m_1,\ldots,n_{l-1}=0 \\ (6.26)}}^{\infty} \sum_{\substack{\tilde{m}_1,\ldots,\tilde{n}_{l-1}=0 \\ (6.28)}}^{\infty} \psi_{m_1}(u_{j_2})\cdots\psi_{n_{l-1}}(u_{j_{l-1}})\psi_{\tilde{m}_1}(u_{j_2})\cdots\psi_{\tilde{n}_{l-1}}(u_{j_{l-1}})$$
$$\exp(-i\lambda_{k_1}(n_1-m_1+\tilde{m}_1-\tilde{n}_1))\cdots\exp(-i\lambda_{k_{l-1}}(n_{l-1}-m_{l-1}+\tilde{m}_{l-1}-\tilde{n}_{l-1})).$$

A factorisation yields $J_{1,T} = L_{1,T} \times L_{3,T} \times \cdots \times L_{l-1,T}$ with

$$L_{i,T} := \frac{1}{8\pi^2 T} \sum_{j_i=1}^{M} \sum_{k_i=-\lfloor (N-1)/2 \rfloor}^{N/2} \phi_{1,2,T}(u_{j_i},\lambda_i)^2 \sum_{\substack{m_i,n_i,\tilde{m}_i,\tilde{n}_i=0 \\ |m_i-n_i+\tilde{n}_i-\tilde{m}_i|<N}}^{\infty} \frac{1}{N} \sum_{\substack{p_i=0 \\ 0\le p_i+m_i-n_i+\tilde{n}_i-\tilde{m}_i \le N-1}}^{N-1}$$
$$\sum_{\substack{j_{i+1}=1 \\ 0\le p_i+m_i-n_i+t_{j_i}-t_{j_{i+1}} \le N-1}}^{M} \psi_{m_i}(u_{j_{i+1}})\psi_{n_i}(u_{j_i})\psi_{\tilde{m}_i}(u_{j_{i+1}})\psi_{\tilde{n}_i}(u_{j_i})\exp(-i\lambda_{k_i}(n_i-m_i+\tilde{m}_i-\tilde{n}_i)).$$

Employing the same arguments as in the proof of a) – c) we see that $j_{i+1} = j_i$ must hold and that we can skip all conditions on $m_i, n_i, \tilde{m}_i, \tilde{n}_i, p_i$. With (2.10) we then obtain the following bound for $L_{i,T}$, namely

$$\frac{1}{T}\sum_{j=1}^{M}\sum_{k=1}^{N/2}\phi_{1,2,T}(u_j,\lambda_k)^2\frac{1}{\lambda_k^{4D}},$$

and this yields $J_{1,T} \leq C^{l/2}\rho_{2,T,D}((v_1,\omega_1),(v_2,\omega_2))$. Since the same upper bound is obtained for $J_{2,T}$ the claim follows analogously to the proof of Theorem 5.1 in Preuß et al. (2012) by employing that $(2l)!2^l$ is an upper bound for the number of indecomposable partitions. $\square$

## 6.2 Proof of Theorem 6.2

The proof works in the same way as the proof of Theorem 6.1 but by employing Lemma 4.3 instead of (2.5) in order to keep error terms uniformly small in probability.

**Proof of a):** At first note that the coefficents in the $MA(\infty)$ representations (4.7) and (4.10) do not depend on the time. Thus, if we write $I_N^*(u,\lambda)$ for the bootstrap analogon of $I_N(u,\lambda)$, we obtain

$$\mathbb{E}\left(\hat{G}_T^*(v,\omega)/\sqrt{T}\Big|X_{1,T},\ldots,X_{T,T}\right) = \mathbb{E}\Big(\frac{1}{T}\sum_{j=1}^{M}\sum_{k=1}^{N/2}\phi_{v,\omega,T}(u_j,\lambda_k)I_N^*(u_j,\lambda_k)|X_{1,T},\ldots,X_{T,T}\Big)$$
$$= \frac{1}{T}\sum_{j=1}^{M}\sum_{k=1}^{N/2}\phi_{v,\omega,T}(u_j,\lambda_k)\frac{1}{2\pi N}\sum_{p,q=0}^{N-1}\sum_{l,m=0}^{\infty}\hat{\psi}_{l,p}\hat{\psi}_{m,q}\mathbb{E}(Z^*_{t_j-N/2+1+p-m}Z^*_{t_j-N/2+1+q-l})\exp(-i\lambda_k(p-q)).$$

The $\hat{\psi}_{l,p}$ possess no time dependence, thus the above expression equals zero by definition of $\phi_{v,\omega,T}$. The same result holds for $\hat{G}^*_{T,2}(v,\omega)$.

**Proof of b):** Because of part a) we obtain

$$\mathrm{Var}\Big((\hat{G}_T^*(v,\omega)-\hat{G}^*_{T,2}(v,\omega))1_{A_T(\alpha)}\Big)/T = \mathbb{E}\Big(\mathrm{Var}(\hat{G}_T^*(v,\omega)-\hat{G}^*_{T,2}(v,\omega)|X_{1,T},\ldots,X_{T,T})1_{A_T(\alpha)}\Big)/T$$
$$= \frac{1}{T^2}\sum_{j_1,j_2=1}^{M}\sum_{k_1,k_2=1}^{N/2}\phi_{v_1,\omega_1,T}(u_{j_1},\lambda_{k_1})\phi_{v_2,\omega_2,T}(u_{j_2},\lambda_{k_2})\frac{1}{(2\pi N)^2}\sum_{p_1,p_2,q_1,q_2=0}^{N-1}\sum_{m_1,m_2,l_1,l_2=0}^{\infty}\mathbb{E}(\hat{\psi}_{m_1,l_1,m_2,l_2,p}1_{A_T(\alpha)})$$
$$\times\ \mathrm{cum}(Z^*_{e_{j_1,N}+p_1-m_1}Z^*_{e_{j_1,N}+q_1-l_1},Z^*_{e_{j_2,N}+p_2-m_2}Z^*_{e_{j_2,N}+q_2-l_2})\exp(-i\lambda_{k_1}(p_1-q_1))\exp(-i\lambda_{k_2}(p_2-q_2))$$

with $\hat{\psi}_{m_1,l_1,m_2,l_2,p} = (\hat{\psi}_{m_1,p}\hat{\psi}_{l_1,p}-\psi_{m_1,p}\psi_{l_1,p})(\hat{\psi}_{m_2,p}\hat{\psi}_{l_2,p}-\psi_{m_2,p}\psi_{l_2,p})$. By using

$$\hat{\psi}_{m_1,p}\hat{\psi}_{l_1,p}-\psi_{m_1,p}\psi_{l_1,p} = (\hat{\psi}_{m_1,p}-\psi_{m_1,p})\psi_{l_1,p}+(\hat{\psi}_{l_1,p}-\psi_{l_1,p})\hat{\psi}_{m_1,p}$$

and the analogue for $\hat{\psi}_{m_2,p}\hat{\psi}_{l_2,p}-\psi_{m_2,p}\psi_{l_2,p}$, we can divide the above expression into the sum of four terms. For the sake of brevity details are presented only for the first one. By using (6.15) the corresponding summand splits into two terms and we restrict ourselves to the first one which we denote with $V^*_{T,1}$. As in the proof of Theorem 6.1 b) we then obtain error terms $V^{+,*}_{T,1}$, $V^{j_2=j_1+,*}_{T,1}$ which are defined as $V^{+}_{T,1}$, $V^{j_2=j_1+}_{T,1}$ but with the $\psi_{m_1}(u_{j_1})\psi_{l_1}(u_{j_1})\psi_{m_2}(u_{j_2})\psi_{l_2}(u_{j_2})$ replaced by $\mathbb{E}\Big((\hat{\psi}_{m_1,p}-\psi_{m_1,p})(\hat{\psi}_{m_2,p}-\psi_{m_2,p})\psi_{l_1,p}\psi_{l_2,p}1_{A_T(\alpha)}\Big)$. In the following we will demonstrate that

$$\max(V^{+,*}_{T,1},V^{j_2=j_1+,*}_{T,1}) = O\left(p^8\log(T)^3\log(N)^2N^{\max(4\underline{d}-1,0)+\alpha}T^{-2}\right).$$

The proof is similar to the one of (6.17) and (6.18) up to employing (6.3). Let us demonstrate this concept in the treatment of $V_{N,T,3}^{+,*}$ which is bounded by

$$\frac{2}{(2\pi T)^2}\sum_{j_1=1}^{M}\sum_{\substack{m_1,m_2,l_1,l_2=0\\|m_1-l_2|\geq N}}^{\infty}\sum_{q_2=0}^{N-1}\sum_{\substack{j_2=1\\0\leq q_2+m_1-l_2+\Delta t\leq N-1}}^{M}\mathbb{E}|(\hat{\psi}_{m_1,p}-\psi_{m_1,p})(\hat{\psi}_{m_2,p}-\psi_{m_2,p})\psi_{l_1,p}\psi_{l_2,p}1_{A_T(\alpha)}|$$
$$\times\sum_{\substack{q_1=0\\0\leq q_1+m_2-l_1-\Delta t\leq N-1}}^{N-1}\Big|\frac{1}{N}\sum_{k_1=1}^{N/2}\phi_{v_1,\omega_1,T}(u_{j_1},\lambda_{k_1})\exp(-i\lambda_{k_1}(q_2-q_1+\Delta t+m_1-l_2))\Big| \tag{6.19}$$
$$\times\Big|\frac{1}{N}\sum_{k_2=1}^{N/2}\phi_{v_2,\omega_2,T}(u_{j_2},\lambda_{k_2})\exp(-i\lambda_{k_2}(q_1-q_2-\Delta t+m_2-l_1))\Big|;$$

compare with (6.20). In the proof of Theorem 6.1 b) we have shown that $V_{N,T,3}^{+}$ is of order $O(\log(N)^2/(TN^{1-4D}))$ by employing (6.5). Here we use (6.3) instead and combine it with the fact that $|\hat{\underline{d}}-\underline{d}|<\alpha/4$ on $A_T(\alpha)$ to obtain

$$|\hat{\psi}_{l,p}-\psi_l|\leq Cp^4\log(T)^{3/2}T^{-1/2}|l|^{\alpha/4+\underline{d}-1}\quad\forall l\in I\!N. \tag{6.29}$$

This together with (6.5) and Assumption 4.2 implies

$$|\hat{\psi}_{l,p}|\leq C|l|^{\alpha/4+\underline{d}-1}\quad\forall l,p\in I\!N. \tag{6.30}$$

Thus the role of $D$ is played by $\underline{d}+\alpha/4$ now, and using (6.29) and (6.30) instead of (6.5) we obtain

$$V_{N,T,3}^{+,*}\leq Cp^8\log(T)^3T^{-1}\times\log(N)^2T^{-1}N^{4\underline{d}+\alpha-1}\leq Cp^8\log(T)^3\log(N)^2N^{\max(4\underline{d}-1,0)+\alpha}T^{-2}.$$

Similarly, the subsequent steps in the proof of Theorem 6.1 b) reveal that $V_{T,1}^{*}$ becomes

$$V_{T,1}^{*}=\frac{1}{T^2}\sum_{j=1}^{M}\sum_{k=1}^{N/2}\phi_{v_1,\omega_1,T}(u_j,\lambda_k)\phi_{v_2,\omega_2,T}(u_j,\lambda_k)\frac{1}{(2\pi)^2}\sum_{\substack{m_1,m_2,l_1,l_2=0\\(+)}}^{\infty}\exp(-i\lambda_k(m_1-l_2+m_2-l_1))$$
$$\times\,\mathbb{E}\left((\hat{\psi}_{m_1,p}-\psi_{m_1,p})(\hat{\psi}_{m_2,p}-\psi_{m_2,p})\psi_{l_1,p}\psi_{l_2,p}1_{A_T(\alpha)}\right)+O\Big(p^8\log(T)^3\log(N)^2N^{\max(4\underline{d}-1,0)+\alpha}T^{-2}\Big).$$

In the proof of Theorem 6.1 b), the analogue of the first quantity on the right hand side above is the main term contributing to the variance. Here, however, it is of the same order as the error terms. This can be seen using (6.29) and (6.30) again plus Lemma 6.4 d).

**Proof of c):** If we employ (6.29) and (6.30) as in the proof of part b) and follow the arguments in the proof of Theorem 6.1 d), we obtain

$$\mathbb{E}\Big(|(\hat{G}_T^*(v_1,\omega_1)-\hat{G}_{T,2}^*(v_1,\omega_1))-(\hat{G}_T^*(v_2,\omega_2)-\hat{G}_{T,2}^*(v_2,\omega_2))|^k1_{A_T(\alpha)}\Big)$$
$$\leq(2k)!C^k\rho_{2,T,\underline{d}+\alpha/4}\left((v_1,\omega_1),(v_2,\omega_2)\right)^k\left(p^8\log(T)^3T^{-1}\right)^{k/2},$$

where $\rho_{2,T,\underline{d}+\alpha/4}(\cdot,\cdot)$ corresponds to the metric defined in (6.1) but with $D$ replaced by $\underline{d}+\alpha/4$ due to $|\hat{\underline{d}}-\underline{d}| \leq \alpha/4$. The claim then follows from

$$\rho_{2,T,\underline{d}+\alpha/4}\left((v_1,\omega_1),(v_2,\omega_2)\right) \leq C\tilde{\rho}\left((v_1,\omega_1),(v_2,\omega_2)\right)\sqrt{N^{\max(4\underline{d}-1,0)+\alpha}}.$$

□

### 6.3 Proofs from Section 3 and 4

**Proof of Theorem 3.1:** To show weak convergence we have to prove the following two claims [see van der Vaart and Wellner (1996), Theorem 1.5.4 and 1.5.7]:

(1) Convergence of the finite dimensional distributions

$$(\hat{G}_T(y_j))_{j=1,\ldots,K} \xrightarrow{D} (G(y_j))_{j=1,\ldots,K} \tag{6.31}$$

where $y_j = (v_j,\omega_j) \in [0,1]^2$ $(j=1,\ldots,K)$ and $K \in I\!N$.

(2) Stochastic equicontinuity, i.e.

$$\forall \eta,\varepsilon > 0 \quad \exists \delta > 0: \lim_{T\to\infty} P\Big(\sup_{y_1,y_2\in[0,1]^2:\rho_{2,D}(y_1,y_2)<\delta} |\hat{G}_T(y_1)-\hat{G}_T(y_2)| > \eta\Big) < \varepsilon,$$

where

$$\rho_{2,D}\left(y_1,y_2\right) := \left(\frac{1}{2\pi}\int_0^1\int_0^\pi (\phi_{v_1,\omega_1}(u,\lambda)-\phi_{v_2,\omega_2}(u,\lambda))^2\frac{1}{\lambda^{4D}}d\lambda du\right)^{1/2}$$

[with the functions $\phi_{v,\omega}$ defined in (6.2) and $y_i = (v_i,\omega_i)$ for $i=1,2$].

The claim (6.31) can be deduced from Theorem 6.1 a)–c), while stochastic equicontinuity can be concluded along the lines of the corresponding result in Preuß et al. (2012). □

**Proof of Lemma 4.3:** If we denote with $\psi_{l,p}$ the coefficents in the MA($\infty$) representation of the process $(1-B)^{-\underline{d}}Y_t^{AR}(p)$ and with $\eta_l$ the coefficent which appears if we replace $\hat{\underline{d}}$ with $\underline{d}$ in $\hat{\eta}_l$, we obtain with (4.8)

$$\hat{\psi}_{l,p}-\psi_{l,p} = \sum_{k=0}^{l}\left(\hat{c}_{k,p}\hat{\eta}_{l-k}-c_{k,p}\eta_{l-k}\right) = \sum_{k=0}^{l}\left(\hat{c}_{k,p}-c_{k,p}\right)\hat{\eta}_{l-k} + \sum_{k=0}^{l}c_{k,p}\left(\hat{\eta}_{l-k}-\eta_{l-k}\right). \tag{6.32}$$

We start with the treatment of the first term and let $l \geq 1$. By employing (4.6) we can apply Cauchy's inequality for holomorphic functions analogously to the proof of Lemma 2.5 in Kreiß et al. (2011) to obtain

$$|\hat{c}_{l,p}-c_{l,p}| = \frac{p}{(1+1/p)^l}\sqrt{\log(T)/T}O_P(1), \quad \text{uniformly in } p,l \in I\!N.$$

With this bound we get $\sum_{k=0}^\infty k^2|\hat{c}_{k,p}-c_{k,p}| = O_P(p^4\sqrt{\log(T)/T})$ which directly yields

$$|\hat{c}_{k,p}-c_{k,p}| = O_P(p^4\sqrt{\log(T)/T}/k^2). \tag{6.33}$$

Using (2.16) and properties of the Gamma function we obtain $\hat{\eta}_l \leq C/l^{1-\hat{\underline{d}}}$, uniformly in $\hat{\underline{d}}$. Therefore we see with (6.33) that the first term in (6.32) is some $O_P(p^4\sqrt{\log(T)/T}/l^{1-\hat{\underline{d}}})$. This works again by replacing the sum through its approximating integral (for $k \neq 0, l$) and applying (6.6) as in the proof of Lemma 6.4. Concerning the second summand in (6.32) note that (4.4) implies $c_{k,p} = O_P\left(1/k^7\right)$, uniformly in $p$. If we combine this with (4.6) and the mean value theorem we obtain that the second summand in (6.32) is of order $O_P(\sqrt{\log(T)/T}/l^{1-\max(\hat{\underline{d}},\underline{d})})$.

Thus to complete the proof it remains to consider $|\psi_{l,p} - \psi_l|$ which is bounded through

$$\sum_{k=0}^{l} |c_{k,p} - c_k||\eta_{l-k}|,$$

where $c_k$ are the coefficents in the MA($\infty$) representation of the process $Y_t = (1-B)^{\underline{d}} X_t$, see Assumption 4.1. It follows from (4.2) and Lemma 2.4 in Kreiß et al. (2011) that $|c_{k,p} - c_k| = O(1/(k^2 p^5))$ which implies that $|\psi_{l,p} - \psi_l|$ is of order $O(1/(l^{1-\underline{d}} p^5))$. This yields the claim since $\sqrt{T/\log(T)} = o(p^9)$. $\square$

**Proof of Theorem 4.4:** First, $\sup_{(v,\omega)\in[0,1]^2} |\hat{G}_{T,a}(v,\omega)|$ and $\sup_{(v,\omega)\in[0,1]^2} |\hat{G}^*_{T,a}(v,\omega)|$ have the same distribution, because $\psi_l = \psi_l(u)$ for all $u \in [0,1]$ under the null hypothesis and since the $Z_t$ and $Z_t^*$ are both independent and standard normal distributed. Furthermore, the proof of Theorem 6.1 reveals that we have

$$\mathbb{E}\Big(\sup_{(v,\omega)\in[0,1]^2} |\hat{G}^*_{T,a}(v,\omega)|^2\Big)^{-1/2} \sup_{v,\omega} |\hat{G}_{T,a}(v,\omega)| = \mathbb{E}\Big(\sup_{(v,\omega)\in[0,1]^2} |\hat{G}^*_{T,a}(v,\omega)|^2\Big)^{-1/2} \sup_{v,\omega} |\hat{G}_{T}(v,\omega)| + o_P(1)$$

due to (2.3), so let us focus on part c). We show

$$\sqrt{T/\mathbb{E}\Big(\sup_{(v,\omega)\in[0,1]^2} \hat{G}^*_{T,a}(v,\omega)^2\Big)} \sup_{(v,\omega)\in[0,1]^2} \Big|\hat{G}^*_{T}(v,\omega) - \hat{G}^*_{T,a}(v,\omega)\Big| / \sqrt{T} = o_P(1)$$

Note that (2.10), Assumption 4.1 and Theorem 6.1 a), b) prove

$$C_1 N^{\max(4\underline{d}-1,0)}/T \leq \mathbb{E}(\hat{G}^*_{T,a}(v,\omega)^2)/T \leq C_2(N^{\max(4\underline{d}-1,0)} + \log(N) 1_{\{\underline{d}=1/4\}})/T, \tag{6.34}$$

Thus d) follows, and according to Newey (1991) we have to show the following two claims:

(1) For every $v, \omega \in [0,1]$ we have

$$(\hat{G}^*_{T}(v,\omega) - \hat{G}^*_{T,a}(v,\omega))/\sqrt{N^{\max(4\underline{d}-1,0)}} = o_P(1).$$

(2) For every $\eta, \varepsilon > 0$ there exists a $\delta > 0$ such that

$$\lim_{T\to\infty} P\Big(\sup_{y_1,y_2\in[0,1]^2:\tilde{\rho}(y_1,y_2)<\delta} |(\hat{G}^*_{T}(y_1) - \hat{G}^*_{T,a}(y_1)) - (\hat{G}^*_{T}(y_2) - \hat{G}^*_{T,a}(y_2))|/\sqrt{N^{\max(4\underline{d}-1,0)}} > \eta\Big) < \varepsilon,$$

where $y_i = (v_i, \omega_i)$ for $i = 1, 2$.

Note that Lemma 4.3 implies $P(A_T(\alpha)) \to 1$ as $T \to \infty$ for every $\alpha > 0$. Therefore part (1) follows from Theorem 6.2 a) and b) and the conditions on the grow rate of $p = p(T)$ which are specified in Assumption 4.2 by choosing $\alpha$ small. The second claim can be shown analogously to the proof of Theorem 3.1 by employing Theorem 6.2 c) instead of Theorem 6.1 d). $\square$

## References


Adak, S. (1998). Time-dependent spectral analysis of nonstationary time series. *Journal of the American Statistical Association*, 93:1488–1501.

Akaike, H. (1973). *Information theory and an extension of the maximum likelihood principle*. Budapest, Akademia Kiado, 267-281.

Beran, J. (2009). On parameter estimation for locally stationary long-memory processes. *Journal of Statistical Planning and Inference*, 139:900–915.

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.

Berkes, I., Horvarth, L., Kokoszka, P., and Shao, Q. M. (2006). On discriminating between long-range dependence and changes in mean. *Annals of Statistics*, 34:1140–1165.

Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory*. McGraw Hill, New York.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*. Springer Verlag, New York.

Chang, C. and Morettin, P. (1999). estimation of time-varying linear systems. *Statistical Inference for Stochastic Processes.*, 2:253–285.

Chen, Y., Hï¿½rdle, W., and Pigorsch, U. (2010). Localized realized volatility modeling. *Journal of the American Statistical Association*, 105(492):1376–1393.

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

Dahlhaus, R. (2009). Local inference for locally stationary time series based on the empirical spectral measure. *Journal of Econometrics*, 151:101–112.

Dahlhaus, R. and Polonik, W. (2006). Nonparametric quasi maximum likelihood estimation for Gaussian locally stationary processes. *Annals of Statistics*, 34(6):2790–2824.

Dahlhaus, R. and Polonik, W. (2009). Empirical spectral processes for locally stationary time series. *Bernoulli*, 15:1–39.

Dehling, H., Rooch, A., and Taqqu, M. S. (2011). Nonparametric change-point tests for long-range dependent data. Prepint, Ruhr-Universität Bochum.

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.

Doukhan, P., Oppenheim, G., and Taqqu, M. S. (2002). *Theory and Applications of Long-Range Dependence*. Birkhäuser, Boston.

Dwivedi, Y. and Subba Rao, S. (2010). A test for second order stationarity of a time series based on the discrete fourier transform. *Journal of Time Series Analysis.*, 32(1):68–91.

Eichler, M. (2008). Testing nonparametric and semiparametric hypotheses in vector stationary processes. *Journal of Multivariate Analysis*, 99:968–1009.

Fox, R. and Taqqu, M. S. (1987). Central limit theorems for quadratic forms in random variables having long-range dependence. *Probability Theory and Related Fields*, 74:213–240.

Fryzlewicz, P., Sapatinas, T., and Subba Rao, S. (2006). A Haar-Fisz technique for locally stationary volatility estimation. *Biometrika*, 93:687–704.

Granger, C. W. J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1:15–29.

Henry, M. and Zaffaroni, P. (2002). The long-range dependence paradigm for macroeconomics and finance. In *Theory and Applications of Long-Range Dependence (P. Doukhan, G. Oppenheim and M. S. Taqqu)*, pages 417–438. Birkhäuser, Boston.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, 68:165–176.

Jensen, M. and Whitcher, B. (2000). Time-varying long memory in volatility: Detection and estimation with wavelets. Technical report, EURANDOM.

Kokoszka, P. and Mikosch, T. (1997). The integrated periodogram for long-memory processes with finite or infinite variance. *Stochastic Process. Appl.*, 66(1):55–78.

Kokoszka, P. S. and Taqqu, M. S. (1995). Fractional ARIMA with stable innovations. *Stochastic Processes and Their Applications*, 60:19–47.

Kreiß, J.-P. (1988). *Asymptotic statistical inference for a class of stochastic processes.* Habilitationsschrift, Fachbereich Mathematik, Universität Hamburg.

Kreiß, J.-P. and Paparoditis, E. (2011). Bootstrapping locally stationary processes. Technical report.

Kreiß, J.-P., Paparoditis, E., and Politis, D. N. (2011). On the range of the validity of the autoregressive sieve bootstrap. *to appear in Annals of statistics.*

Mikosch, T. and Starica, C. (2004). Non-stationarities in financial time series, the long range dependence and the IGARCH effects. *The Review of Economics and Statistics*, 86:378–390.

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

Newey, W. K. (1991). Uniform convergence in probability and stochastic equicontinuity. *Econometrica*, 59(4):1161–1167.

Palma, W. (2007). *Long-Memory Time Series: Theory and Methods.* Wiley Series in Probability and Statistics.

Palma, W. and Olea, R. (2010). An efficient estimator for locally stationary Gaussian long-memory processes. *Annals of Statistics.*, 38(5):2958–2997.

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

Paparoditis, E. (2010). Validating stationarity assumptions in time series analysis by rolling local periodograms. *Journal of the American Statistical Association*, 105(490):839–851.

Park, K. and Willinger, W. (2000). *Self-Similar Network Traffic and Performance Evaluation.* Wiley, New York.

Perron, P. and Qu, Z. (2010). Long-memory and level shifts in the volatility of stock market return indices. *Journal of Business and Economic Statistics*, 23(2):275–290.

Pipiras, V. and Taqqu, M. S. (2011). *Long-Range Dependence and Self-Similarity.* Boston.

Preuß, P., Vetter, M., and Dette, H. (2012). A test for stationarity based on empirical processes. *To appear in Bernoulli.*

Roueff, F. and von Sachs, R. (2011). Locally stationary long memory estimation. *Stochastic Processes and their Applications*, 121:813–844.

Sakiyama, K. and Taniguchi, M. (2004). Discriminant analysis for locally stationary processes. *Journal of Multivariate Analysis*, 90:282–300.

Sibbertsen, P. and Kruse, R. (2009). Testing for a change in persistence under long-range dependencies. *Journal of Time Series Analysis*, 30:263–285.

Starica, C. and Granger, C. (2005). Nonstationarities in stock returns. *The Review of Economics and Statistics*, 87:503–522.

Van Bellegem, S. and von Sachs, R. (2008). Locally adaptive estimation of evolutionary wavelet spectra. *Annals of Statistics*, 36(4):1879–1924.

van der Vaart, A. and Wellner, J. (1996). *Weak Convergence and Empirical Processes.* Springer, Berlin.

von Sachs, R. and Neumann, M. H. (2000). A wavelet-based test for stationarity. *Journal of Time Series Analysis*, 21:597–613.

Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis.* HUppsala: Almqvist and Wiksell.

Zygmund, A. (1959). *Trigonometric Series.* Cambridge University Press.

| | | | $\phi = -0.9$ | | $\phi = -0.5$ | | $\phi = 0$ | | $\phi = 0.5$ | | $\phi = 0.9$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $T$ | $N$ | $M$ | 5% | 10% | 5% | 10% | 5% | 10% | 5% | 10% | 5% | 10% |
| 128 | 16 | 8 | 0.131 | 0.179 | 0.064 | 0.098 | 0.054 | 0.087 | 0.078 | 0.122 | 0.104 | 0.17 |
| 128 | 8 | 16 | 0.129 | 0.167 | 0.069 | 0.11 | 0.056 | 0.102 | 0.086 | 0.127 | 0.095 | 0.151 |
| 256 | 32 | 8 | 0.093 | 0.129 | 0.056 | 0.099 | 0.039 | 0.072 | 0.051 | 0.083 | 0.087 | 0.152 |
| 256 | 16 | 16 | 0.069 | 0.107 | 0.057 | 0.088 | 0.041 | 0.086 | 0.068 | 0.124 | 0.08 | 0.118 |
| 256 | 8 | 32 | 0.067 | 0.112 | 0.046 | 0.093 | 0.046 | 0.09 | 0.077 | 0.118 | 0.051 | 0.096 |
| 512 | 64 | 8 | 0.051 | 0.099 | 0.047 | 0.086 | 0.039 | 0.087 | 0.031 | 0.07 | 0.062 | 0.108 |
| 512 | 32 | 16 | 0.058 | 0.109 | 0.048 | 0.097 | 0.043 | 0.087 | 0.051 | 0.1 | 0.077 | 0.14 |
| 512 | 16 | 32 | 0.056 | 0.109 | 0.046 | 0.085 | 0.062 | 0.115 | 0.066 | 0.112 | 0.054 | 0.122 |
| 512 | 8 | 64 | 0.052 | 0.092 | 0.05 | 0.1 | 0.033 | 0.086 | 0.065 | 0.118 | 0.041 | 0.091 |

Table 1: *Rejection probabilities of the test* (4.9) *under the null hypothesis. The data was generated according to model* (5.1) *with* $d = 0.2$, $\theta = 0$ *and different values for* $\phi$.

| | | | $\theta = -0.9$ | | $\theta = -0.5$ | | $\theta = 0.5$ | | $\theta = 0.9$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $T$ | $N$ | $M$ | 5% | 10% | 5% | 10% | 5% | 10% | 5% | 10% |
| 128 | 16 | 8 | 0.075 | 0.124 | 0.066 | 0.124 | 0.061 | 0.106 | 0.059 | 0.092 |
| 128 | 8 | 16 | 0.064 | 0.112 | 0.058 | 0.101 | 0.066 | 0.109 | 0.069 | 0.112 |
| 256 | 32 | 8 | 0.046 | 0.107 | 0.056 | 0.105 | 0.044 | 0.097 | 0.056 | 0.094 |
| 256 | 16 | 16 | 0.047 | 0.094 | 0.058 | 0.115 | 0.037 | 0.085 | 0.064 | 0.108 |
| 256 | 8 | 16 | 0.059 | 0.098 | 0.061 | 0.109 | 0.047 | 0.085 | 0.046 | 0.085 |
| 512 | 64 | 8 | 0.057 | 0.096 | 0.041 | 0.084 | 0.041 | 0.088 | 0.049 | 0.094 |
| 512 | 32 | 16 | 0.041 | 0.089 | 0.056 | 0.107 | 0.052 | 0.101 | 0.058 | 0.091 |
| 512 | 16 | 32 | 0.046 | 0.084 | 0.046 | 0.098 | 0.057 | 0.095 | 0.048 | 0.087 |
| 512 | 8 | 64 | 0.036 | 0.089 | 0.05 | 0.091 | 0.043 | 0.083 | 0.055 | 0.1 |

Table 2: *Rejection probabilities of the test* (4.9) *under the null hypothesis. The data was generated according to model* (5.1) *with* $d = 0.2$, $\phi = 0$ *and different values for* $\theta$.

| | | | $\phi = -0.9$ | | $\phi = -0.5$ | | $\phi = 0$ | | $\phi = 0.5$ | | $\phi = 0.9$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $T$ | $N$ | $M$ | 5% | 10% | 5% | 10% | 5% | 10% | 5% | 10% | 5% | 10% |
| 128 | 16 | 8 | 0.138 | 0.174 | 0.056 | 0.104 | 0.06 | 0.091 | 0.096 | 0.138 | 0.18 | 0.256 |
| 128 | 8 | 16 | 0.126 | 0.168 | 0.083 | 0.124 | 0.059 | 0.107 | 0.088 | 0.139 | 0.153 | 0.219 |
| 256 | 32 | 8 | 0.08 | 0.116 | 0.044 | 0.078 | 0.05 | 0.087 | 0.047 | 0.099 | 0.12 | 0.196 |
| 256 | 16 | 16 | 0.082 | 0.125 | 0.043 | 0.075 | 0.052 | 0.09 | 0.055 | 0.101 | 0.111 | 0.173 |
| 256 | 8 | 32 | 0.071 | 0.107 | 0.055 | 0.096 | 0.045 | 0.097 | 0.064 | 0.112 | 0.084 | 0.13 |
| 512 | 64 | 8 | 0.051 | 0.1 | 0.041 | 0.089 | 0.044 | 0.083 | 0.029 | 0.067 | 0.061 | 0.124 |
| 512 | 32 | 16 | 0.053 | 0.104 | 0.049 | 0.094 | 0.038 | 0.09 | 0.057 | 0.097 | 0.082 | 0.145 |
| 512 | 16 | 32 | 0.063 | 0.111 | 0.053 | 0.105 | 0.056 | 0.112 | 0.049 | 0.086 | 0.074 | 0.129 |
| 512 | 8 | 64 | 0.051 | 0.096 | 0.051 | 0.094 | 0.042 | 0.089 | 0.056 | 0.11 | 0.067 | 0.117 |

Table 3: *Rejection probabilities of the test* (4.9) *under the null hypothesis. The data was generated according to model* (5.1) *with* $d = 0.4$, $\theta = 0$ *and different values for* $\phi$.

| | | | $\theta = -0.9$ | | $\theta = -0.5$ | | $\theta = 0.5$ | | $\theta = 0.9$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $T$ | $N$ | $M$ | 5% | 10% | 5% | 10% | 5% | 10% | 5% | 10% |
| 128 | 16 | 8 | 0.086 | 0.136 | 0.081 | 0.13 | 0.053 | 0.084 | 0.069 | 0.099 |
| 128 | 8 | 16 | 0.085 | 0.128 | 0.065 | 0.11 | 0.069 | 0.11 | 0.073 | 0.11 |
| 256 | 32 | 8 | 0.07 | 0.116 | 0.059 | 0.096 | 0.039 | 0.07 | 0.05 | 0.096 |
| 256 | 16 | 16 | 0.069 | 0.119 | 0.076 | 0.133 | 0.053 | 0.09 | 0.04 | 0.089 |
| 256 | 8 | 16 | 0.043 | 0.087 | 0.068 | 0.111 | 0.051 | 0.099 | 0.051 | 0.112 |
| 512 | 64 | 8 | 0.052 | 0.109 | 0.037 | 0.079 | 0.051 | 0.085 | 0.046 | 0.105 |
| 512 | 32 | 16 | 0.068 | 0.119 | 0.05 | 0.103 | 0.053 | 0.099 | 0.042 | 0.095 |
| 512 | 16 | 32 | 0.056 | 0.101 | 0.054 | 0.106 | 0.045 | 0.084 | 0.056 | 0.11 |
| 512 | 8 | 64 | 0.056 | 0.102 | 0.065 | 0.101 | 0.054 | 0.098 | 0.043 | 0.082 |

Table 4: *Rejection probabilities of the test* (4.9) *under the null hypothesis. The data was generated according to model* (5.1) *with* $d = 0.4$, $\phi = 0$ *and different values for* $\theta$.

| | | | | (5.4) | | (5.5) | | (5.6) | |
|---|---|---|---|---|---|---|---|---|---|
| $T$ | $N$ | $M$ | $d$ | 5% | 10% | 5% | 10% | 5% | 10% |
| 128 | 16 | 8 | 0.2 | 0.694 | 0.811 | 0.198 | 0.303 | 0.028 | 0.084 |
| 128 | 8 | 16 | 0.2 | 0.702 | 0.824 | 0.169 | 0.266 | 0.023 | 0.071 |
| 256 | 32 | 8 | 0.2 | 0.909 | 0.968 | 0.211 | 0.332 | 0.132 | 0.262 |
| 256 | 16 | 16 | 0.2 | 0.946 | 0.978 | 0.197 | 0.312 | 0.121 | 0.3 |
| 256 | 8 | 16 | 0.2 | 0.942 | 0.98 | 0.158 | 0.264 | 0.164 | 0.32 |
| 512 | 64 | 8 | 0.2 | 0.997 | 1.0 | 0.519 | 0.791 | 0.557 | 0.742 |
| 512 | 32 | 16 | 0.2 | 0.999 | 1.0 | 0.477 | 0.702 | 0.575 | 0.764 |
| 512 | 16 | 32 | 0.2 | 1.0 | 1.0 | 0.362 | 0.564 | 0.648 | 0.808 |
| 512 | 8 | 64 | 0.2 | 1.0 | 1.0 | 0.258 | 0.39 | 0.664 | 0.823 |
| 128 | 16 | 8 | 0.4 | 0.517 | 0.659 | 0.217 | 0.326 | 0.027 | 0.056 |
| 128 | 8 | 16 | 0.4 | 0.649 | 0.769 | 0.188 | 0.262 | 0.022 | 0.067 |
| 256 | 32 | 8 | 0.4 | 0.639 | 0.771 | 0.198 | 0.308 | 0.115 | 0.246 |
| 256 | 16 | 16 | 0.4 | 0.795 | 0.903 | 0.162 | 0.292 | 0.11 | 0.271 |
| 256 | 8 | 16 | 0.4 | 0.907 | 0.963 | 0.137 | 0.236 | 0.138 | 0.312 |
| 512 | 64 | 8 | 0.4 | 0.731 | 0.861 | 0.275 | 0.525 | 0.471 | 0.652 |
| 512 | 32 | 16 | 0.4 | 0.925 | 0.974 | 0.355 | 0.602 | 0.531 | 0.718 |
| 512 | 16 | 32 | 0.4 | 0.989 | 0.995 | 0.355 | 0.564 | 0.662 | 0.784 |
| 512 | 8 | 64 | 0.4 | 0.997 | 1.0 | 0.221 | 0.386 | 0.677 | 0.819 |

Table 5: *Rejection probabilities of the test* (4.9) *for the models* (5.4)–(5.6).

| | | | $d_1(u)$ | | $d_2(u)$ | |
|---|---|---|---|---|---|---|
| $T$ | $N$ | $M$ | 5% | 10% | 5% | 10% |
| 128 | 16 | 8 | 0.058 | 0.108 | 0.037 | 0.075 |
| 128 | 8 | 16 | 0.078 | 0.129 | 0.07 | 0.114 |
| 256 | 32 | 8 | 0.054 | 0.108 | 0.049 | 0.125 |
| 256 | 16 | 16 | 0.074 | 0.147 | 0.047 | 0.109 |
| 256 | 8 | 16 | 0.094 | 0.143 | 0.085 | 0.128 |
| 512 | 64 | 8 | 0.175 | 0.288 | 0.283 | 0.439 |
| 512 | 32 | 16 | 0.131 | 0.218 | 0.218 | 0.356 |
| 512 | 16 | 32 | 0.074 | 0.145 | 0.096 | 0.179 |
| 512 | 8 | 64 | 0.104 | 0.172 | 0.099 | 0.181 |

Table 6: *Rejection probabilities of the test* (4.9) *for the models from* (5.7).

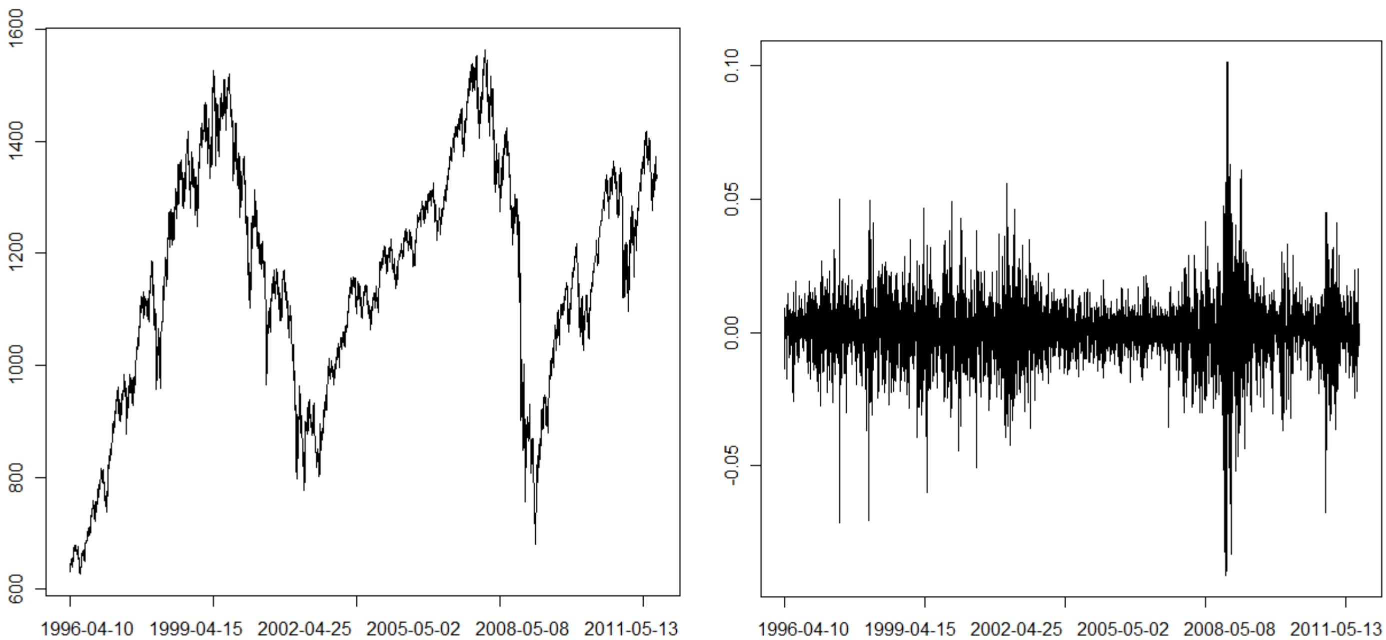


Figure 1: *The left panel displays the price of the S&P 500 between April 10th 1996 and July 13th 2012 whereas the log returns of the S&P 500 in the same period are shown in the right panel.*

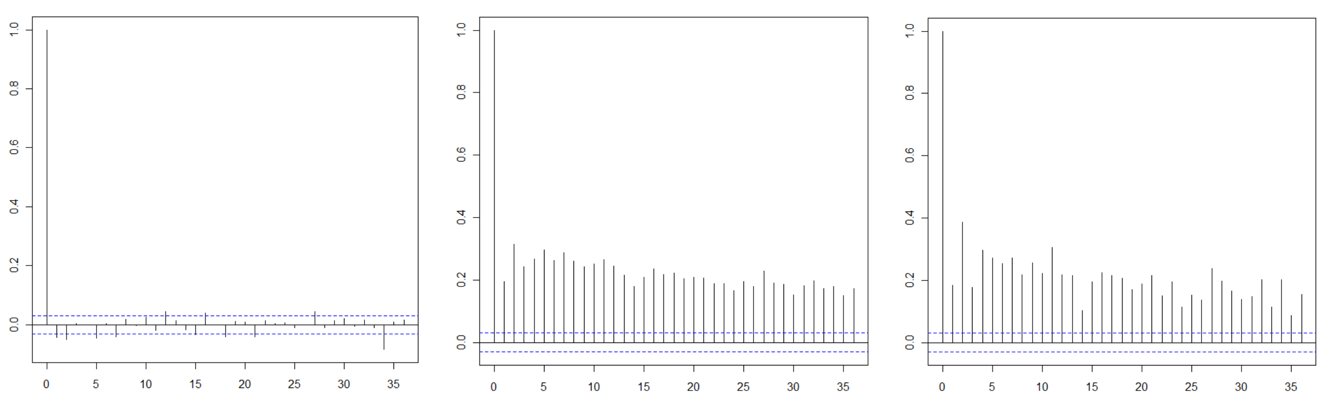

Figure 2: *Left panel: ACF (autocorrelation function) of the log returns* $X_t$*, middle panel: ACF of the absolute log returns* $|X_t|$*, right panel: ACF of the squared log returns* $X_t^2$*.*

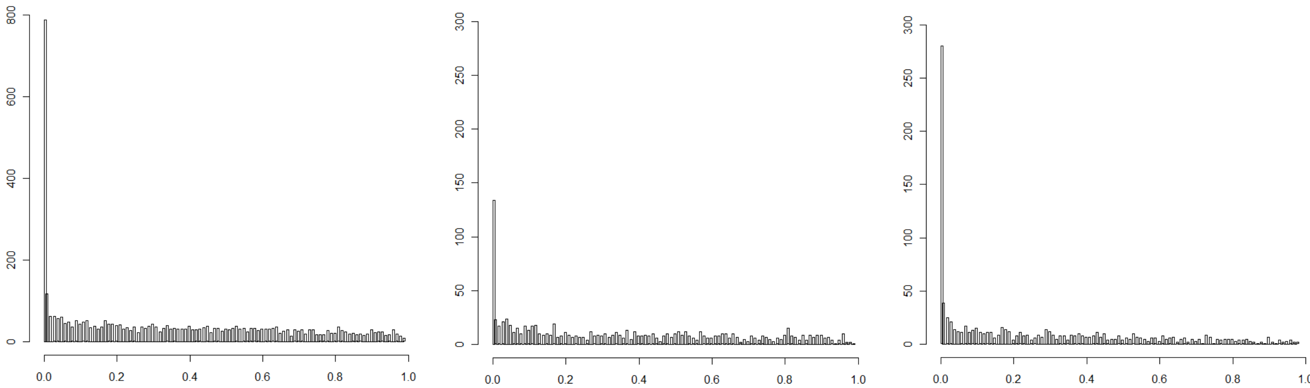


Figure 3: *The left panel displays the histogram of the p-values if the test* (4.9) *(with* $T = 64$ *and* $N = 8$*) is applied on a rolling window of the* 4096 *datapoints. In the middle panel we present the histogram of the p-values if the test* (4.9) *(with* $T = 64$ *and* $N = 8$*) is applied on a rolling window of the first* 1000 *datapoints. The right panel shows the corresponding histogram if the last* 1000 *datapoints are used.*

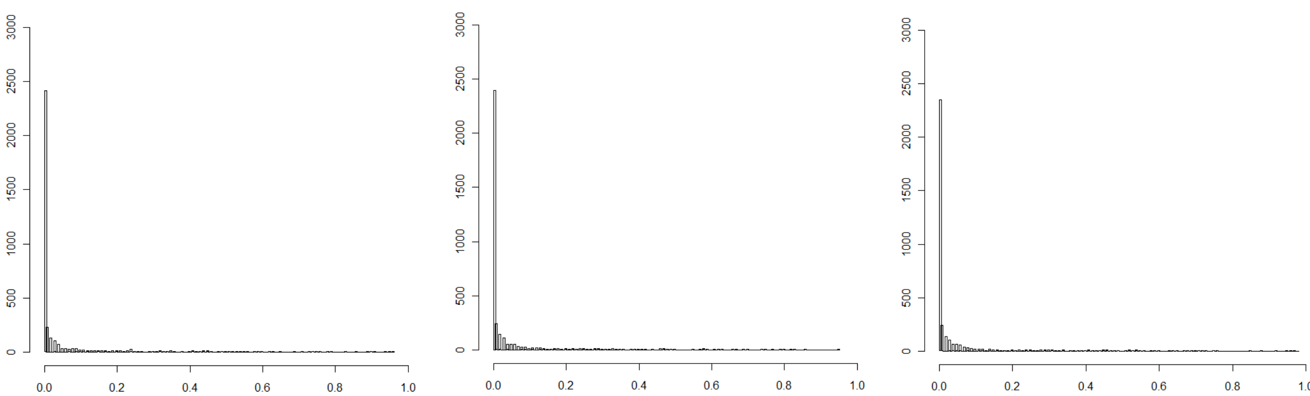


Figure 4: *Histograms of the p-values if the test* (4.9) *with* $T = 256$ *and different choice for* $N$ *is applied on a rolling window of the* 4096 *datapoints. Left panel:* $N = 32$*, middle panel:* $N = 16$*, right panel:* $N = 8$*.*