

MEASURING STATIONARITY IN LONG-MEMORY PROCESSES

Kemal Sen, Philip Preuß and Holger Dette

Ruhr-Universität Bochum

Abstract: In this paper we consider the problem of measuring stationarity in locally stationary long-memory processes. We introduce an L_2 -distance between the spectral density of the locally stationary process and its best approximation under the assumption of stationarity. The distance is estimated by a numerical approximation of the integrated spectral periodogram and asymptotic normality of the resulting estimate is established. The results can be used to construct a simple test for the hypothesis of stationarity in locally stationary long-range dependent processes. We also propose a bootstrap procedure to improve the approximation of the nominal level and prove its consistency. Throughout the paper, we work with Riemann sums of a squared periodogram instead of integrals (as it is usually done in the literature) and as a by-product of independent interest it is demonstrated that the two approaches behave differently in the limit.

Key words and phrases: Bootstrap, empirical spectral measure, goodness-of-fit tests, integrated periodogram, locally stationary process, long-memory, non-stationary processes, spectral density.

1. Introduction

The assumption of (second-order) stationarity is quite common in the analysis of time series data like wind speeds, computer network traffic or stock returns. This condition allows for a well developed statistical analysis, and there exist numerous books and articles dealing with parameter estimation or forecasting techniques. However, many real world phenomena cannot be adequately described by stationarity models and locally stationary processes have been proposed as an attractive alternative class of models, as they offer a more realistic theoretical framework for the analysis of time series with varying second-order characteristics. Early work on this subject can be found in Priestley (1965), who considered oscillating processes. Neumann and von Sachs (1997) and Nason, von Sachs, and Kroisandt (2000) discussed the estimation of evolutionary spectra by wavelet methods. Dahlhaus (1997) gave a definition of locally stationary processes on the basis of a time varying spectral representation and established the asymptotic theory for statistical inference in such cases (see also Dahlhaus

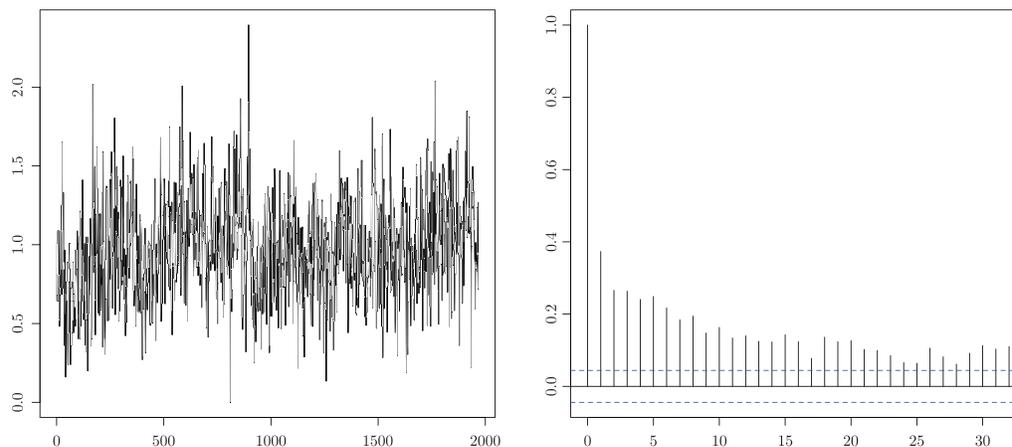


Figure 1. Left panel: Plot of the 1967 standardized annual tree ring width measurements at 2,926m in Bristlecone Pine, Nevada from 1 A.D. to 1969 A.D.; right panel: Sample autocovariance function.

(2000), Dahlhaus and Polonik (2006), and Dahlhaus (2009)). Alternative concepts to model time varying dependencies have recently been introduced by Zhou and Wu (2009, 2010) and cover a wide range of non-stationary processes. Most of this literature considers the case of short-range dependent data, while Palma and Olea (2010), Palma (2010), and Wu and Zhou (2014) developed statistical methodology for long-range dependent non-stationary processes.

A typical example of data from a long-range dependent process can be found in the left panel of Figure 1 which shows 1967 standardized annual tree ring width measurements at 2,926m in Bristlecone Pine, Nevada. In dendrochronology, time series long-memory occurs frequently and in a “pure” form, and therefore tree ring series are often used as prime examples of long-memory (see for example p.5 in Beran et al. (2013)). In the right part of the Figure 1 we display the sample autocovariance function of the data, which clearly indicates long-range dependence. In the present paper we are interested in the question whether data of this type can be analyzed under the assumption of stationarity. For this purpose we propose a measure of stationarity in long-range dependent locally stationary processes, which is used for the construction of a consistent test for the hypothesis of stationarity.

There exist several procedures to validate this condition in the context of short-memory processes. One of the first tests for stationarity dates back to Priestley and Subba Rao (1968) and is based on evolutionary spectral analysis. Von Sachs and Neumann (2000) proposed a test that is based on the estimation of wavelet coefficients by a localized version of the periodogram. Recently, Nason (2013) used the same principle and examined the constancy of a wavelet spectrum

by examining its Haar wavelet coefficients over a finite set of wavelet scales. Paparoditis (2009, 2010) suggested an L_2 -distance between the estimated spectral densities under the assumptions of stationarity and of local stationarity. Many authors have used the localized periodogram to construct a test for stationarity. Dwivedi and Subba Rao (2010) used the entire length of the time series and developed a Portmanteau-type test statistic to detect deviations from stationarity. Jentsch and Subba Rao (2015) utilized the fact that, under the assumption of stationarity, discrete Fourier transforms (DFT) are asymptotically uncorrelated for this purpose, while the test of Jin, Wang and Wang (2015) is based on the Walsh transform. All these methods require the choice of at least one smoothing parameter, such as the order of the wavelet expansion, a bandwidth for the estimation of the spectral density or the order in a Portmanteau-type test. Dette, Preuß and Vetter (2011a) developed tests for the hypothesis of stationarity in the framework of locally stationary processes which is based on an estimate of the L_2 -distance

$$D^2 := \int_0^1 \int_{-\pi}^{\pi} \left(f(u, \lambda) - \int_0^1 f(v, \lambda) dv \right)^2 d\lambda du \quad (1.1)$$

between the time varying spectral density $f(u, \lambda)$ and its approximation $\lambda \mapsto \int_0^1 f(v, \lambda) dv$ through a spectral density corresponding to a stationary process. It is easy to see that the process is stationary (i.e. the time varying spectral density does not depend on u) if and only if $D^2 = 0$, and D^2 can be considered as a measure of deviation from stationarity in the frequency domain. This quantity corresponds to the measure used in Paparoditis (2009), but unlike to this author, Dette, Preuß and Vetter (2011a) estimated D^2 directly via Riemann sums of the (squared) local periodogram instead of a smoothed local periodogram and thus avoided the choice of a smoothing parameter. Preuß and Vetter and Dette (2013) proposed an alternative measure for deviations from stationarity based on the Kolmogorov-Smirnov distance

$$D_{KS} := \sup_{(v, \omega) \in [0, 1]^2} \frac{1}{\pi} \left| \int_0^v \int_{-\pi\omega}^{\pi\omega} f(u, \lambda) d\lambda du - v \int_{-\pi\omega}^{\pi\omega} \int_0^1 f(u, \lambda) du d\lambda \right| \quad (1.2)$$

(see also Dahlhaus (2009)). Both approaches have pros and cons. In particular, tests based on the distance (1.2) are \sqrt{T} -consistent (here T denotes the sample size). On the other hand it is well known that - although such tests are consistent against alternatives converging to the null hypothesis at a parametric rate - Kolmogorov-type and related tests greatly weigh down contributions from high frequency components (see Ghosh and Huang (1991), Eubank and LaRiccia (1992) or Fan (1996)). Moreover, the limiting distribution of Kolmogorov-Smirnov-type test statistics is usually not known. In principle this problem can

be solved by bootstrap methods. However in many cases this yields to a loss of power. On the other hand, tests based on the L_2 -approach can often use critical values from the standard normal distribution.

As all procedures which have been suggested so far for discriminating between stationarity and non-stationarity, the tests proposed by Dette, Preuß and Vetter (2011a) and Preuß and Vetter and Dette (2013) are only applicable to short-memory processes, and the development of a corresponding methodology in the context of long-range dependence is missing. In fact, although stationary long-memory models are often employed in practice, there do not exist many tests for the hypothesis of stationarity which include these processes. Berkes et al. (2006), Sibbertsen and Kruse (2009) and Dehling, Rooch, and Taqqu (2013) consider CUSUM and Wilcoxon type tests to discriminate between long-range dependence and one change in mean. A change with respect to the mean is of course only the simplest possible deviation from stationarity and there is particular interest in measuring deviations in the dependency structure over time as well.

Recently, Preuß and Vetter (2013) developed a test for stationarity which includes the long-range dependent case and is based on the distance (1.2). There exist several situations where this approach is not the best and for this reason we consider an alternative test that is based on the measure defined in (1.1). For this purpose, we estimate the integrals in the distance D^2 in (1.1) by Riemann sums when the unknown spectral densities are replaced by periodograms. For the resulting statistic we show that an appropriately standardized statistic converges to a standard normal distribution if the (possibly time varying) long-memory parameter $d(u)$ is smaller than $1/8$. These results are used to develop a bootstrap procedure for the approximation of the limit distribution and to prove its consistency in the general case.

The proof of asymptotic normality is of a more theoretical nature, because the bootstrap procedure can in principle also be applied in the case $d(u) < 1/4$. However the results are of interest from several perspectives. Several arguments used in the proof of asymptotic normality are required in the proof of consistency of the bootstrap procedure and are illustrated more easily in the unconditional case. The estimate \hat{D}_T^2 of D^2 is based on estimates of the integrated and integrated squared spectral density $\int_0^1 f(u, \lambda) du$ and $\int_0^1 \int_{-\pi}^{\pi} f^2(u, \lambda) d\lambda du$. We use Riemann sums of the squared periodogram here instead of non-computable integrals as it is usually done in the literature (see Taniguchi (1980), Fox and Taqqu (1987) and Palma and Olea (2010), among others). Although one might expect that both estimators exhibit similar behavior with respect to weak convergence, it is demonstrated in Section 3 that this is not the case in the present context. A

similar observation was also made by Deo and Chen (2000) in the case of short-memory stationary processes. To the best of our knowledge, even in the (much simpler) stationary case, Riemann sums of a squared periodogram have not been considered in the literature for the long-range dependent case.

The paper is organized as follows. In Section 2, we introduce the necessary notation and define an empirical measure of stationarity. In Section 3, we prove that an appropriately standardized version of this measure converges weakly to a standard normal distribution if the time varying long-memory parameter is smaller than 1/8. In Section 4, we present a bootstrap procedure to approximate the distribution of \hat{D}_T^2 in the more general case and prove its consistency. The finite sample properties are investigated in Section 5. We defer all technical details to appendices in Section 6 and 7.

2. Measuring Stationarity in Locally Stationary Long-Memory Processes

In order to obtain a measure of stationarity, including the long-range dependent case, we require a set-up that is flexible enough to cover stationary long-memory processes and a reasonable time-varying extension of it as well. For this reason, we consider a theoretical framework of a locally stationary long-memory process (similar approaches can be found in Beran (2009), Palma and Olea (2010), and Roueff and von Sachs (2011)).

Let $\{X_{t,T}\}_{t=1,\dots,T}$ be a sequence of stochastic processes that have a MA(∞) representation of the form

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

where $\{Z_t\}_{t \in \mathbb{Z}}$ are independent and standard normal random variables. We assume the following.

Assumption 1.

- (1) *There exist twice continuously differentiable functions $\psi_l : [0, 1] \rightarrow \mathbb{R}$ ($l \in \mathbb{Z}$) such that*

$$\sup_{t=1,\dots,T} \left| \psi_{t,T,l} - \psi_l(t/T) \right| \leq \frac{C}{T I(l)^{1-d_\infty}}, \quad \forall l \in \mathbb{N}, \tag{2.2}$$

and

$$\psi_l(u) = \frac{a(u)}{I(l)^{1-d(u)}} + O\left(\frac{1}{I(l)^{2-d_\infty}}\right) \tag{2.3}$$

holds uniformly in u as $l \rightarrow \infty$, where $d : [0, 1] \rightarrow [0, 1/2)$ and $a : [0, 1] \rightarrow \mathbb{R}_+$ are twice differentiable functions, $C \in \mathbb{R}^+$ and $d_\infty = \sup_{u \in [0,1]} d(u) < 1/4$ are constants and $I(x) := |x| \cdot 1_{\{x \neq 0\}} + 1_{\{x=0\}}$.

(2) *The time varying spectral density*

$$f(u, \lambda) := \frac{1}{2\pi} \left| \sum_{l=0}^{\infty} \psi_l(u) \exp(-i\lambda l) \right|^2 \quad (2.4)$$

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

(3) *There exists a constant $C \in \mathbb{R}^+$, independent of u and λ , such that for $l \neq 0$*

$$\sup_{u \in (0,1)} |\psi'_l(u)| \leq \frac{C \log(l)}{l^{1-d_\infty}}, \quad (2.5)$$

$$\sup_{u \in (0,1)} |\psi''_l(u)| \leq \frac{C \log^2(l)}{l^{1-d_\infty}}, \quad (2.6)$$

$$\sup_{u \in (0,1)} \left| \frac{\partial}{\partial u} f(u, \lambda) \right| \leq \frac{C \log(\lambda)}{\lambda^{2d_\infty}}, \quad (2.7)$$

$$\sup_{u \in (0,1)} \left| \frac{\partial^2}{\partial u^2} f(u, \lambda) \right| \leq \frac{C \log^2(\lambda)}{\lambda^{2d_\infty}}. \quad (2.8)$$

For the sake of a transparent notation, we use $C \in \mathbb{R}^+$ as a universal constant throughout. Note that the process is stationary if $\psi_{l,t,T} = \psi_l$ for all $l, t, T \in \mathbb{N}$. Condition (2.2) and (2.3) ensure that the infinite sum in (2.1) exists in the L_2 sense, and (2.2) means that the process $X_{t,T}$ can be approximated by a stationary model on a small time interval. The assumption of Gaussianity is only imposed here to simplify technical arguments; see Remark 3 for more details. The main difference of this approach to that of Palma and Olea (2010) is that we do not assume that $\psi_{t,T,l} - \psi_l(t/T)$ is absolutely summable in l . Thus, a more general framework than usually considered, allowing for a rigorous proof of time-varying FARIMA models being included (see the discussion following (2.11) for more details).

We consider the process

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

in order to visualize some properties of a locally stationary long-memory process. Here $X_t(u)$ is stationary for every fixed $u \in [0, 1]$ and, analogously to the stationary case, (2.3) implies the existence of bounded functions $y_i : [0, 1] \rightarrow \mathbb{R}$ ($i = 1, 2$) such that

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

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

(see Palma and Olea (2010)). Consequently, the autocovariance function $\gamma(u, k) = \text{Cov}(X_0(u), X_k(u))$ is not absolutely summable and the time varying spectral density $f(u, \lambda)$ has a pole at $\lambda = 0$ for any $u \in [0, 1]$.

As an example we consider the time varying FARIMA(p, d, q) model defined by

$$a\left(\frac{t}{T}, B\right)X_{t,T} = b\left(\frac{t}{T}, B\right)(1 - B)^{-d(t/T)}Z_t, \quad t = 1, \dots, T, \tag{2.11}$$

where B denotes the backshift operator,

$$a(u, z) := 1 - \sum_{j=1}^p a_j(u)z^j, \quad b(u, z) := 1 + \sum_{j=1}^q b_j(u)z^j$$

for given functions $a_j, b_j : [0, 1] \rightarrow \mathbb{R}$, and $d : [0, 1] \rightarrow (0, d_\infty]$ is twice continuously differentiable with $d_\infty < 1/2$. It is shown in Preuß and Vetter (2013) that under certain regularity conditions on the functions a_j, b_j , these equations have a locally stationary solution in the sense of Assumption 1. If the functions a_j, b_j , and d do not depend on u , (2.11) corresponds to the common FARIMA(p, d, q) equation (see for example Palma (2007) for conditions for the existence of a solution), which is included in our theoretical framework.

For the construction of an estimate of the quantity (1.1) we note that

$$D^2 = 2\pi F_1 - 4\pi F_2, \tag{2.12}$$

where

$$F_1 := \frac{1}{2\pi} \int_0^1 \int_{-\pi}^\pi f^2(u, \lambda) d\lambda du, \tag{2.13}$$

$$F_2 := \frac{1}{4\pi} \int_{-\pi}^\pi \left(\int_0^1 f(u, \lambda) du \right)^2 d\lambda. \tag{2.14}$$

Consequently, it follows from (2.10) that the distance D^2 is only well defined if $d_\infty < 1/4$. We assume without loss of generality that the sample size T can be decomposed into M blocks with length N , $T = NM$ where N and M are positive integers and N is even. A rough estimator for the time-varying spectral density $f(u, \lambda)$ is then given by the local periodogram at the rescaled time point $u \in [0, 1]$, defined by

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

where $X_{j,T} = 0$ for $j \notin \{1, \dots, T\}$ (see Dahlhaus (1997)). This is the usual periodogram computed from the observations $X_{\lfloor uT \rfloor - N/2 + 1, T}, \dots, X_{\lfloor uT \rfloor + N/2, T}$, and it can be shown that it is asymptotically unbiased for the time-varying spectral density $f(u, \lambda)$ if $N \rightarrow \infty$ and $N = o(T)$. However, $I_N(u, \lambda)$ is not consistent just as the usual periodogram. In addition, $I_N(u, \lambda)^2$ is an unbiased (but not consistent) estimate of the quantity $2f^2(u, \lambda)$ instead of $f^2(u, \lambda)$.

We construct empirical versions of (2.13) and (2.14) by replacing the integrals through appropriate Riemann-sums and substitute $f(u, \lambda)$ and $f^2(u, \lambda)$ by $I_N(u, \lambda)$ and $I_N(u, \lambda)^2/2$, respectively. For this purpose, let

$$u_j := \frac{t_j}{T} := \frac{N(j-1) + N/2}{T}, \quad j = 1, \dots, M,$$

and consider the statistics

$$\hat{F}_{1,T} := \frac{1}{T} \sum_{j=1}^M \sum_{k=1}^{\lfloor N/2 \rfloor} I_N(u_j, \lambda_{k,N})^2, \quad (2.15)$$

$$\hat{F}_{2,T} := \frac{1}{N} \sum_{k=1}^{\lfloor \frac{N}{2} \rfloor} \left(\frac{1}{M} \sum_{j=1}^M I_N(u_j, \lambda_{k,N}) \right)^2, \quad (2.16)$$

where the $\lambda_{k,N} = 2\pi k/N$ denote the usual Fourier frequencies. The empirical measure of stationarity (1.1) is then

$$\hat{D}_T^2 := 2\pi \hat{F}_{1,T} - 4\pi \hat{F}_{2,T}. \quad (2.17)$$

It is far from obvious that \hat{D}_T^2 is a consistent estimator of D^2 . In general it is not true that an integrated non-linear function of the periodogram converges to the corresponding integrated function of the spectral density. This - at a first glance - is a counterintuitive property of the integrated periodogram, and was already observed by Taniguchi (1980) in the context of stationary short-memory processes. These problems are also visible here as we require a multiple of $I_N(u, \lambda)^2$ to obtain an asymptotically unbiased estimator for $f^2(u, \lambda)$. In the following section we prove consistency of \hat{D}_T^2 and study its weak convergence.

3. Consistency and Weak Convergence

Throughout this paper, the symbols $\xrightarrow{\mathcal{P}}$ and $\xrightarrow{\mathcal{D}}$ denote convergence in probability and weak convergence, respectively. In order to specify the bias of $\hat{F}_{1,T}$ and $\hat{F}_{2,T}$ we define

$$F_{1,T} := \frac{1}{2\pi M} \sum_{j=1}^M \int_{-\pi}^{\pi} f^2(u_j, \lambda) d\lambda, \quad F_{2,T} := \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\frac{1}{M} \sum_{j=1}^M f(u_j, \lambda) \right)^2 d\lambda,$$

and obtain the following results.

Theorem 1. *Suppose Assumption 1 holds with $\sup_{u \in [0,1]} d(u) < 1/4$ and that the conditions*

$$N \rightarrow \infty, \quad \frac{N}{T} \rightarrow 0 \quad (3.1)$$

are satisfied. Then $\hat{F}_{1,T} \xrightarrow{\mathcal{P}} F_1$, $\hat{F}_{2,T} \xrightarrow{\mathcal{P}} F_2$ and, in particular, $\hat{D}_T^2 \xrightarrow{\mathcal{P}} D^2$ as $T \rightarrow \infty$.

Remark 1. In the Gaussian case there are at least two heuristic explanations of why the factor 2 is required for the squared integrated periodogram. First, $I_N(u, \lambda_j) = |J_N(u, \lambda_j)|^2$, where $J_N(u, \lambda_j)$ denotes the localized discrete Fourier transform (DFT). An expansion in terms of cumulants then yields

$$\begin{aligned} E(I_n^2(u, \lambda_j)) &= 2E(I_N(u, \lambda_j))^2 \\ &\quad + \text{cum}(J_N(u, \lambda_j), J_N(u, \lambda_j)) \text{cum}(J_N(u, -\lambda_j), J_N(u, -\lambda_j)) \\ &\quad + \text{cum}(J_N(u, \lambda_j), J_N(u, -\lambda_j), J_N(u, \lambda_j), J_N(u, -\lambda_j)), \end{aligned}$$

where the third term vanishes (because of the Gaussian assumption) and the second term is of order $o(1)$ as $N \rightarrow \infty, M \rightarrow \infty$. This implies that $I_n^2(u, \lambda_j)/2$ is an unbiased estimate of the squared spectral density,

$$E[I_n^2(u, \lambda_j)] = 2f^2(u, \lambda_j) + o(1). \quad (3.2)$$

A second explanation for (3.2) can be obtained from the approximation

$$I_n(u_j, \lambda_k) \approx f(u_j, \lambda_k) Z_{jk}, \quad (3.3)$$

where Z_{jk} are independent identically exponentially distributed random variables, in particular $E[Z_{jk}^2] = 2$.

Theorem 2. *Suppose Assumption 1 holds with $d_\infty = \sup_{u \in [0,1]} d(u) < 1/8$ and that the conditions*

$$N \rightarrow \infty, \quad \frac{N}{T} \rightarrow 0, \quad \frac{\sqrt{T}}{N^{1-4d_\infty}} \rightarrow 0 \quad (3.4)$$

are satisfied. Then as $T \rightarrow \infty$ we have

$$\sqrt{T} \{ (\hat{F}_{1,T}, \hat{F}_{2,T})^T - (F_{1,T}, F_{2,T} + d_{N,T})^T - \mathbf{C}_T \} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \mathbf{\Sigma}),$$

where the covariance matrix $\mathbf{\Sigma}$ and the constant $d_{N,T}$ are

$$\mathbf{\Sigma} = \begin{pmatrix} \frac{5}{\pi} \int_{-\pi}^{\pi} \int_0^1 f^4(u, \lambda) du d\lambda & \frac{2}{\pi} \int_{-\pi}^{\pi} \left(\int_0^1 f(u, \lambda) du \int_0^1 f^3(u, \lambda) du \right) d\lambda \\ \frac{2}{\pi} \int_{-\pi}^{\pi} \left(\int_0^1 f(u, \lambda) du \int_0^1 f^3(u, \lambda) du \right) d\lambda & \frac{1}{\pi} \int_{-\pi}^{\pi} \left(\left(\int_0^1 f(u, \lambda) du \right)^2 \int_0^1 f^2(u, \lambda) du \right) d\lambda \end{pmatrix}, \quad (3.5)$$

$$d_{N,T} = \frac{1}{4\pi M^2} \sum_{j=1}^M \int_{-\pi}^{\pi} f^2(u_j, \lambda) d\lambda, \quad (3.6)$$

and the vector $\mathbf{C}_T = (C_{l,1}, C_{l,2})^T \in \mathbb{R}^2$ is of order $O(N^2/T^2 + \log(N)/(MN^{1-4d_\infty}))$, vanishing if the functions $\psi_l(u)$ are independent of u for all $l \in \mathbb{Z}$.

A similar result for the short-memory situation has been derived by Dette, Preuß and Vetter (2011a) under the condition $\sqrt{T}/N \rightarrow 0$. A comparison with (3.4) shows that a larger size of the window length N is necessary in the case of long-range dependence, which corresponds to intuition. In contrast to their result, there appears an additional bias term \mathbf{C}_T in Theorem 2. This term is negligible if $N^2/T^{3/2} \rightarrow 0$, which holds under the stronger restriction $d_\infty < 1/12$ due to (3.4). Under the null hypothesis of a time independent spectral density,

$$H_0 : f(u, \lambda) \text{ is independent of } u, \quad (3.7)$$

we have $\mathbf{C}_T = 0$ (this follows from the proof of Theorem 2 in the Appendix). Since the covariance matrix (3.5) contains the integrated fourth power of the spectral density, we obtain from (2.10) that Theorem 2 is not valid whenever $d_\infty \geq 1/8$.

Remark 2. The proof of Theorem 2 is based on a careful calculation of the higher order cumulants of the statistic $(\hat{F}_{1,T}, \hat{F}_{2,T})$ (see Theorem A.1 in the Appendix). As pointed out by a referee, an alternative proof could be obtained by generalizing recent results of Lahiri (2003) in several directions. Lahiri (2003) considered long-range dependent stationary processes, say $(\tilde{X}_t)_{t \in \mathbb{Z}}$, and investigated conditions, such that $\text{Cov}(J_T(\lambda_k), J_T(\lambda_{k'}))$ is asymptotically negligible, where $J_T(\lambda_k)$ denotes the DFT of $(\tilde{X}_t)_{t=1}^T$, at the k th Fourier frequency. We expect that similar results can be obtained for the localized DFT as it appears in the definition of the localized periodogram. In particular, bounds for $\text{Cov}(J_N(u, \lambda_{j_1}), J_N(v, \lambda_{j_2}))$ and $\text{Cov}(J_N(u, \lambda_{j_1}), J_N(v, \lambda_{j_1}))$ could be derived, such that the covariances are of small order if λ_{j_1} and λ_{j_2} have enough frequency separation and the distance between u and v is sufficiently large. An approximation for $\text{Cov}(J_N(u, \lambda_{j_1}), J_N(u, \lambda_{j_2}))$ has to be found if λ_{j_1} and λ_{j_2} are close. With results of this type it is possible to derive an approximation for $E[\hat{F}_{2,T}]$. For the calculation of $E[\hat{F}_{1,T}]$ and the higher order cumulants of $(\hat{F}_{1,T}, \hat{F}_{2,T})$, it is necessary to obtain results of this type for the cumulants

$$\text{cum}(J_N(u_{i_1}, \lambda_{j_1}), \dots, J_N(u_{i_\ell}, \lambda_{j_\ell})) \quad (3.8)$$

for any $\ell \in \mathbb{N}$. The derivation of such bounds in the stationary (generalizing the results of Lahiri (2003) in a non-trivial direction) and its extension to the

non-stationary long-memory case is an interesting topic for future research. If sharp approximations for these cumulants can be derived, it is possible to write the higher order cumulants of $(\hat{F}_{1,T}, \hat{F}_{2,T})$ in terms of the cumulants of $J_N(u_i, \lambda_j)$ and to prove their convergence (if the statistics are appropriately standardized).

Corollary 1. *Under the assumptions of Theorem 2,*

$$\sqrt{T} \left(\hat{D}_T^2 - D_T^2 + 4\pi d_{N,T} + 4\pi C_{2,T} - 2\pi C_{1,T} \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \tau^2), \tag{3.9}$$

where $D_T^2 := 2\pi F_{1,T} - 4\pi F_{2,T}$ and the asymptotic variance is given by

$$\begin{aligned} \tau^2 := & 20\pi \int_{-\pi}^{\pi} \int_0^1 f^4(u, \lambda) dud\lambda - 32\pi \int_{-\pi}^{\pi} \left(\int_0^1 f(u, \lambda) du \int_0^1 f^3(u, \lambda) du \right) d\lambda \\ & + 16\pi \int_{-\pi}^{\pi} \left(\left(\int_0^1 f(u, \lambda) du \right)^2 \int_0^1 f^2(u, \lambda) du \right) d\lambda. \end{aligned} \tag{3.10}$$

Under the null hypothesis (3.7) we have $D_T^2 = C_{1,T} = C_{2,T} = 0$ and the asymptotic variance in (3.10) reduces to $\tau_{H_0}^2 := 4\pi \int_{-\pi}^{\pi} f^4(\lambda) d\lambda$. The asymptotic bias $4\pi d_{N,T} = 2\pi N/T F_{1,T}$ can easily be estimated by the statistic $\hat{B}_T := 2\pi N/T \hat{F}_{1,T}$, and we infer from Theorem 2 that

$$\sqrt{T} \left(\hat{B}_T - 4\pi d_{N,T} \right) = \frac{2\pi N}{T} \sqrt{T} \left(\hat{F}_{1,T} - F_{1,T} \right) \xrightarrow{\mathcal{P}} 0. \tag{3.11}$$

Thus Slutsky’s Lemma, together with (3.9), yields

$$\sqrt{T} \left(\hat{D}_T^2 + \hat{B}_T \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \tau_{H_0}^2) \tag{3.12}$$

under the null hypothesis. To construct an asymptotic level α -test for stationarity, it remains to estimate the variance $\tau_{H_0}^2$ in (3.12), and an estimator for this quantity is given by $\hat{\tau}_{H_0}^2 := 4\pi^2 \hat{\tau}_1^2$ with

$$\hat{\tau}_1^2 := \frac{1}{6T} \sum_{k=1}^{\lfloor N/2 \rfloor} \sum_{j=1}^M I_N(u_j, \lambda_{k,N})^4.$$

Theorem 3. *If the assumptions of Theorem 2 are satisfied, we have*

$$\hat{\tau}_1^2 \xrightarrow{\mathcal{P}} \frac{1}{\pi} \int_{-\pi}^{\pi} \int_0^1 f^4(u, \lambda) dud\lambda.$$

Combining (3.12) with Theorem 3 yields

$$\frac{\sqrt{T} \left(\hat{D}_T^2 + \hat{B}_T \right)}{\sqrt{\hat{\tau}_{H_0}^2}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1), \tag{3.13}$$

and therefore an asymptotic level α -test is obtained by rejecting the null hypothesis (3.7) whenever

$$\frac{\sqrt{T}(\hat{D}_T^2 + \hat{B}_T)}{\sqrt{\hat{\tau}_{H_0}^2}} \geq u_{1-\alpha}, \tag{3.14}$$

where $u_{1-\alpha}$ denotes the $(1 - \alpha)$ -quantile of the standard normal. It follows from Theorem 2 that this test is consistent, because the left hand side of (3.14) converges to infinity whenever there exists a $\tilde{\lambda} \in [-\pi, \pi]$ such that the function $u \mapsto f(u, \tilde{\lambda})$ is not constant.

Remark 3. If the innovation process $(Z_t)_{t \in \mathbb{Z}}$ in (2.1) is not Gaussian, it can be shown that Corollary 1 is still valid, where the asymptotic variance τ^2 in (3.10) has to be replaced by

$$\begin{aligned} \tau_g^2 = & \tau^2 + \frac{\kappa_4}{\kappa_2^2} \left\{ 4 \int_0^1 \left(\int_{-\pi}^{\pi} f^2(u, \lambda) d\lambda \right)^2 du \right. \\ & + 4 \int_0^1 \left(\int_{-\pi}^{\pi} f(u, \lambda) \left(\int_0^1 f(\nu, \lambda) d\nu \right) d\lambda \right)^2 du \\ & \left. - 8 \int_0^1 \left(\int_{-\pi}^{\pi} f^2(u, \lambda) d\lambda \int_{-\pi}^{\pi} f(u, \lambda) \left(\int_0^1 f(\nu, \lambda) d\nu \right) d\lambda \right) du \right\}, \end{aligned}$$

and κ_2 and κ_4 denote the second and fourth cumulants of the innovation process. In particular, under the null hypothesis of stationarity, it follows that $\tau_g^2 = \tau^2 = \tau_{H_0}^2$, and hence no adjustments in the asymptotic level α -test in (3.14) are necessary to address non normal distributed innovations.

Remark 4. The quantity D^2 in (1.1) can be considered as measure of stationarity, because it quantifies the distance between the spectral density $f(u, \lambda)$ and its best approximation by a spectral density of a stationary process. Theorem 1 provides a consistent estimate of this quantity. Moreover, Corollary 1 can be used to construct an asymptotic confidence interval for D^2 . Under the condition $d_\infty < 1/12$, the discussion in the paragraph following Theorem 2 shows that $\sqrt{TC}_{1,T}$ and $\sqrt{TC}_{2,T}$ are asymptotically negligible. Due to Assumption 1(2) it follows that $D^2 = D_T^2 + O(1/M^2)$, and by similar arguments as given in the proof of Theorem 3 it can be shown that $\hat{\tau}_{H_1}^2 = 20\pi^2\hat{\tau}_1^2 - 32\pi^2\hat{\tau}_2 + 16\pi^2\hat{\tau}_3^2$ is a consistent estimate of the asymptotic variance τ^2 , where

$$\hat{\tau}_2^2 = \frac{2}{3NM^2} \sum_{k=1}^{\lfloor N/2 \rfloor} \sum_{j_1, j_2=1}^M I_N(u_{j_1}, \lambda_{k,N}) I_N(u_{j_2}, \lambda_{k,N})^3,$$

$$\hat{\tau}_3^2 = \frac{2}{NM^3} \sum_{k=1}^{\lfloor \frac{N}{2} \rfloor} \sum_{j_1, j_2, j_3=1}^M I_N(u_{j_1}, \lambda_{k,N}) I_N(u_{j_2}, \lambda_{k,N}) I_N(u_{j_3}, \lambda_{k,N})^2.$$

Therefore from (3.9) and (3.11), an asymptotic $(1 - \alpha)$ confidence interval for the measure of stationarity D^2 is

$$\left[\hat{D}_T - \hat{B}_T - \frac{\hat{\tau}_{H_1}}{\sqrt{T}} u_{1-\alpha/2}, \hat{D}_T - \hat{B}_T + \frac{\hat{\tau}_{H_1}}{\sqrt{T}} u_{1-\alpha/2} \right].$$

Remark 5. For locally stationary long-range dependent models the asymptotic variances of the statistics

$$\tilde{F}_{1,T} = \frac{1}{4\pi M} \sum_{j=1}^M \int_{-\pi}^{\pi} I_N(u_j, \lambda)^2 d\lambda$$

and of $\hat{F}_{1,T}$, defined in (2.15), are different. In fact we prove in Section A.2 that

$$\begin{aligned} \lim_{T \rightarrow \infty} T \text{Var}(\tilde{F}_{1,T}) &= \frac{14}{3\pi} \int_{-\pi}^{\pi} \int_0^1 f^4(u, \lambda) du d\lambda, \\ \lim_{T \rightarrow \infty} T \text{Var}(\hat{F}_{1,T}) &= \frac{5}{\pi} \int_{-\pi}^{\pi} \int_0^1 f^4(u, \lambda) du d\lambda \end{aligned}$$

follows by Theorem 2. Moreover, similar arguments show that, even in the stationary case, the asymptotic variance of the statistic $\int_{-\pi}^{\pi} I_T(\lambda)^2 d\lambda$ and its discretized version $(2\pi/T) \sum_{k=1}^T I_T(\lambda_{k,T})^2$ are not the same (here $I_T(\lambda)$ denotes the usual periodogram and $\lambda_{k,T} = 2\pi k/T$ are the Fourier frequencies). Deo and Chen (2000) observed the same effect in the context of stationary short-memory processes.

4. Critical Values by Resampling

We now consider the more general set-up with $d_\infty < 1/4$ as specified in Assumption 1. We show that in this case a bootstrap procedure can be used to approximate the distribution of \hat{D}_T^2 under the null hypothesis (3.7). We employ the FARI(∞) bootstrap, recently introduced by Preuß and Vetter (2013), that fits an FARIMA($p, \underline{d}, 0$)-model to the data, where $p = p(T)$ converges to infinity with increasing sample size T . The test statistic investigated by these authors is different from that in this paper and we therefore require different growth conditions on p , N , and T to obtain a valid testing bootstrap procedure. These are summarized in the following.

Assumption 2. For the stationary process $\{X_t\}_{t \in \mathbb{Z}}$ with strictly positive spectral density $\lambda \mapsto \int_0^1 f(u, \lambda) du$, there exists a constant $\underline{d} \in (0, 1/4)$ such that the process

$$Y_t = (1 - B)^{\underline{d}} X_t \quad (4.1)$$

has an $AR(\infty)$ -representation of the form

$$Y_t = \sum_{j=1}^{\infty} a_j Y_{t-j} + Z_t^{AR}, \quad (4.2)$$

where $\{Z_j^{AR}\}_{j \in \mathbb{Z}}$ denotes a Gaussian White Noise process with variance $\sigma^2 > 0$, and the coefficients in the representation (4.2) satisfy

$$\sum_{j=1}^{\infty} |a_j| |j|^7 < \infty, \quad (4.3)$$

$$1 - \sum_{j=1}^{\infty} a_j z^j \neq 0 \text{ for } |z| \leq 1. \quad (4.4)$$

Under the null hypothesis of a time independent spectral density, it follows that $\underline{d} = d_\infty = d(u)$ for all $u \in [0, 1]$, but under the alternative we usually have $\underline{d} \neq d_\infty$. The FARI(∞) bootstrap incorporates the following steps. Choose a $p = p(T) \in \mathbb{N}$ to construct an estimator, say $\hat{\underline{d}}$, of the long-range dependence parameter \underline{d} in model (4.1); calculate an estimator of

$$(a_{1,p}, \dots, a_{p,p}) = \underset{b_{1,p}, \dots, b_{p,p}}{\operatorname{argmin}} E \left(Y_t - \sum_{j=1}^p b_{j,p} Y_{t-j} \right)^2, \quad (4.5)$$

by fitting an $AR(p)$ -model to the data. To describe the main idea of our procedure in detail, take the “true” approximating process to be

$$Y_t^{AR}(p) = \sum_{j=1}^p a_{j,p} Y_{t-j}^{AR}(p) + Z_t^{AR}, \quad (4.6)$$

where the parameters $a_{j,p}$ are defined in (4.5), and $\{Z_t^{AR}\}_{t \in \mathbb{Z}}$ is a Gaussian White Noise process with mean zero and variance $\sigma_p^2 = E(Y_t - \sum_{j=1}^p a_{j,p} Y_{t-j})^2$. If $p = p(T) \rightarrow \infty$ the process $Y_t^{AR}(p)$ approximates 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 by $\lambda \mapsto \int_0^1 f(u, \lambda) du$. Under the null hypothesis of stationarity, this function coincides with the spectral density of $\{X_{t,T}\}_{t=1, \dots, T}$. Hence, observing the data $X_{1,T}, \dots, X_{T,T}$, the FARI(∞) bootstrap precisely works as follows.

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

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

where $\theta_p = (d, \sigma_p^2, a_{1,p}, \dots, a_{p,p})$,

$$I_T(\lambda) = \frac{1}{2\pi T} \left| \sum_{t=1}^T X_{t,T} \exp(-i\lambda t) \right|^2$$

is the usual periodogram, and

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

is the spectral density of a stationary FARIMA($p, d, 0$)-model. (The estimator $\hat{\theta}_{T,p}$ is the classical Whittle estimator of a stationary process (see Whittle (1951))).

- (2) Simulate the pseudo-series $(X_t^*)_{t \in \mathbb{Z}}$ according to a FARIMA($p, \hat{d}, 0$) model with AR-parameters $\hat{a}_{1,p}, \dots, \hat{a}_{p,p}$ and $\hat{\sigma}_p^2$ as the variance of the Gaussian innovations.
- (3) Compute $\hat{D}_T^{2,*}$ in the same way as \hat{D}_T^2 , where the original observations $X_{1,T}, \dots, X_{T,T}$ are replaced by the bootstrap replicates $X_{1,T}^*, \dots, X_{T,T}^*$.

Theorem 4. *Assume the null hypothesis (3.7) and Assumption 1 and 2 hold. Suppose the conditions*

$$N \rightarrow \infty, \quad \frac{N}{T} \rightarrow 0, \quad \frac{T}{N^{1+\delta}} \rightarrow 0 \tag{4.7}$$

are satisfied for some $0 < \delta < 1/2$, and assume $p = p(T)$ satisfies the following.

- (i) *There exist sequences $p_{max}(T) \geq p_{min}(T) \xrightarrow{T \rightarrow \infty} \infty$ such that $p(T) \in [p_{min}(T), p_{max}(T)]$,*

$$p_{max}^9(T) \log(T)^3 N^\delta T^{-1} = O(1), \tag{4.8}$$

$$\frac{\sqrt{T} p_{min}^{-9}(T)}{\sqrt{\log(T)}} = o(1). \tag{4.9}$$

- (ii) *The condition*

$$\|\hat{\theta}_{T,p} - \theta_p\|_\infty = O_P \left(\sqrt{\frac{\log(T)}{T}} \right) \tag{4.10}$$

is fulfilled uniformly with respect to p , where $\hat{\theta}_{T,p}$ denotes the estimator used in step (1) of the bootstrap procedure and $\theta_p = (d_p, \sigma_p^2, a_{1,p}, \dots, a_{p,p})$ are the corresponding “true” parameters.

Then there exist random variables $\hat{D}_{T,a}^2$ and $\hat{D}_{T,a}^{2,*}$ such that

- (a) $\hat{D}_{T,a}^2 \stackrel{\mathcal{D}}{=} \hat{D}_{T,a}^{2,*}$,
- (b) $\text{Var}(\hat{D}_T^2)^{-1/2}(\hat{D}_T^2 - \hat{D}_{T,a}^2) = o_P(1)$,
- (c) $\text{Var}(\hat{D}_T^{2,*})^{-1/2}(\hat{D}_T^{2,*} - \hat{D}_{T,a}^{2,*}) = o_P(1)$,
- (d) $E|\hat{D}_{T,a}^{2,*}| = O\left(\frac{N^{\max(4d-1/2,0)}}{\sqrt{T}} + \frac{\sqrt{\log(N)}1_{\{d=1/8\}}}{\sqrt{T}} + \frac{1}{N^{1-4d}}\right)$.

The estimate in (d) also holds if the null hypothesis (3.7) is not satisfied.

Conditions like (4.8)–(4.10) are standard in the context of parametric bootstraps [see for example Berg, Paparoditis, and Politis (2010) or Kreiß, Paparoditis, and Politis (2011)] and a detailed discussion of them is given in Preuß and Vetter (2013). There it is also discussed why (4.10) might hold in the general framework considered here, but a rigorous proof of such a statement is an open problem so far.

We obtain an asymptotic level α -test based on \hat{D}_T^2 as follows. Calculate B bootstrap replicates $\hat{D}_T^{2,*}$, denote by $(\hat{D}_T^{2,*})_{T,1}, \dots, (\hat{D}_T^{2,*})_{T,B}$ the resulting order statistic and reject the null hypothesis whenever

$$\hat{D}_T^2 > (\hat{D}_T^{2,*})_{T, \lfloor (1-\alpha)B \rfloor}. \quad (4.11)$$

Theorem 4 and the arguments in Paparoditis (2010) indicate that this procedure yields an asymptotic level α -test. To prove this, we follow Bickel and Freedman (1981) by considering the Mallows metric $d_2(F, G) = \inf \sqrt{E(X - Y)^2}$ between two distributions F and G , where the infimum is taken over all pairs (X, Y) of random variables with marginal distributions F and G . Theorem 4 then yields the following.

Theorem 5. *Suppose the null hypothesis (3.7) and the assumptions of Theorem 4 are satisfied. Then, as $T \rightarrow \infty$, the Mallows distance d_2 between the distributions of the random variables*

$$\frac{\hat{D}_T^2}{\sqrt{\text{Var}(\hat{D}_T^2)}} \quad \text{and} \quad \frac{\hat{D}_T^{2,*}}{\sqrt{\text{Var}(\hat{D}_T^{2,*})}}$$

converges to zero in probability.

Consistency under the alternative follows since Theorem 4(d) yields that each bootstrap statistic $\hat{D}_T^{2,*}$ converges to zero while \hat{D}_T^2 exceeds some positive constant (for T sufficiently large) due to Theorem 1.

Remark 6. The condition $d_\infty < 1/4$ is required to guarantee the existence of the integral $\int_0^1 \int_{-\pi}^\pi f^2(u, \lambda) d\lambda du$. A way to avoid restrictions of this type is the consideration of the modified L_2 -distance

$$D_2^2 := \int_0^1 \int_{-\pi}^\pi \lambda^2 \left(f(u, \lambda) - \int_0^1 f(v, \lambda) dv \right)^2 d\lambda du$$

to incorporate the case $1/4 \leq d_\infty < 1/2$. Here the term λ^2 compensates for the possible pole of the squared time varying spectral density at the origin. With this definition the restriction $d_\infty < 1/8$ in the decision rule (3.14) (and $d_\infty < 1/4$ in the bootstrap version (4.11)) can be omitted and the corresponding test statistic becomes $\hat{D}_{2,T}^2 := 2\pi\hat{F}_{1,2,T} - 4\pi\hat{F}_{2,2,T}$, where

$$\begin{aligned} \hat{F}_{1,2,T} &:= \frac{1}{T} \sum_{j=1}^M \sum_{k=1}^{\lfloor N/2 \rfloor} \lambda_{k,N}^2 I_N(u_j, \lambda_{k,N})^2, \\ \hat{F}_{2,2,T} &:= \frac{1}{N} \sum_{k=1}^{\lfloor N/2 \rfloor} \lambda_{k,N}^2 \left(\frac{1}{M} \sum_{j=1}^M I_N(u_j, \lambda_{k,N}) \right)^2. \end{aligned}$$

By using similar arguments as in the proof of Theorems 2 and 3 one obtains that, under H_0 , the term $\sqrt{T}(\hat{D}_{2,T}^2 + \hat{B}_{2,T})/\hat{\tau}_2$ is asymptotically normal with variance one, where

$$\begin{aligned} \hat{B}_{2,T} &:= \frac{2\pi N}{T^2} \sum_{j=1}^M \sum_{k=1}^{\lfloor N/2 \rfloor} \lambda_{k,N}^2 I_N(u_j, \lambda_{k,N})^2 \\ \hat{\tau}_2 &:= \frac{2\pi^2}{3} \frac{1}{T} \sum_{k=1}^{\lfloor N/2 \rfloor} \sum_{j=1}^M \lambda_{k,N}^4 I_N(u_j, \lambda_{k,N})^4 \end{aligned}$$

are the estimators for the corresponding bias and variance term (cf. B_T and $\hat{\tau}_1$ in (3.14)). This yields a test for stationarity in the long-range dependent framework that is based on quantiles of the standard normal and does not impose any restrictions on the range of the time-varying long memory parameter. However, since the small frequencies are weighted down by the additional factor λ^2 , this approach lacks power in detecting changes in the long-range dependence parameter itself.

5. Finite Sample Properties

In this section we examine the finite sample properties of the proposed decision rule (4.11). An important problem is the choice of the window length N for the calculation of the local periodogram and the choice of the AR parameter p in

the bootstrap procedure. Throughout this section we choose p as the minimizer of the AIC criterion (see Akaike (1973)),

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

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]. We therefore restrict ourselves to an analysis of the sensitivity with respect to N in the following, and it will turn out that the test (4.11) using the FARI(∞) bootstrap exhibits a remarkable robustness with respect to the choice of N . All reported results of this section are based on 1,000 simulation runs and 200 bootstrap replications.

5.1. Size and power of the test

In order to investigate the approximation of the nominal level we simulated data from the FARIMA(1, d , 0) model

$$(1 - \phi B)(1 - B)^d X_t = Z_t, \quad (5.1)$$

and the FARIMA(0, d , 1) process

$$(1 - B)^d X_t = (1 + \theta B) Z_t \quad (5.2)$$

for different values of ϕ, θ , and d where Z_t are independent standard normal. The rejection probabilities for the bootstrap test (4.11) are displayed in Tables 1–4 where $d \in \{0.1, 0.2\}$. We observe a very precise approximation of the nominal level in nearly all cases that is rather robust with respect to different choices of the parameter M and N .

To study the power of the test we considered the alternatives

$$X_{t,T} = Z_t^{(d)} + 0.8 \cos \left(1.5 - \cos \left(\frac{4\pi t}{T} \right) \right) Z_{t-1}^{(d)}, \quad (5.3)$$

$$\left(1 - 0.6 \sin \left(\frac{4\pi t}{T} \right) B \right) X_{t,T} = Z_t^{(d)}, \quad (5.4)$$

$$X_{t,T} = \sqrt{\sin \left(\frac{\pi t}{T} \right)} Z_t^{(d)}, \quad (5.5)$$

where $Z_t^{(d)} = (1 - B)^{-d} Z_t$ and $d = 0.2$. These kinds of alternatives were investigated by several authors in the context of locally stationary short-memory processes (see Paparoditis (2010) and Dahlhaus (1997)). The rejection frequencies for the bootstrap test (4.11) are depicted in Figures 2–4

Table 1. Rejection probabilities of the bootstrap test (4.11) under H_0 for different choices of T, N and M . The data was generated according to model (5.1) with $d = 0.1$ and different values for ϕ .

	T	N	M	$\phi = -0.9$		$\phi = -0.5$		$\phi = 0$		$\phi = 0.5$		$\phi = 0.9$	
				5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
A1	128	16	8	0.126	0.182	0.072	0.123	0.036	0.074	0.073	0.126	0.084	0.147
A2	128	8	16	0.140	0.200	0.085	0.132	0.041	0.090	0.073	0.128	0.084	0.118
B1	256	32	8	0.065	0.135	0.064	0.119	0.042	0.088	0.062	0.113	0.075	0.147
B2	256	16	16	0.080	0.132	0.056	0.108	0.040	0.082	0.051	0.109	0.062	0.109
B3	256	8	32	0.068	0.111	0.045	0.095	0.046	0.097	0.072	0.147	0.049	0.114
C1	512	64	8	0.054	0.109	0.049	0.106	0.039	0.089	0.049	0.114	0.082	0.134
C2	512	32	16	0.038	0.093	0.043	0.086	0.039	0.085	0.059	0.108	0.065	0.132
C3	512	16	32	0.061	0.095	0.051	0.102	0.045	0.081	0.059	0.109	0.043	0.104
C4	512	8	64	0.060	0.107	0.053	0.098	0.045	0.083	0.060	0.116	0.042	0.093
D1	1024	128	8	0.039	0.104	0.042	0.093	0.042	0.085	0.035	0.093	0.079	0.132
D2	1024	64	16	0.053	0.104	0.058	0.097	0.050	0.110	0.057	0.101	0.068	0.126
D3	1024	32	32	0.033	0.076	0.058	0.114	0.046	0.086	0.070	0.107	0.062	0.115
D4	1024	16	64	0.046	0.089	0.036	0.091	0.044	0.084	0.054	0.109	0.044	0.099
D5	1024	8	128	0.037	0.073	0.041	0.091	0.041	0.091	0.061	0.131	0.045	0.097

Table 2. Rejection probabilities of the bootstrap test (4.11) under H_0 for different choices of T, N and M . The data was generated according to model (5.1) with $d = 0.2$ and different values for ϕ .

	T	N	M	$\phi = -0.9$		$\phi = -0.5$		$\phi = 0$		$\phi = 0.5$		$\phi = 0.9$	
				5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
A1	128	16	8	0.107	0.164	0.063	0.114	0.050	0.108	0.072	0.121	0.108	0.166
A2	128	8	16	0.106	0.160	0.064	0.118	0.041	0.085	0.073	0.124	0.078	0.138
B1	256	32	8	0.064	0.123	0.048	0.104	0.042	0.094	0.075	0.131	0.079	0.137
B2	256	16	16	0.058	0.125	0.051	0.101	0.040	0.101	0.065	0.112	0.055	0.116
B3	256	8	32	0.079	0.124	0.047	0.089	0.051	0.091	0.053	0.106	0.050	0.105
C1	512	64	8	0.050	0.093	0.048	0.090	0.051	0.103	0.047	0.104	0.075	0.133
C2	512	32	16	0.047	0.104	0.044	0.087	0.039	0.085	0.053	0.109	0.068	0.124
C3	512	16	32	0.042	0.097	0.044	0.087	0.057	0.106	0.046	0.105	0.060	0.104
C4	512	8	64	0.050	0.102	0.053	0.101	0.052	0.088	0.058	0.121	0.062	0.114
D1	1024	128	8	0.044	0.090	0.046	0.102	0.051	0.107	0.039	0.092	0.076	0.140
D2	1024	64	16	0.043	0.082	0.040	0.088	0.050	0.098	0.046	0.098	0.060	0.106
D3	1024	32	32	0.045	0.089	0.054	0.097	0.057	0.103	0.060	0.104	0.066	0.115
D4	1024	16	64	0.044	0.087	0.038	0.087	0.049	0.094	0.059	0.106	0.051	0.101
D5	1024	8	128	0.041	0.082	0.041	0.089	0.038	0.086	0.061	0.103	0.054	0.103

for different combinations of T and N . Additionally, the results for the Kolmogorov-Smirnov approach of Preuß and Vetter (2013) are presented. We observe that the new procedure clearly outperforms the test of Preuß and Vetter (2013) for the models (5.3) and (5.4) while the Kolmogorov-Smirnov test works

Table 3. Rejection probabilities of the bootstrap test (4.11) under H_0 for different choices of T, N and M . The data was generated according to model (5.2) with $d = 0.1$ and different values for θ .

	T	N	M	$\phi = -0.9$		$\phi = -0.5$		$\phi = 0$		$\phi = 0.5$		$\phi = 0.9$	
				5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
A1	128	16	8	0.072	0.116	0.054	0.107	0.041	0.085	0.044	0.088	0.077	0.123
A2	128	8	16	0.068	0.122	0.054	0.112	0.059	0.125	0.073	0.117	0.070	0.133
B1	256	32	8	0.045	0.100	0.060	0.101	0.041	0.081	0.042	0.082	0.036	0.084
B2	256	16	16	0.053	0.096	0.058	0.104	0.045	0.094	0.045	0.102	0.060	0.104
B3	256	8	32	0.064	0.123	0.057	0.113	0.049	0.101	0.042	0.092	0.061	0.130
C1	512	64	8	0.043	0.089	0.043	0.095	0.044	0.086	0.045	0.088	0.041	0.095
C2	512	32	16	0.046	0.109	0.067	0.112	0.052	0.093	0.051	0.096	0.043	0.086
C3	512	16	32	0.048	0.099	0.055	0.095	0.062	0.114	0.050	0.102	0.051	0.098
C4	512	8	64	0.038	0.097	0.055	0.100	0.047	0.100	0.046	0.093	0.042	0.092
D1	1024	128	8	0.053	0.103	0.060	0.099	0.051	0.099	0.071	0.118	0.044	0.094
D2	1024	64	16	0.044	0.100	0.062	0.124	0.048	0.090	0.068	0.119	0.042	0.093
D3	1024	32	32	0.053	0.107	0.064	0.116	0.044	0.082	0.045	0.094	0.043	0.098
D4	1024	16	64	0.044	0.096	0.038	0.084	0.042	0.093	0.045	0.087	0.042	0.087
D5	1024	8	128	0.049	0.109	0.042	0.085	0.054	0.109	0.048	0.083	0.042	0.096

Table 4. Rejection probabilities of the bootstrap test (4.11) under H_0 for different choices of T, N and M . The data was generated according to model (5.2) with $d = 0.2$ and different values for θ .

	T	N	M	$\theta = -0.9$		$\theta = -0.5$		$\theta = 0$		$\theta = 0.5$		$\theta = 0.9$	
				5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
A1	128	16	8	0.068	0.112	0.060	0.103	0.030	0.081	0.060	0.108	0.053	0.111
A2	128	8	16	0.060	0.117	0.051	0.103	0.061	0.110	0.062	0.114	0.068	0.117
B1	256	32	8	0.059	0.122	0.048	0.102	0.045	0.094	0.038	0.078	0.040	0.083
B2	256	16	16	0.053	0.109	0.041	0.095	0.047	0.093	0.040	0.080	0.048	0.091
B3	256	8	32	0.060	0.100	0.048	0.098	0.057	0.119	0.050	0.092	0.061	0.106
C1	512	64	8	0.059	0.110	0.064	0.122	0.052	0.099	0.056	0.099	0.055	0.101
C2	512	32	16	0.060	0.122	0.044	0.107	0.041	0.103	0.043	0.113	0.046	0.086
C3	512	16	32	0.061	0.116	0.056	0.122	0.049	0.089	0.046	0.088	0.052	0.099
C4	512	8	64	0.056	0.095	0.057	0.118	0.057	0.110	0.047	0.100	0.055	0.102
D1	1024	128	8	0.063	0.125	0.054	0.102	0.039	0.086	0.044	0.101	0.051	0.098
D2	1024	64	16	0.051	0.109	0.061	0.112	0.047	0.107	0.056	0.106	0.047	0.100
D3	1024	32	32	0.055	0.092	0.057	0.111	0.048	0.095	0.057	0.106	0.047	0.119
D4	1024	16	64	0.065	0.124	0.061	0.116	0.043	0.092	0.048	0.087	0.049	0.098
D5	1024	8	128	0.059	0.116	0.052	0.095	0.049	0.093	0.035	0.075	0.050	0.115

better for the process (5.5). In addition, we observe that the new decision rule is less sensitive with respect to different choices of N than the test based on the Kolmogorov-Smirnov distance.

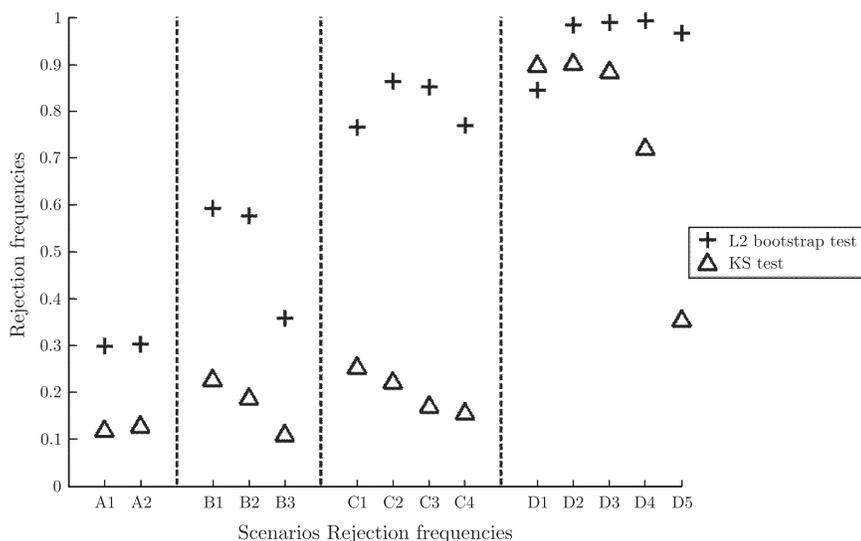


Figure 2. Power of the test (4.11) and the Kolmogorov-Smirnov test of Preuß and Vetter (2013) at 5% level for the model (5.3) under the scenarios A1–D5 from Table 1.

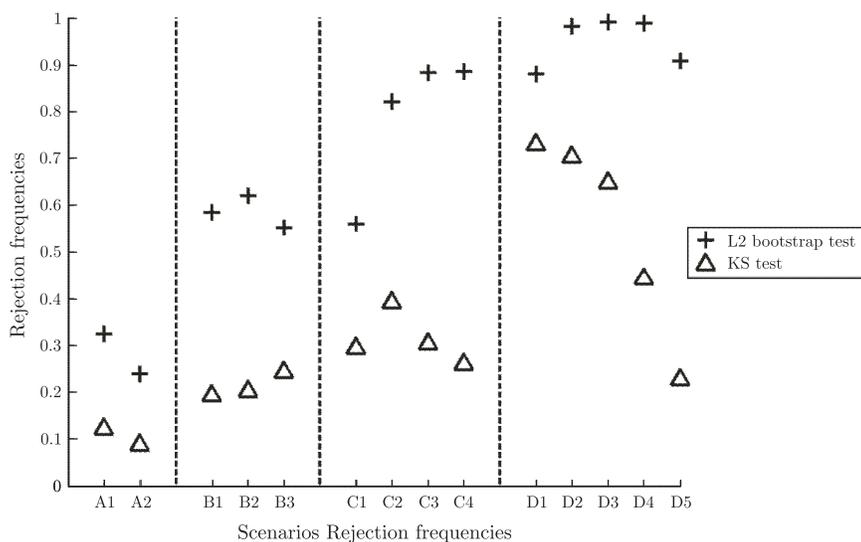


Figure 3. Power of the test (4.11) and the Kolmogorov-Smirnov test of Preuß and Vetter (2013) at 5% level for the model (5.4) under the scenarios A1–D5 from Table 1.

5.2. Data example

As an illustration we applied the L_2 -test (3.14) and the bootstrap test (4.11)

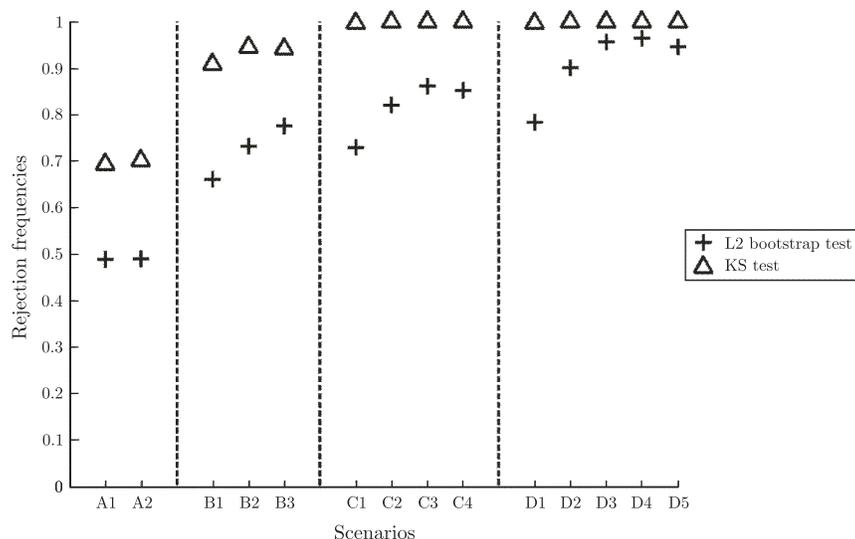


Figure 4. Power of the test (4.11) and the Kolmogorov-Smirnov test of Preuß and Vetter (2013) at 5% level for the model (5.5) under the scenarios A1–D5 from Table 1.

to the annual tree ring width measurements described in the introduction. The p-values for the bootstrap test were calculated by replicating the bootstrap procedure 5,000 times. The p-values for the L_2 -test for $M = 4$ and $M = 8$ blocks were 0.27 and 0.43, respectively, and the p-values for the bootstrap test for $M = 4$ and $M = 8$ were 0.18 and 0.33. Hence the null hypothesis of stationarity can not be rejected.

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).

Appendix A: Technical Details

We state two results are central for the proof of the statements in Sections 3 and 4.

Theorem A.1. If Assumption 1 is satisfied with $d_\infty < 1/4$, the following hold.

(a)
$$E((\hat{F}_{1,T}, \hat{F}_{2,T})^T) = (F_{1,T}, F_{2,T} + d_{N,T})^T + \tilde{C}_T + O\left(\frac{1}{N^{1-4d_\infty}}\right),$$

where the vector $\tilde{C}_T \in \mathbb{R}^2$ is of order $O(N^2/T^2 + \log(N)/(MN^{1-4d_\infty}))$ and vanishes if the functions $\psi_l(u)$ are independent of u for all $l \in \mathbb{Z}$. Here the big O notation is to be understood componentwise.

(b)
$$\text{Cov}((\hat{F}_{1,T}, \hat{F}_{2,T})^T) = \tilde{\Sigma}_T + O(T, d_\infty),$$
 where

$$O(T, d_\infty) := O\left(\frac{\log(N)}{N^{1-8d_\infty T}}\right) + O\left(\frac{N^2}{T^3} + \frac{N^{2+4d_\infty}}{T^3} 1_{\{\frac{1}{8} \leq d_\infty < \frac{1}{4}\}}\right)$$

and $\tilde{\Sigma}_T$ is the same as the matrix Σ in (3.5) except that the integral \int_0^1 is replaced by $\frac{1}{TM} \sum_{j=1}^M$.

(c) If $d_\infty < 1/8$ and $l_1, l_2 \in \mathbb{N}_0$ with $l_1 + l_2 \geq 3$, then

$$\text{cum}(\sqrt{T}\hat{F}_{1,T}\mathbf{1}_{l_1}^T, \sqrt{T}\hat{F}_{2,T}\mathbf{1}_{l_2}^T) = O(T^{(1-\frac{l_1+l_2}{2})(1-8d_\infty)}),$$

where $\mathbf{1}_{l_i} \in \mathbb{R}^{l_i}$ denotes a vector containing merely ones ($i = 1, 2$).

It follows by the same arguments as given in Section 4 of Preuß and Vetter (2013), that there exist parameters $\hat{\psi}_{l,p}$ such that the bootstrap process $X_{t,T}^*$ defined in part 2) of the bootstrap description in Section 4 can be represented as

$$X_{t,T}^* = \sum_{l=0}^{\infty} \hat{\psi}_{l,p} Z_{t-l}^*, \tag{A.1}$$

where $\{Z_t^*\}_{t \in \mathbb{Z}}$ denotes an independent sequence of standard normals. Assume that the null hypothesis (3.7) holds, and consider the process

$$X_{t,T,2}^* = \sum_{l=0}^{\infty} \psi_l Z_{t-l}^*, \tag{A.2}$$

where the coefficients $\psi_l = \psi_l(u)$ are the coefficients in (2.9). We define $\hat{D}_{T,2}^{2,*}$ as \hat{D}_T^2 in (2.17) whereby the random variables $X_{t,T}$ are replaced by $X_{t,T,2}^*$.

Theorem A.2. Let $\alpha > 0$ be fixed and denote by $A_T(\alpha)$ the event that $|\hat{d} - \underline{d}| \leq \alpha/4$. If Assumption 1 and the inequality

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

are satisfied, then

- (a)
$$E\left((\hat{D}_T^{2,*} - \hat{D}_{T,2}^{2,*}) 1_{A_T(\alpha)}\right) = O\left(p^4 \log(T)^{3/2} N^{4\hat{d}-1+\alpha} T^{-1/2}\right),$$
- (b)
$$\text{Var}\left((\hat{D}_T^{2,*} - \hat{D}_{T,2}^{2,*}) 1_{A_T(\alpha)}\right) = O\left(p^8 \log(T)^3 \log(N)^2 N^{\max(8\hat{d}-1, 0)+2\alpha} T^{-2}\right).$$

A.1. Proof of Theorem A.1.

Proof of part (a). Let $\tilde{t}_j := t_j - N/2 + 1$, $\tilde{\psi}_l(u_{j,p}) := \psi_l\left(\frac{\tilde{t}_j+p}{T}\right)$, $Z_{a,b} := Z_{a-N/2+1+b}$ and, because of

$$E[\hat{F}_{1,T}] = \frac{1}{2}E\left(\frac{1}{T} \sum_{j=1}^M \sum_{k=-\lfloor(N-1)/2\rfloor}^{\lfloor N/2\rfloor} I_N(u_j, \lambda_{k,N})^2\right) \left(1 + o\left(\frac{1}{\sqrt{T}}\right)\right),$$

consider similar to Dette, Preuß and Vetter (2011b),

$$\begin{aligned} & \frac{1}{2}E\left(\frac{1}{T} \sum_{j=1}^M \sum_{k=-\lfloor(N-1)/2\rfloor}^{\lfloor N/2\rfloor} I_N(u_j, \lambda_{k,N})^2\right) \\ &= \frac{1}{2T} \sum_{j=1}^M \sum_{k=-\lfloor(N-1)/2\rfloor}^{\lfloor N/2\rfloor} \frac{1}{(2\pi N)^2} \sum_{p,q,r,s=0}^{N-1} \sum_{l,m,n,o=0}^{\infty} e^{-i(p-q+r-s)\lambda_{k,N}} \\ & \quad \psi_{\tilde{t}_j+p,T,l} \psi_{\tilde{t}_j+q,T,m} \psi_{\tilde{t}_j+r,T,n} \psi_{\tilde{t}_j+s,T,o} E[Z_{t_j,p-l} Z_{t_j,q-m} Z_{t_j,r-n} Z_{t_j,s-o}] \\ &= E_{N,T}^1 + E_{N,T}^2 + A_{N,T} + B_{N,T}, \end{aligned}$$

where we use the notation

$$\begin{aligned} E_{N,T}^1 &= \frac{1}{2T} \sum_{j=1}^M \sum_{k=-\lfloor(N-1)/2\rfloor}^{\lfloor N/2\rfloor} \frac{1}{(2\pi N)^2} \sum_{p,q,r,s=0}^{N-1} \sum_{l,m,n,o=0}^{\infty} \psi_l(u_j) \psi_m(u_j) \psi_n(u_j) \psi_o(u_j) \\ & \quad e^{-i(p-q+r-s)\lambda_{k,N}} \left(E[Z_{t_j,p-l} Z_{t_j,q-m}] E[Z_{t_j,r-n} Z_{t_j,s-o}] \right. \\ & \quad \left. + E[Z_{t_j,p-l} Z_{t_j,s-o}] E[Z_{t_j,q-m} Z_{t_j,r-n}] \right), \\ E_{N,T}^2 &= \frac{1}{2T} \sum_{j=1}^M \sum_{k=-\lfloor(N-1)/2\rfloor}^{\lfloor N/2\rfloor} \frac{1}{(2\pi N)^2} \sum_{p,q,r,s=0}^{N-1} \sum_{l,m,n,o=0}^{\infty} \psi_l(u_j) \psi_m(u_j) \psi_n(u_j) \psi_o(u_j) \\ & \quad e^{-i(p-q+r-s)\lambda_{k,N}} E[Z_{t_j,p-l} Z_{t_j,r-n}] E[Z_{t_j,q-m} Z_{t_j,s-o}], \\ A_{N,T} &= \frac{1}{2T} \sum_{j=1}^M \sum_{k=-\lfloor(N-1)/2\rfloor}^{\lfloor N/2\rfloor} \frac{1}{(2\pi N)^2} \sum_{p,q,r,s=0}^{N-1} \sum_{l,m,n,o=0}^{\infty} e^{-i(p-q+r-s)\lambda_{k,N}} \\ & \quad E[Z_{t_j,p-l} Z_{t_j,q-m} Z_{t_j,r-n} Z_{t_j,s-o}] \left\{ \left(\tilde{\psi}_l(u_{j,p}) - \psi_l(u_j) \right) \tilde{\psi}_m(u_{j,q}) \tilde{\psi}_n(u_{j,r}) \tilde{\psi}_o(u_{j,s}) \right. \\ & \quad \left. + \psi_l(u_j) \left(\tilde{\psi}_m(u_{j,q}) - \psi_m(u_j) \right) \tilde{\psi}_n(u_{j,r}) \tilde{\psi}_o(u_{j,s}) \right. \\ & \quad \left. + \psi_l(u_j) \psi_m(u_j) \left(\tilde{\psi}_n(u_{j,r}) - \psi_n(u_j) \right) \tilde{\psi}_o(u_{j,s}) \right. \\ & \quad \left. + \psi_l(u_j) \psi_m(u_j) \psi_n(u_j) \left(\tilde{\psi}_o(u_{j,s}) - \psi_o(u_j) \right) \right\}, \end{aligned}$$

$$\begin{aligned}
 B_{N,T} = & \frac{1}{2T} \sum_{j=1}^M \sum_{k=-\lfloor(N-1)/2\rfloor}^{\lfloor N/2\rfloor} \frac{1}{(2\pi N)^2} \sum_{p,q,r,s=0}^{N-1} \sum_{l,m,n,o=0}^{\infty} e^{-i(p-q+r-s)\lambda_{k,N}} \\
 & \times E[Z_{t_j,p-l}Z_{t_j,q-m}Z_{t_j,r-n}Z_{t_j,s-o}] \\
 & \times \left\{ \left(\psi_{\tilde{t}_j+p,T,l} - \tilde{\psi}_l(u_{j,p}) \right) \tilde{\psi}_m(u_{j,q}) \tilde{\psi}_n(u_{j,r}) \tilde{\psi}_o(u_{j,s}) \right. \\
 & + \psi_{\tilde{t}_j+p,T,l} \left(\psi_{\tilde{t}_j+q,T,m} - \tilde{\psi}_m(u_{j,q}) \right) \tilde{\psi}_n(u_{j,r}) \tilde{\psi}_o(u_{j,s}) \\
 & + \psi_{\tilde{t}_j+p,T,l} \psi_{\tilde{t}_j+q,T,m} \left(\psi_{\tilde{t}_j+r,T,n} - \tilde{\psi}_n(u_{j,r}) \right) \tilde{\psi}_o(u_{j,s}) \\
 & \left. + \psi_{\tilde{t}_j+p,T,l} \psi_{\tilde{t}_j+q,T,m} \psi_{\tilde{t}_j+r,T,n} \left(\psi_{\tilde{t}_j+s,T,o} - \tilde{\psi}_o(u_{j,s}) \right) \right\}.
 \end{aligned}$$

Here $B_{N,T}$ corresponds to the error that occurs if the coefficients $\psi_{t,T,l}$ are replaced by $\psi_l(t/T)$ and that $A_{N,T}$ contains the approximation error of $\psi_l(t/T)$ through $\psi_l(t_j/T)$ with t_j denoting the midpoint of the j -th block. The proof for $E[\hat{F}_{1,T}]$ follows from

$$E_{N,T}^1 = \frac{1}{2\pi M} \sum_{j=1}^M \int_{-\pi}^{\pi} f^2(u_j, \lambda) d\lambda + O\left(\frac{1}{N^{1-4d_\infty}}\right), \tag{A.4}$$

$$E_{N,T}^2 = O\left(\frac{1}{N^{1-4d_\infty}}\right), \tag{A.5}$$

$$A_{N,T} = O\left(\frac{\log N}{MN^{1-4d_\infty}}\right) + O\left(\frac{N^2}{T^2}\right), \tag{A.6}$$

$$B_{N,T} = O\left(\frac{1}{T}\right). \tag{A.7}$$

Proof of (A.4). Without loss of generality, we only consider the first summand in $E_{N,T}^1$. Due to the independence of the random variables Z_t , only those terms contribute to the sum where the conditions $0 \leq p = q + l - m \leq N - 1$ and $0 \leq r = s + n - o \leq N - 1$ are satisfied, which implies $\max\{|l - m|, |n - o|\} \leq N - 1$. Thus, the first term in $E_{N,T}^1$ can be expressed as

$$\begin{aligned}
 & \frac{1}{2T} \sum_{j=1}^M \sum_{k=-\lfloor(N-1)/2\rfloor}^{\lfloor N/2\rfloor} \frac{1}{(2\pi N)^2} \sum_{\substack{l,m,n,o=0 \\ |l-m|\leq N-1 \\ |n-o|\leq N-1}}^{\infty} \psi_l(u_j)\psi_m(u_j)\psi_n(u_j)\psi_o(u_j)e^{-i(l-m+n-o)\lambda_{k,N}} \\
 & \times (N - |l - m|)(N - |n - o|) \\
 & = \frac{1}{2M} \sum_{j=1}^M \sum_{h=-1}^1 \frac{1}{(2\pi N)^2} \sum_{\substack{l,m,n,o=0 \\ |l-m|\leq N-1 \\ |n-o|\leq N-1 \\ l-m+n-o=hN}}^{\infty} \psi_l(u_j)\psi_m(u_j)\psi_n(u_j)\psi_o(u_j) \\
 & \times (N - |l - m|)(N - |n - o|),
 \end{aligned}$$

where we used the well-known identity,

$$\sum_{k=0}^{N-1} \exp(-i\lambda_{k,N}r) = \begin{cases} N & \text{if } r = Nh \text{ for some } h \in \mathbb{Z}, \\ 0 & \text{else,} \end{cases} \tag{A.8}$$

(we have only to consider three possible values of h since $\max\{|l-m|, |n-o|\} \leq N-1$). It is easy to see that $E_{N,T}^1 = E_{N,T,0}^1 + E_{N,T,1}^1 + E_{N,T,2}^1 + E_{N,T,3}^1$ where

$$\begin{aligned} E_{N,T,0}^1 &= \frac{1}{2M} \sum_{j=1}^M \frac{1}{(2\pi)^2} \sum_{\substack{l,m,n,o=0 \\ l-m+n-o=0}}^{\infty} \psi_l(u_j)\psi_m(u_j)\psi_n(u_j)\psi_o(u_j) \\ &= \frac{1}{4\pi M} \sum_{j=1}^M \int_{-\pi}^{\pi} f^2(u_j, \lambda) d\lambda, \\ E_{N,T,1}^1 &= \frac{1}{MN^2} \sum_{j=1}^M \sum_{h=-1}^1 \sum_{\substack{l,m,n,o=0 \\ |l-m| \leq N-1 \\ |n-o| \leq N-1 \\ l-m+n-o=hN}}^{\infty} \psi_l(u_j)\psi_m(u_j)\psi_n(u_j)\psi_o(u_j) \tag{A.9} \\ &\quad \times (-N|l-m| - N|n-o| + |l-m||n-o|), \end{aligned}$$

$$E_{N,T,2}^1 = \frac{1}{M} \sum_{j=1}^M \sum_{h \in \{-1,1\}} \sum_{\substack{l,m,n,o=0 \\ |l-m| \leq N-1 \\ |n-o| \leq N-1 \\ l-m+n-o=hN}}^{\infty} \psi_l(u_j)\psi_m(u_j)\psi_n(u_j)\psi_o(u_j), \tag{A.10}$$

$$E_{N,T,3}^1 = \frac{1}{M} \sum_{j=1}^M \sum_{\substack{l,m,n,o=0 \\ N \leq |l-m| \\ N \leq |n-o| \\ l-m+n-o=0}}^{\infty} \psi_l(u_j)\psi_m(u_j)\psi_n(u_j)\psi_o(u_j). \tag{A.11}$$

In order to complete the proof of (A.4), it therefore suffices to demonstrate that the last three expressions are of order $O(1/N^{1-4d_\infty})$. We commence with (A.9). Setting $o = l - m + n - hN \geq 0$ and using (2.3), it follows that there exists a constant $C \in \mathbb{R}$ such that

$$E_{N,T,1}^1 \leq \frac{C}{MN} \sum_{j=1}^M \sum_{h=-1}^1 \sum_{\substack{l,m,n=1 \\ 1 \leq |l-m-hN| \leq N-1 \\ 1 \leq l-m+n-hN}}^{\infty} \frac{1}{l^{1-d_\infty}} \frac{1}{m^{1-d_\infty}} \frac{1}{n^{1-d_\infty}} \frac{|l-m-hN|}{(l-m+n-hN)^{1-d_\infty}} \tag{A.12}$$

(all terms where one of the variables l, m, n or $l - m + n - hN$ vanishes are of smaller or the same order). This argument is employed throughout this proof

without mentioning it explicitly. The summand $|l - m|$ does not occur in the numerator of the above expression due to the symmetry of $|l - m|$ and $|n - o|$ in (A.9), while $C \in \mathbb{R}^+$ denotes a universal constant throughout the proof. Setting $z := l - m - hN$, we obtain $|z| = |l - m - hN| \leq N - 1$ and the expression on the right hand side of (A.12) can be written as

$$\begin{aligned} & \frac{C}{N} \sum_{h=-1}^1 \sum_{\substack{z \in \mathbb{Z} \\ 1 \leq |z| \leq N-1}} \sum_{\substack{m,n=1 \\ 1 \leq z+m+hN \\ 1 \leq n+z}}^{\infty} \frac{1}{(z+m+hN)^{1-d_\infty}} \frac{1}{m^{1-d_\infty}} \frac{1}{n^{1-d_\infty}} \frac{|z|}{(n+z)^{1-d_\infty}} \\ & \stackrel{(B.2)}{\lesssim} \frac{1}{N} \sum_{h=-1}^1 \sum_{\substack{z \in \mathbb{Z} \\ 1 \leq |z| \leq N-1}} \frac{|z|^{2d_\infty}}{|z+hN|^{1-2d_\infty}} \\ & \lesssim \frac{1}{N^{1-2d_\infty}} \sum_{h=-1}^1 \sum_{\substack{z \in \mathbb{Z} \\ 1 \leq |z| \leq N-1}} \frac{1}{|z+hN|^{1-2d_\infty}} \lesssim \frac{1}{N^{1-4d_\infty}}, \end{aligned}$$

where $a_n \lesssim b_n$ means that a_n/b_n is bounded by some finite constant for all $n \in \mathbb{N}$. By using (2.3), (A.8), (B.2), and similar arguments, we obtain that (A.10) is bounded by

$$\sum_{h \in \{-1,1\}} \sum_{\substack{z \in \mathbb{Z} \\ 1 \leq |z| \leq N-1}} \frac{1}{|z+hN|^{1-2d_\infty}} \frac{1}{|z|^{1-2d_\infty}} \stackrel{(B.1),(B.2)}{\lesssim} \frac{1}{N^{1-4d_\infty}},$$

and, since (A.11) is shown analogously, we therefore conclude the proof of (A.4).

Proof of (A.5). The result follows by similar arguments as used in the treatment of (A.9)–(A.11).

Proof of (A.6). Without loss of generality, we only consider the first summand and replace $\tilde{\psi}_m(u_{j,q})\tilde{\psi}_n(u_{j,r})\tilde{\psi}_o(u_{j,s})$ by $\psi_m(u_j)\psi_n(u_j)\psi_o(u_j)$ (the error due to this replacement is negligible, which follows by analogous arguments as given for the term $A_{N,T,1}^{(2)}$ at a later stage of this proof). Due to the independence of the random variables Z_t , we obtain the sum of three terms (compare the definition of $E_{N,T}^1$ for the first two summands and the definition of $E_{N,T}^2$ for the third one) and we restrict ourselves to the first one,

$$\begin{aligned} A_{N,T,1} & := \frac{1}{2T} \sum_{j=1}^M \sum_{k=-\lfloor (N-1)/2 \rfloor}^{\lfloor N/2 \rfloor} \frac{1}{(2\pi N)^2} \sum_{p,q,r,s=0}^{N-1} \sum_{l,m,n,o=0}^{\infty} e^{-i(p-q+r-s)\lambda_{k,N}} \\ & E[Z_{t_j,p-l}Z_{t_j,q-m}]E[Z_{t_j,r-n}Z_{t_j,s-o}] \left(\tilde{\psi}_l(u_{j,p}) - \psi_l(u_j) \right) \psi_m(u_j)\psi_n(u_j)\psi_o(u_j). \end{aligned}$$

Using a Taylor expansion, we can write

$$\tilde{\psi}_l(u_{j,p}) - \psi_l(u_j) = \psi'_l(u_j) \left(\frac{-N/2 + 1 + p}{T} \right) + \frac{\psi''(\eta_{l,j,p})}{2} \left(\frac{-N/2 + 1 + p}{T} \right)^2$$

with $\eta_{l,j,p} \in (u_j - N/(2T), u_j + N/(2T))$, and therefore $A_{N,T,1}$ splits into two terms that are denoted by $A_{N,T,1}^{(1)}$ and $A_{N,T,1}^{(2)}$ in the following discussion. We start with the treatment of the first summand. Employing the independence of the innovations we obtain that the indices corresponding to non-vanishing terms must satisfy $q = p + m - l$ and $n = o + r - s$. Applying (A.8) subsequently yields $0 \leq m = l + r - s - hN$ with $h \in \{-1, 0, 1\}$ and this, combined with (2.3) and (2.5), implies

$$A_{N,T,1}^{(1)} \lesssim \frac{1}{N^2} \sum_{h=-1}^1 \sum_{r,s=0}^{N-1} \sum_{\substack{l,o=1 \\ 1 \leq o+r-s \\ 1 \leq l+r-s-hN \\ 0 \leq |r-s-hN| \leq N-1}}^{\infty} \frac{\log(l)}{l^{1-d_\infty}} \frac{1}{(l+r-s-hN)^{1-d_\infty}} \left| \frac{1}{(o+r-s)^{1-d_\infty}} \frac{1}{o^{1-d_\infty}} \sum_{\substack{p=0 \\ 0 \leq p+r-s-hN \leq N-1}}^{N-1} \left(\frac{-N/2 + 1 + p}{T} \right) \right|.$$

We restrict ourselves to the cases $|r - s| \geq 1$ and $|r - s - hN| \geq 1$ since the remaining terms are of smaller order. A straightforward calculation yields

$$\begin{aligned} & \left| \sum_{\substack{p=0 \\ 0 \leq p+q \leq N-1}}^{N-1} \left(\frac{-N/2 + 1 + p}{T} \right) \right| \\ &= \frac{N}{2T} \times 1_{\{q=0\}} + \min \left(\frac{N|q|}{T}, \frac{(N - |q|)|q|}{T} \right) O(1) \times 1_{\{1 \leq |q| \leq N-1\}} \quad (\text{A.13}) \end{aligned}$$

and, by using the second summand, it follows that $A_{N,T,1}^{(1)}$ is bounded by

$$\begin{aligned} & \frac{C}{NT} \sum_{h=-1}^1 \sum_{r,s=0}^{N-1} \sum_{\substack{l,o=1 \\ 1 \leq o+r-s \\ 1 \leq l+r-s-hN \\ 1 \leq |r-s-hN| \leq N-1}}^{\infty} \frac{\log(l)}{l^{1-d_\infty}} \frac{1}{(l+r-s-hN)^{1-d_\infty}} \\ & \frac{1}{(o+r-s)^{1-d_\infty}} \frac{1}{o^{1-d_\infty}} |r-s-hN| \\ & \lesssim \frac{\log(N)}{N^{1-2d_\infty} T} \sum_{r,s=0}^{N-1} \sum_{\substack{o=1 \\ 1 \leq o+r-s}}^{\infty} \frac{1}{(o+r-s)^{1-d_\infty}} \frac{1}{o^{1-d_\infty}} \end{aligned}$$

$$\lesssim \frac{\log(N)}{N^{1-2d_\infty} T} \sum_{r,s=0}^{N-1} \frac{1}{|r-s|^{1-2d_\infty}} \lesssim \frac{\log(N)}{MN^{1-4d_\infty}},$$

where we used Lemma B.1(c) and B.1(b) for the first and second inequalities, respectively. Next, we show that $A_{N,T,1}^{(2)}$ is of order $O(N^2 T^{-2})$, and for this reason we choose $\epsilon > 0$ such that $1 - 4d_\infty - \epsilon > 0$. Using (2.3), (2.5), (2.6) and $h \in \{-1, 0, 1\}$, the claim then follows by a further application of Lemma B.1(b),

$$\begin{aligned} & \frac{C}{N^2} \sum_{r,s=0}^{N-1} \sum_{\substack{l,o=1 \\ 1 \leq o+r-s \\ 1 \leq l+r-s-hN}}^{\infty} \frac{\log^2(l)}{l^{1-d_\infty}} \frac{1}{(l+r-s-hN)^{1-d_\infty}} \frac{1}{(o+r-s)^{1-d_\infty}} \frac{1}{o^{1-d_\infty}} \\ & \times \sum_{\substack{p=0 \\ 0 \leq p+r-s-hN \leq N-1}}^{N-1} \left(\frac{-N/2 + 1 + p}{T} \right)^2 \\ & \lesssim \frac{N}{T^2} \sum_{r,s=0}^{N-1} \sum_{\substack{l,o=1 \\ 1 \leq o+r-s \\ 1 \leq l+r-s-hN}}^{\infty} \frac{1}{l^{1-d_\infty-\epsilon}} \frac{1}{(l+r-s-hN)^{1-d_\infty}} \frac{1}{(o+r-s)^{1-d_\infty}} \frac{1}{o^{1-d_\infty}} \\ & \lesssim \frac{N}{T^2} \sum_{r,s=0}^{N-1} \frac{1}{|r-s-hN|^{1-2d_\infty-\epsilon}} \frac{1}{|r-s|^{1-2d_\infty}} \lesssim \frac{N^2}{T^2}. \end{aligned}$$

Proof of (A.7). The statement follows from (2.2) and similar arguments as given in the proofs of (A.4) and (A.6).

In order to prove the assertion for $\hat{F}_{2,T}$, one proceeds in the same way and the details are omitted. However, it turns out that the expression corresponding to $E_{N,T}^2$ does not vanish in this case and there appears an additional bias, which is denoted by $d_{N,T}$.

Proof of part (b). We restrict ourselves to the proof of

$$\text{Var} \left(\frac{1}{T} \sum_{j=1}^M \sum_{k=-\lfloor (N-1)/2 \rfloor}^{\lfloor N/2 \rfloor} I_N(u_j, \lambda_{k,N})^2 \right) = \frac{20}{\pi} \frac{1}{TM} \sum_{j=1}^M \int_{-\pi}^{\pi} f^4(u_j, \lambda) d\lambda + O(T, d_\infty),$$

and recall the definition of the remainder

$$O(T, d_\infty) = O\left(\frac{\log(N)}{N^{1-8d_\infty} T}\right) + O\left(\frac{N^2}{T^3} + \frac{N^{2+4d_\infty}}{T^3} 1_{\{\frac{1}{8} \leq d_\infty < \frac{1}{4}\}}\right).$$

All other statements can be verified completely analogously and the details are omitted. By combining the arguments from the proof of part a) and from

Dette, Preuß and Vetter (2011b), we obtain that

$$\text{Var} \left(\frac{1}{T} \sum_{j=1}^M \sum_{k=-\lfloor (N-1)/2 \rfloor}^{\lfloor N/2 \rfloor} I_N(u_j, \lambda_{k,N})^2 \right) = 32V^*(\nu') + 8V^*(\nu'') + O(T, d_\infty),$$

where

$$\begin{aligned} V^*(\nu') &= \frac{1}{T^2} \sum_{j_1, j_2=1}^M \sum_{k_1, k_2=-\lfloor (N-1)/2 \rfloor}^{\lfloor N/2 \rfloor} \frac{1}{(2\pi N)^4} \\ &\quad \sum_{\substack{p_1, q_1, r_1, s_1=0 \\ p_2, q_2, r_2, s_2=0 \\ v_1, w_1, x_1, y_1=0 \\ v_2, w_2, x_2, y_2=0}}^{N-1} \sum_{\substack{p_1, q_1, r_1, s_1=0 \\ p_2, q_2, r_2, s_2=0 \\ v_1, w_1, x_1, y_1=0 \\ v_2, w_2, x_2, y_2=0}}^{N-1} \sum_{\substack{p_1, q_1, r_1, s_1=0 \\ p_2, q_2, r_2, s_2=0 \\ v_1, w_1, x_1, y_1=0 \\ v_2, w_2, x_2, y_2=0}}^{\infty} \sum_{\substack{p_1, q_1, r_1, s_1=0 \\ p_2, q_2, r_2, s_2=0 \\ v_1, w_1, x_1, y_1=0 \\ v_2, w_2, x_2, y_2=0}}^{\infty} \\ &\quad \psi_{v_1}(u_{j_1}) \psi_{w_1}(u_{j_1}) \psi_{x_1}(u_{j_1}) \psi_{y_1}(u_{j_1}) \psi_{v_2}(u_{j_2}) \psi_{w_2}(u_{j_2}) \psi_{x_2}(u_{j_2}) \psi_{y_2}(u_{j_2}) \\ &\quad e^{-i(p_1 - q_1 + r_1 - s_1)\lambda_{k_1}} e^{-i(p_2 - q_2 + r_2 - s_2)\lambda_{k_2}} E[Z_{t_{j_1}, p_1 - v_1} Z_{t_{j_1}, q_1 - w_1}] \\ &\quad E[Z_{t_{j_1}, r_1 - x_1} Z_{t_{j_2}, p_2 - v_2}] E[Z_{t_{j_1}, s_1 - y_1} Z_{t_{j_2}, q_2 - w_2}] E[Z_{t_{j_2}, r_2 - x_2} Z_{t_{j_2}, s_2 - y_2}], \\ V^*(\nu'') &= \frac{1}{T^2} \sum_{j_1, j_2=1}^M \sum_{k_1, k_2=-\lfloor (N-1)/2 \rfloor}^{\lfloor N/2 \rfloor} \frac{1}{(2\pi N)^4} \\ &\quad \sum_{\substack{p_1, q_1, r_1, s_1=0 \\ p_2, q_2, r_2, s_2=0 \\ v_1, w_1, x_1, y_1=0 \\ v_2, w_2, x_2, y_2=0}}^{N-1} \sum_{\substack{p_1, q_1, r_1, s_1=0 \\ p_2, q_2, r_2, s_2=0 \\ v_1, w_1, x_1, y_1=0 \\ v_2, w_2, x_2, y_2=0}}^{N-1} \sum_{\substack{p_1, q_1, r_1, s_1=0 \\ p_2, q_2, r_2, s_2=0 \\ v_1, w_1, x_1, y_1=0 \\ v_2, w_2, x_2, y_2=0}}^{\infty} \sum_{\substack{p_1, q_1, r_1, s_1=0 \\ p_2, q_2, r_2, s_2=0 \\ v_1, w_1, x_1, y_1=0 \\ v_2, w_2, x_2, y_2=0}}^{\infty} \\ &\quad \psi_{v_1}(u_{j_1}) \psi_{w_1}(u_{j_1}) \psi_{x_1}(u_{j_1}) \psi_{y_1}(u_{j_1}) \psi_{v_2}(u_{j_2}) \psi_{w_2}(u_{j_2}) \psi_{x_2}(u_{j_2}) \psi_{y_2}(u_{j_2}) \\ &\quad e^{-i(p_1 - q_1 + r_1 - s_1)\lambda_{k_1}} e^{-i(p_2 - q_2 + r_2 - s_2)\lambda_{k_2}} E[Z_{t_{j_1}, p_1 - v_1} Z_{t_{j_2}, p_2 - v_2}] \\ &\quad E[Z_{t_{j_1}, r_1 - x_1} Z_{t_{j_2}, r_2 - x_2}] E[Z_{t_{j_1}, q_1 - w_1} Z_{t_{j_2}, q_2 - w_2}] E[Z_{t_{j_1}, s_1 - y_1} Z_{t_{j_2}, s_2 - y_2}]. \end{aligned}$$

We start with $V^*(\nu')$. Because of the independence of the random variables Z_t , the restrictions $p_1 = q_1 + v_1 - w_1$, $p_2 = r_1 + v_2 - x_1 + (j_1 - j_2)N$, $q_2 = s_1 + w_2 - y_1 + (j_1 - j_2)N$ and $s_2 = r_2 + y_2 - x_2$ are necessary for a non-vanishing term. Consider $h_1, h_2 \in \{-1, 0, 1\}$ and sum over k_1, k_2 by using (A.8). Then, $V^*(\nu')$ can be written as

$$\begin{aligned} V^*(\nu') &= \frac{1}{T^2} \sum_{j_1, j_2=1}^M \sum_{k_1, k_2=-\lfloor (N-1)/2 \rfloor}^{\lfloor N/2 \rfloor} \frac{1}{(2\pi N)^4} \sum_{\substack{q_1, r_1, s_1, r_2=0 \\ v_1, w_1, x_1, y_1=0 \\ 0 \leq q_1 + v_1 - w_1 \leq N-1}}^{N-1} \sum_{\substack{v_2, w_2, x_2, y_2=0 \\ 0 \leq r_1 + v_2 - x_1 + (j_1 - j_2)N \leq N-1 \\ 0 \leq s_1 + w_2 - y_1 + (j_1 - j_2)N \leq N-1 \\ 0 \leq r_2 + y_2 - x_2 \leq N-1}}^{\infty} \\ &\quad \psi_{v_1}(u_{j_1}) \psi_{w_1}(u_{j_1}) \psi_{x_1}(u_{j_1}) \psi_{y_1}(u_{j_1}) \psi_{v_2}(u_{j_2}) \psi_{w_2}(u_{j_2}) \psi_{x_2}(u_{j_2}) \psi_{y_2}(u_{j_2}) \end{aligned}$$

$$\begin{aligned}
 & e^{-i(v_1-w_1+r_1-s_1)\lambda_{k_1}} e^{-i(r_1+v_2-x_1-s_1-w_2+y_1-y_2+x_2)\lambda_{k_2}} \\
 = & \frac{1}{M^2} \sum_{j_1, j_2=1}^M \sum_{h_1, h_2=-1}^1 \frac{1}{(2\pi N)^4} \sum_{r_1, s_1=0}^{N-1} \sum_{\substack{v_1, w_1, x_1, y_1=0 \\ v_1-w_1+r_1-s_1=h_1 N \\ 0 \leq |v_1-w_1| \leq N-1}}^{\infty} \\
 & \sum_{\substack{v_2, w_2, x_2, y_2=0 \\ r_1+v_2-x_1-s_1-w_2+y_1-y_2+x_2=h_2 N \\ 0 \leq r_1+v_2-x_1+(j_1-j_2)N \leq N-1 \\ 0 \leq s_1+w_2-y_1+(j_1-j_2)N \leq N-1 \\ 0 \leq |y_2-x_2| \leq N-1}}^{\infty} \\
 & \psi_{v_1}(u_{j_1}) \psi_{w_1}(u_{j_1}) \psi_{x_1}(u_{j_1}) \psi_{y_1}(u_{j_1}) \psi_{v_2}(u_{j_2}) \psi_{w_2}(u_{j_2}) \psi_{x_2}(u_{j_2}) \psi_{y_2}(u_{j_2}) \\
 & (N^2 - N|y_2 - x_2| - N|v_1 - w_1| + |v_1 - w_1||y_2 - x_2|).
 \end{aligned}$$

An application of (2.3) yields, similar to the proof of part a), that the above expression is of order $O(1/N^{1-8d_\infty} T)$, if

- (i) $h_1, h_2 \in \{-1, 1\}$ (compare (A.10)),
- (ii) $j_1 \neq j_2$ (we prove this claim in Lemma B.2 in the appendix since this kind of restriction did not occur in the proof of part a)),
- (iii) we drop $-N|y_2 - x_2| - N|v_1 - w_1| + |v_1 - w_1||y_2 - x_2|$ (compare (A.9)),
- (iv) we drop $0 \leq |v_1 - w_1| \leq N - 1$ and $0 \leq |y_2 - x_2| \leq N - 1$ (compare (A.11)),
- (v) we drop $0 \leq r_1 + v_2 - x_1 + (j_1 - j_2)N \leq N - 1$ and $0 \leq s_1 + w_2 - y_1 + (j_1 - j_2)N \leq N - 1$ (compare (A.9)).

By rearranging the equation $v_1 - w_1 + r_1 - s_1 = 0$ to $0 \leq s_1 = r_1 + v_1 - w_1 \leq N - 1$, it follows that

$$\begin{aligned}
 V^*(\nu') &= \frac{1}{M^2 N^2} \sum_{j_1=1}^M \frac{1}{(2\pi)^4} \sum_{r_1=0}^{N-1} \sum_{\substack{v_1, w_1, x_1, y_1=0 \\ 0 \leq r_1 + v_1 - w_1 \leq N-1}}^{\infty} \sum_{\substack{v_2, w_2, x_2, y_2=0 \\ w_1 - v_1 + v_2 - x_1 - w_2 + y_1 - y_2 + x_2 = 0}}^{\infty} \\
 & \psi_{v_1}(u_{j_1}) \psi_{w_1}(u_{j_1}) \psi_{x_1}(u_{j_1}) \psi_{y_1}(u_{j_1}) \psi_{v_2}(u_{j_1}) \psi_{w_2}(u_{j_1}) \psi_{x_2}(u_{j_1}) \psi_{y_2}(u_{j_1}) \\
 & + O(T, d_\infty) \\
 &= \frac{1}{M^2 N} \sum_{j_1=1}^M \frac{1}{(2\pi)^4} \sum_{v_1, w_1, x_1, y_1=0}^{\infty} \sum_{\substack{v_2, w_2, x_2, y_2=0 \\ w_1 - v_1 + v_2 - w_2 + y_1 - x_1 + x_2 - y_2 = 0}}^{\infty} \\
 & \psi_{v_1}(u_{j_1}) \psi_{w_1}(u_{j_1}) \psi_{x_1}(u_{j_1}) \psi_{y_1}(u_{j_1}) \psi_{v_2}(u_{j_1}) \psi_{w_2}(u_{j_1}) \psi_{x_2}(u_{j_1}) \psi_{y_2}(u_{j_1}) \\
 & + O(T, d_\infty) \\
 &= \frac{1}{TM} \frac{1}{2\pi} \sum_{j_1=1}^M \int_{-\pi}^{\pi} f^4(u_{j_1}, \lambda) d\lambda + O(T, d_\infty).
 \end{aligned}$$

By using the same techniques as in $V^*(\nu')$, we obtain

$$\begin{aligned}
 V^*(\nu'') &= \frac{1}{M^2} \sum_{j_1=1}^M \sum_{h_1=-1}^1 \frac{1}{(2\pi N)^4} \sum_{\substack{q_2, r_2, s_2=0 \\ 0 \leq q_2 - r_2 + s_2 + h_1 N \leq N-1}}^{N-1} \sum_{\substack{v_1, w_1, x_1, y_1=0 \\ v_2, w_2, x_2, y_2=0 \\ v_1 - v_2 + w_2 - w_1 + x_1 - x_2 + y_2 - y_1=0}}^{\infty} \\
 &\quad \psi_{v_1}(u_{j_1}) \psi_{w_1}(u_{j_1}) \psi_{x_1}(u_{j_1}) \psi_{y_1}(u_{j_1}) \psi_{v_2}(u_{j_1}) \psi_{w_2}(u_{j_1}) \psi_{x_2}(u_{j_1}) \psi_{y_2}(u_{j_1}) \\
 &\quad + O(T, d_\infty).
 \end{aligned}$$

In contrast to the term $V^*(\nu')$, the cases where $h_1 \in \{-1, 1\}$ do not vanish. In fact, using

$$\begin{aligned}
 \sum_{\substack{q_2, r_2, s_2=0 \\ 0 \leq q_2 - r_2 + s_2 \leq N-1}}^{N-1} &= \frac{2}{3} N^3 + O(N^2), & \sum_{\substack{q_2, r_2, s_2=0 \\ 0 \leq q_2 - r_2 + s_2 + N \leq N-1}}^{N-1} &= \frac{1}{6} N^3 + O(N^2), \\
 \sum_{\substack{q_2, r_2, s_2=0 \\ 0 \leq q_2 - r_2 + s_2 - N \leq N-1}}^{N-1} &= \frac{1}{6} N^3 + O(N^2)
 \end{aligned}$$

we deduce

$$\begin{aligned}
 V^*(\nu'') &= \left(\frac{2}{3} + \frac{1}{6} + \frac{1}{6} \right) \frac{1}{M^2 N} \sum_{j_1=1}^M \frac{1}{(2\pi)^4} \sum_{v_1, w_1, x_1, y_1=0}^{\infty} \sum_{\substack{v_2, w_2, x_2, y_2=0 \\ v_1 - v_2 + w_2 - w_1 + x_1 - x_2 + y_2 - y_1=0}}^{\infty} \\
 &\quad \psi_{v_1}(u_{j_1}) \psi_{w_1}(u_{j_1}) \psi_{x_1}(u_{j_1}) \psi_{y_1}(u_{j_1}) \psi_{v_2}(u_{j_1}) \psi_{w_2}(u_{j_1}) \psi_{x_2}(u_{j_1}) \psi_{y_2}(u_{j_1}) \\
 &\quad + O(T, d_\infty) \\
 &= \frac{1}{TM} \frac{1}{2\pi} \sum_{j_1=1}^M \int_{-\pi}^{\pi} f^4(u_{j_1}, \lambda) d\lambda + O(T, d_\infty). \tag{A.14}
 \end{aligned}$$

Proof of part (c) Exemplarily we consider the case $l_2 = 0$ and $l := l_1 \geq 3$. The other cases can be treated similarly with an additional amount of notation. Following the same lines as in the proof of Theorem 3.1 in Dette, Preuß and Vetter (2011a), it is sufficient to choose an arbitrary indecomposable partition

$$\left\{ (Z_{t_{i_1}, a_1 - v_1} Z_{t_{i_2}, a_2 - w_1}), (Z_{t_{i_3}, a_3 - x_1} Z_{t_{i_4}, a_4 - y_1}), \dots, (Z_{t_{i_{4l-1}}, a_{4l-1} - x_l} Z_{t_{i_{4l}}, a_{4l} - y_l}) \right\}$$

of the table

$$\begin{array}{cccc}
 Z_{t_{j_1}, p_1 - g_1} & Z_{t_{j_1}, q_1 - m_1} & Z_{t_{j_1}, r_1 - n_1} & Z_{t_{j_1}, s_1 - o_1} \\
 \vdots & \vdots & \vdots & \vdots \\
 Z_{t_{j_l}, p_l - g_l} & Z_{t_{j_l}, q_l - m_l} & Z_{t_{j_l}, r_l - n_l} & Z_{t_{j_l}, s_l - o_l}
 \end{array} \tag{A.15}$$

(see Brillinger (1981)) and to treat the term

$$\begin{aligned} & \frac{T^{l/2}}{T^l} \frac{1}{N^{2l}} \sum_{j_1, \dots, j_l=1}^M \sum_{k_1, \dots, k_l=-\lfloor (N-1)/2 \rfloor}^{\lfloor N/2 \rfloor} \sum_{p_1, \dots, s_l=0}^{N-1} \sum_{v_1, w_1, x_1, y_1=0}^{\infty} \dots \sum_{v_l, w_l, x_l, y_l=0}^{\infty} \\ & \tilde{\psi}_{v_1}(u_{j_1, p_1}) \tilde{\psi}_{w_1}(u_{j_1, q_1}) \tilde{\psi}_{x_1}(u_{j_1, r_1}) \tilde{\psi}_{y_1}(u_{j_1, s_1}) \\ & \tilde{\psi}_{v_2}(u_{j_2, p_2}) \tilde{\psi}_{w_2}(u_{j_2, q_2}) \tilde{\psi}_{x_2}(u_{j_2, r_2}) \tilde{\psi}_{y_2}(u_{j_2, s_2}) \\ & \dots \\ & \tilde{\psi}_{v_l}(u_{j_l, p_l}) \tilde{\psi}_{w_l}(u_{j_l, q_l}) \tilde{\psi}_{x_l}(u_{j_l, r_l}) \tilde{\psi}_{y_l}(u_{j_l, s_l}) e^{-i(p_1 - q_1 + r_1 - s_1)\lambda_{k_1}} \dots e^{-i(p_l - q_l + r_l - s_l)\lambda_{k_l}} \\ & E[Z_{t_{i_1}, a_1 - v_1} Z_{t_{i_2}, a_2 - w_1}] E[Z_{t_{i_3}, a_3 - x_1} Z_{t_{i_4}, a_4 - y_1}] \\ & E[Z_{t_{i_5}, a_5 - v_2} Z_{t_{i_6}, a_6 - w_2}] E[Z_{t_{i_7}, a_7 - x_2} Z_{t_{i_8}, a_8 - y_2}] \\ & \dots \\ & E[Z_{t_{i_{4l-3}}, a_{4l-3} - v_l} Z_{t_{i_{4l-2}}, a_{4l-2} - w_l}] E[Z_{t_{i_{4l-1}}, a_{4l-1} - x_l} Z_{t_{i_{4l}}, a_{4l} - y_l}] \end{aligned}$$

with $\{a_1, a_2, \dots, a_{4l}\} \in \{p_1, \dots, p_l, q_1, \dots, q_l, r_1, \dots, r_l, s_1, \dots, s_l\}$, $a_i \neq a_j$ for $i \neq j$, $\{i_1, i_2, \dots, i_{4l}\} \in \{j_1, j_2, \dots, j_l\}$, and $|\{i_1, i_2, \dots, i_{4l}\}| = l$. We now discuss the conditions which yield a contribution different from 0 in this sum. Some of the i_k are equal to each other and we therefore write j_1, \dots, j_l for the l different values and consider i_k as a function depending on j_1, \dots, j_l . Using the independence of the random variables Z_t and summing with respect to k_1, \dots, k_l , the conditions

$$a_{4m+1} - a_{4m+2} + w_{m+1} - v_{m+1} + (i_{4m+1} - i_{4m+2})N = 0 \text{ for } m=0, \dots, l-1, \tag{A.16}$$

$$a_{4m+3} - a_{4m+4} + y_{m+1} - x_{m+1} + (i_{4m+3} - i_{4m+4})N = 0 \text{ for } m=0, \dots, l-1, \tag{A.17}$$

$$p_i - q_i + r_i - s_i = h_i N \text{ for } i=1, 2, \dots, l \text{ and } h_i \in \{-1, 0, 1\} \tag{A.18}$$

follow. Rearranging the equations in (A.18) for a variable and plugging them into the l equations (A.16) (where in every equation only one variable is replaced) yields, due to the indecomposability of the partition and $v_{m+1}, x_{m+1} \geq 0$, that the conditions

$$\begin{aligned} (1) \quad & 0 \leq v_1 = \tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + w_1 + (i_1 - i_2 + h_1)N, \\ (2) \quad & 0 \leq v_2 = \tilde{a}_7 - \tilde{a}_8 + \tilde{a}_9 - \tilde{a}_{10} + w_2 + (i_5 - i_6 + h_2)N, \\ & \vdots \\ (l) \quad & 0 \leq v_l = \tilde{a}_{6l-5} - \tilde{a}_{6l-4} + \tilde{a}_{6l-3} - \tilde{a}_{6l-2} + w_l + (i_{4l-3} - i_{4l-2} + h_l)N, \tag{A.19} \\ (l+1) \quad & 0 \leq x_1 = \tilde{a}_5 - \tilde{a}_6 + y_1 + (i_3 - i_4)N, \\ (l+2) \quad & 0 \leq x_2 = \tilde{a}_{11} - \tilde{a}_{12} + y_2 + (i_7 - i_8)N, \\ & \vdots \\ (2l) \quad & 0 \leq x_l = \tilde{a}_{6l-1} - \tilde{a}_{6l} + y_l + (i_{4l-1} - i_{4l})N \end{aligned}$$

must hold, where $\{\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_{6l}\} \in \{p_1, \dots, s_l\}$ and $|\{\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_{6l}\}| = 3l$. By employing (2.3), we can bound the above expression up to a constant by

$$\begin{aligned}
& \frac{1}{M^l} \frac{T^{l/2}}{N^{2l}} \sum_{j_1, \dots, j_l=1}^M \sum_{h_1, \dots, h_l=-1}^1 \sum_{\tilde{a}_1, \dots, \tilde{a}_{6l}=0}^{N-1} \\
& \quad \sum_{\substack{w_1, y_1=1 \\ \tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + w_1 + (i_1 - i_2 + h_1)N \geq 1 \\ \tilde{a}_5 - \tilde{a}_6 + y_1 + (i_3 - i_4)N \geq 1}}^{\infty} \quad \dots \quad \sum_{\substack{w_l, y_l=1 \\ \tilde{a}_{6l-5} - \tilde{a}_{6l-4} + \tilde{a}_{6l-3} - \tilde{a}_{6l-2} + w_l + (i_{4l-3} - i_{4l-2} + h_l)N \geq 1 \\ \tilde{a}_{6l-1} - \tilde{a}_{6l} + y_l + (i_{4l-1} - i_{4l})N \geq 1}}^{\infty} \\
& \quad \frac{1}{(\tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + w_1 + (i_1 - i_2 + h_1)N)^{1-d_\infty}} \\
& \quad \frac{1}{w_1^{1-d_\infty}} \frac{1}{(\tilde{a}_5 - \tilde{a}_6 + y_1 + (i_3 - i_4)N)^{1-d_\infty}} \frac{1}{y_1^{1-d_\infty}} \\
& \quad \frac{1}{(\tilde{a}_7 - \tilde{a}_8 + \tilde{a}_9 - \tilde{a}_{10} + w_2 + (i_5 - i_6 + h_2)N)^{1-d_\infty}} \\
& \quad \frac{1}{w_2^{1-d_\infty}} \frac{1}{(\tilde{a}_{11} - \tilde{a}_{12} + y_2 + (i_7 - i_8)N)^{1-d_\infty}} \frac{1}{y_2^{1-d_\infty}} \\
& \quad \vdots \\
& \quad \frac{1}{(\tilde{a}_{6l-5} - \tilde{a}_{6l-4} + \tilde{a}_{6l-3} - \tilde{a}_{6l-2} + w_l + (i_{4l-3} - i_{4l-2} + h_l)N)^{1-d_\infty}} \frac{1}{w_l^{1-d_\infty}} \\
& \quad \frac{1}{(\tilde{a}_{6l-1} - \tilde{a}_{6l} + y_l + (i_{4l-1} - i_{4l})N)^{1-d_\infty}} \frac{1}{y_l^{1-d_\infty}}.
\end{aligned}$$

Using Lemma B.1(b) in the Appendix, this term can be (up to a constant) bounded by

$$\begin{aligned}
& \frac{1}{M^l} \frac{T^{l/2}}{N^{2l}} \sum_{j_1, \dots, j_l=1}^M \sum_{h_1, \dots, h_l=-1}^1 \sum_{\substack{\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_6=0 \\ |\tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + (i_1 - i_2 + h_1)N| \geq 1 \\ |\tilde{a}_5 - \tilde{a}_6 + (i_3 - i_4)N| \geq 1}}^{N-1} \quad \dots \\
& \quad \sum_{\substack{\tilde{a}_{6l-5}, \tilde{a}_{6l-4}, \dots, \tilde{a}_{6l}=0 \\ |\tilde{a}_{6l-5} - \tilde{a}_{6l-4} + \tilde{a}_{6l-3} - \tilde{a}_{6l-2} + (i_{4l-3} - i_{4l-2} + h_l)N| \geq 1 \\ |\tilde{a}_{6l-1} - \tilde{a}_{6l} + (i_{4l-1} - i_{4l})N| \geq 1}}^{N-1} \\
& \quad \frac{1}{|\tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + (i_1 - i_2 + h_1)N|^{1-2d_\infty}} \frac{1}{|\tilde{a}_5 - \tilde{a}_6 + (i_3 - i_4)N|^{1-2d_\infty}} \\
& \quad \frac{1}{|\tilde{a}_7 - \tilde{a}_8 + \tilde{a}_9 - \tilde{a}_{10} + (i_5 - i_6 + h_2)N|^{1-2d_\infty}} \frac{1}{|\tilde{a}_{11} - \tilde{a}_{12} + (i_7 - i_8)N|^{1-2d_\infty}} \\
& \quad \dots
\end{aligned}$$

$$\frac{1}{|\tilde{a}_{6l-5} - \tilde{a}_{6l-4} + \tilde{a}_{6l-3} - \tilde{a}_{6l-2} + (i_{4l-3} - i_{4l-2} + h_l)N|^{1-2d_\infty}} \cdot \frac{1}{|\tilde{a}_{6l-1} - \tilde{a}_{6l} + (i_{4l-1} - i_{4l})N|^{1-2d_\infty}}.$$

We now assume without loss of generality that

$$\begin{aligned} \tilde{a}_{6m+1} - \tilde{a}_{6m+2} + \tilde{a}_{6m+3} - \tilde{a}_{6m+4} + (i_{4m+1} - i_{4m+2} + h_{m+1})N &\geq 1, \\ \tilde{a}_{6m+5} - \tilde{a}_{6m+6} + (i_{4m+3} - i_{4m+4})N &\geq 1 \end{aligned} \tag{A.20}$$

holds for $m = 0, 1, 2, \dots, l - 1$ (the more general case follows analogously with an additional amount of notation). In this case the absolute values in the above expression can be skipped. It follows, as in Dette, Preuß and Vetter (2011a), that the conditions on the \tilde{a}_i imply that, if i_1 is chosen, there are only finitely many possible choices for $i_k, k = 2, \dots, l$. Thus it suffices to consider the sum

$$\begin{aligned} &\frac{1}{M^{l-1}} \frac{T^{l/2}}{N^{2l}} \sum_{\substack{\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_6=0 \\ \tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + C_1 N \geq 1 \\ \tilde{a}_5 - \tilde{a}_6 + C_{l+1} N \geq 1}}^{N-1} \dots \sum_{\substack{\tilde{a}_{6l-5}, \tilde{a}_{6l-4}, \dots, \tilde{a}_{6l}=0 \\ \tilde{a}_{6l-5} - \tilde{a}_{6l-4} + \tilde{a}_{6l-3} - \tilde{a}_{6l-2} + C_l N \geq 1 \\ \tilde{a}_{6l-1} - \tilde{a}_{6l} + C_{2l} N \geq 1}}^{N-1} \\ &\frac{1}{(\tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + C_1 N)^{1-2d_\infty}} \frac{1}{(\tilde{a}_5 - \tilde{a}_6 + C_{l+1} N)^{1-2d_\infty}} \\ &\frac{1}{(\tilde{a}_7 - \tilde{a}_8 + \tilde{a}_9 - \tilde{a}_{10} + C_2 N)^{1-2d_\infty}} \frac{1}{(\tilde{a}_{11} - \tilde{a}_{12} + C_{l+2} N)^{1-2d_\infty}} \\ &\dots \\ &\frac{1}{(\tilde{a}_{6l-5} - \tilde{a}_{6l-4} + \tilde{a}_{6l-3} - \tilde{a}_{6l-2} + C_l N)^{1-2d_\infty}} \frac{1}{(\tilde{a}_{6l-1} - \tilde{a}_{6l} + C_{2l} N)^{1-2d_\infty}} \end{aligned}$$

with $C_1, C_2, \dots, C_l \in \{-1, 0, \dots, M\}$ and $C_{l+1}, C_{l+2}, \dots, C_{2l} \in \{0, 1, \dots, M - 1\}$ (because of (A.20) and $\tilde{a}_i \in \{0, 1, 2, \dots, N - 1\}$, there are no other possible values for C_i). We remind that (due to the indecomposability of the partition) the $2l$ -fractions inside the addend are hooked. This means that for two different fractions there exists a chain of fractions (starting with the first considered fraction and ending with the second one), such that in every element of the chain there exists at least one element \tilde{a}_i which also occurs in the consecutive fraction. We perform a summation in a particular way and in order to illustrate this, we consider the first two fractions and assume that \tilde{a}_1 and \tilde{a}_6 are (up to a the algebraic sign) the same. We distinguish two cases.

(i) If $\tilde{a}_1 = \tilde{a}_6$, we obtain from Lemma B.1 (a) that

$$\sum_{\substack{\tilde{a}_1=0 \\ \tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + C_1 N \geq 1 \\ \tilde{a}_5 - \tilde{a}_6 + C_{l+1} N \geq 1}}^{N-1} \frac{1}{(\tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + C_1 N)^{1-2d_\infty}} \frac{1}{(\tilde{a}_5 - \tilde{a}_6 + C_{l+1} N)^{1-2d_\infty}}$$

$$\lesssim \frac{1}{(-\tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + \tilde{a}_5 + (C_1 + C_{l+1})N)^{1-4d_\infty}} \quad (\text{A.21})$$

$$\lesssim \frac{T^{2d_\infty}}{(-\tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + \tilde{a}_5 + (C_1 + C_{l+1})N)^{1-2d_\infty}}. \quad (\text{A.22})$$

Furthermore we have $-\tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + \tilde{a}_5 + (C_1 + C_{l+1})N \geq 2$ which follows from the conditions $\tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + C_1N \geq 1$ and $\tilde{a}_5 - \tilde{a}_6 + C_{l+1}N = \tilde{a}_5 - \tilde{a}_1 + C_{l+1}N \geq 1$.

(ii) If $\tilde{a}_1 = -\tilde{a}_6$ and $-\tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 - \tilde{a}_5 + (C_1 - C_{l+1})N \neq 0$, it follows from Lemma B.1(b) that

$$\sum_{\substack{\tilde{a}_1=0 \\ \tilde{a}_1-\tilde{a}_2+\tilde{a}_3-\tilde{a}_4+C_1N \geq 1 \\ \tilde{a}_5-\tilde{a}_6+C_{l+1}N \geq 1}}^{N-1} \frac{1}{(\tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + C_1N)^{1-2d_\infty}} \frac{1}{(\tilde{a}_5 - \tilde{a}_6 + C_{l+1}N)^{1-2d_\infty}} \\ \lesssim \frac{1}{|-\tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 - \tilde{a}_5 + (C_1 - C_{l+1})N|^{1-4d_\infty}} \quad (\text{A.23})$$

$$\lesssim \frac{T^{2d_\infty}}{|-\tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 - \tilde{a}_5 + (C_1 - C_{l+1})N|^{1-2d_\infty}}. \quad (\text{A.24})$$

In both cases, it is possible that variables cancel out, for example if $\tilde{a}_4 = \tilde{a}_5$ and $\tilde{a}_3 = \tilde{a}_6$ in the first and second case, respectively. We apply (A.21)–(A.24) in total $2l - 2$ -times. In the first $2l - 4$ -applications, we use (A.22) and (A.24) (depending on the algebraic sign of the variable that appears in both fractions) and in the $(2l - 3)$ th and $(2l - 2)$ th application we employ (A.21) and (A.23). We furthermore assume that h variables cancel out while utilizing these inequalities. Then $3l - (2l - 2) - h = l + 2 - h$ variables remain with $0 \leq h \leq l$, namely \tilde{a}_{6l-1} , \tilde{a}_{6l} and $l - h$ other variables with values in $\{0, 1, 2, \dots, N - 1\}$. Denoting these $l - h$ variables by b_1, b_2, \dots, b_{l-h} , we obtain

$$\frac{1}{M^{l-1}} \frac{T^{1/2}}{N^{2l}} \sum_{\substack{\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_6=0 \\ \tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + C_1N \geq 1 \\ \tilde{a}_5 - \tilde{a}_6 + C_{l+1}N \geq 1}}^{N-1} \dots \sum_{\substack{\tilde{a}_{6l-5}, \tilde{a}_{6l-4}, \dots, \tilde{a}_{6l}=0 \\ \tilde{a}_{6l-5} - \tilde{a}_{6l-4} + \tilde{a}_{6l-3} - \tilde{a}_{6l-2} + C_lN \geq 1 \\ \tilde{a}_{6l-1} - \tilde{a}_{6l} + C_{2l}N \geq 1}}^{N-1} \\ \frac{1}{(\tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + C_1N)^{1-2d_\infty}} \frac{1}{(\tilde{a}_5 - \tilde{a}_6 + C_{l+1}N)^{1-2d_\infty}} \\ \frac{1}{(\tilde{a}_7 - \tilde{a}_8 + \tilde{a}_9 - \tilde{a}_{10} + C_2N)^{1-2d_\infty}} \frac{1}{(\tilde{a}_{11} - \tilde{a}_{12} + C_{l+2}N)^{1-2d_\infty}} \\ \dots \\ \frac{1}{(\tilde{a}_{6l-5} - \tilde{a}_{6l-4} + \tilde{a}_{6l-3} - \tilde{a}_{6l-2} + C_lN)^{1-2d_\infty}} \frac{1}{(\tilde{a}_{6l-1} - \tilde{a}_{6l} + C_{2l}N)^{1-2d_\infty}}$$

$$\begin{aligned} &\lesssim \frac{1}{M^{l-1}} \frac{T^{1/2}}{N^{2l}} \sum_{\substack{\tilde{a}_{6l-1}, \tilde{a}_{6l}=0 \\ \tilde{a}_{6l-1} - \tilde{a}_{6l} + C_{2l}N \geq 1}}^{N-1} \sum_{\substack{b_1, b_2, \dots, b_{l-h}=0 \\ 1 \leq |\tilde{a}_{6l-1} - \tilde{a}_{6l} + \sum_{j=1}^{l-h} (-1)^{k_j} 2b_j + \sum_{j=1}^{2l-1} (-1)^{k_j} C_j N|}}^{N-1} \\ &\frac{N^h T^{(2l-4)2d_\infty}}{|\tilde{a}_{6l-1} - \tilde{a}_{6l} + \sum_{j=1}^{l-h} (-1)^{k_j} 2b_j + \sum_{j=1}^{2l-1} (-1)^{k_j} C_j N|^{1-6d_\infty}} \frac{1}{(\tilde{a}_{6l-1} - \tilde{a}_{6l} + C_{2l}N)^{1-2d_\infty}} \end{aligned}$$

with some $k_j \in \{0, 1\}$. We first consider the case $h = l$. If $\sum_{j=1}^{2l-1} (-1)^{k_j} C_j N = 0$ and $C_{2l} = 0$, it follows that this term is

$$\begin{aligned} &\frac{1}{M^{l-1}} \frac{T^{1/2} N^l T^{(2l-4)2d_\infty}}{N^{2l}} \sum_{\substack{\tilde{a}_{6l-1}, \tilde{a}_{6l}=0 \\ 1 \leq \tilde{a}_{6l-1} - \tilde{a}_{6l}}}^{N-1} \frac{1}{(\tilde{a}_{6l-1} - \tilde{a}_{6l})^{2-8d_\infty}} \\ &\lesssim \frac{1}{M^{l-1}} \frac{T^{1/2} N^{l+1} T^{(2l-4)2d_\infty}}{N^{2l}} = T^{(1-\frac{1}{2})(1-8d_\infty)}. \end{aligned}$$

If $|\sum_{j=1}^{2l-1} (-1)^{k_j} C_j N| \geq 1$ or $C_{2l} = 1$, we apply Lemma B.1(a) and (b) in order to obtain the same upper bound (it can be shown that, in this case, there appears an additional factor N^{1-8d_∞} in the denominator, so the corresponding term is, in fact, of smaller order). The same upper bound holds for $h \leq l - 1$.

A.2. Proof of Remark 5.

If we replace $\hat{F}_{1,T}$ by the corresponding integrated version

$$\tilde{F}_{1,T} = \frac{1}{4\pi M} \sum_{j=1}^M \int_{-\pi}^{\pi} I_N(u_j, \lambda)^2 d\lambda,$$

the derivation of the asymptotic variance can be carried out almost analogously to the proof of Theorem A.1(b), except that the term where the variable h_1 in $V^*(v'')$ is -1 or 1 , does not occur, because for the integrated version one can use

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(-i\lambda r) d\lambda = \begin{cases} 1 & \text{if } r = 0, \\ 0 & \text{else,} \end{cases} \quad (\text{A.25})$$

(for $r \in \mathbb{Z}$) instead of (A.8). Therefore, in the integrated case, we obtain

$$\begin{aligned} V^*(v'') &= \frac{2}{3} \frac{1}{M^2 N} \sum_{j_1=1}^M \frac{1}{(4\pi)^2} \frac{1}{(2\pi)^2} \sum_{\substack{v_1, w_1, x_1, y_1=0 \\ v_1 - v_2 + w_2 - w_1 + x_1 - x_2 + y_2 - y_1=0}}^{\infty} \sum_{\substack{v_2, w_2, x_2, y_2=0 \\ v_1 - v_2 + w_2 - w_1 + x_1 - x_2 + y_2 - y_1=0}}^{\infty} \\ &\psi_{v_1}(u_{j_1}) \psi_{w_1}(u_{j_1}) \psi_{x_1}(u_{j_1}) \psi_{y_1}(u_{j_1}) \psi_{v_2}(u_{j_1}) \psi_{w_2}(u_{j_1}) \psi_{x_2}(u_{j_1}) \psi_{y_2}(u_{j_1}) \end{aligned}$$

$$\begin{aligned}
& +O(T, d_\infty) \\
& = \frac{2}{3} \frac{1}{TM} \frac{1}{8\pi} \sum_{j_1=1}^M \int_{-\pi}^{\pi} f^4(u_{j_1}, \lambda) d\lambda + O(T, d_\infty)
\end{aligned}$$

instead of (A.14), and we recall that the order $O(T, d_\infty)$ is defined in (A.1). This yields that the asymptotic variance of $\sqrt{T}\tilde{F}_{1,T}$ is $(14/3\pi) \int_{-\pi}^{\pi} \int_0^1 f^4(u, \lambda) dud\lambda$ and does not coincide with the asymptotic variance of $\sqrt{T}\hat{F}_{1,T}$.

A.3. Proof of Theorem A.2

Proof of part (a). We define $\hat{F}_{1,T}^*$ and $\hat{F}_{1,T,2}^*$ as $\hat{F}_{1,T}$ where the observed data $X_{t,T}$ are replaced by $X_{t,T}^*$ and $X_{t,T,2}^*$, respectively. By using (A.1) and writing $I_N^*(u, \lambda)$ for the bootstrap analogue of $I_N(u, \lambda)$, we get

$$\begin{aligned}
& E((\hat{F}_{1,T}^* - \hat{F}_{1,T,2}^*)1_{A_T(\alpha)} | X_{1,T}, \dots, X_{T,T}) \\
& = \frac{1}{2T} \sum_{j=1}^M \sum_{k=-\lfloor(N-1)/2\rfloor}^{\lfloor N/2\rfloor} \frac{1}{(2\pi N)^2} \sum_{p,q,r,s=0}^{N-1} \sum_{l,m,n,o=0}^{\infty} e^{-i(p-q+r-s)\lambda_{k,N}} \hat{\psi}_{l,m,n,o,p} 1_{A_T(\alpha)} \\
& \quad E[Z_{t_j,p-l}^* Z_{t_j,q-m}^* Z_{t_j,r-n}^* Z_{t_j,s-o}^*],
\end{aligned}$$

(compare the first set of equalities in the proof of Theorem A.1(a)), where $\hat{\psi}_{l,m,n,o,p} = \hat{\psi}_{l,p} \hat{\psi}_{m,p} \hat{\psi}_{n,p} \hat{\psi}_{o,p} - \psi_l \psi_m \psi_n \psi_o$. By using the decomposition

$$\begin{aligned}
\hat{\psi}_{l,m,n,o,p} & = (\hat{\psi}_{l,p} - \psi_l) \hat{\psi}_{m,p} \hat{\psi}_{n,p} \hat{\psi}_{o,p} + \psi_l (\hat{\psi}_{m,p} - \psi_m) \hat{\psi}_{n,p} \hat{\psi}_{o,p} \\
& \quad + \psi_l \psi_m (\hat{\psi}_{n,p} - \psi_n) \hat{\psi}_{o,p} + \psi_l \psi_m \psi_n (\hat{\psi}_{o,p} - \psi_o)
\end{aligned}$$

this expression splits into four terms and for the sake of brevity we only consider the first one, other cases being treated similarly. As in the proof of Theorem A.1(a), we obtain terms $E_{N,T}^{1,*}$ and $E_{N,T}^{2,*}$ which are defined as $E_{N,T}^1$, $E_{N,T}^2$ where the coefficients $\psi_l(u_j) \psi_m(u_j) \psi_n(u_j) \psi_o(u_j)$ are replaced by $(\hat{\psi}_{l,p} - \psi_l) \hat{\psi}_{m,p} \hat{\psi}_{n,p} \hat{\psi}_{o,p}$ (note that $A_{N,T}^* = B_{N,T}^* = 0$ since the coefficients of the bootstrap process do not possess any time dependence). If we employ (A.3) and combine it with the fact that $|\hat{d} - \underline{d}| < \alpha/4$ on the set $A_T(\alpha)$, we get

$$|\hat{\psi}_{l,p} - \psi_l| \leq Cp^4 \log(T)^{3/2} T^{-1/2} |l|^{\alpha/4 + \underline{d} - 1} \quad \forall l \in \mathbb{N}, \quad (\text{A.26})$$

which, together with (2.3) and the assumptions of the theorem, implies

$$|\hat{\psi}_{l,p}| \leq C |l|^{\alpha/4 + \underline{d} - 1} \quad \forall l, p \in \mathbb{N}. \quad (\text{A.27})$$

The coefficients in the MA(∞) representation of the bootstrap processes do not depend on time and that for such processes we only require

$$\sup_u |\psi_l(u)| \leq C |l|^{\underline{d} - 1} \quad (\text{A.28})$$

in the proof of Theorem A.1(a) to obtain appropriate bounds for the error. By using (A.26) and (A.27), instead of (A.28) and similar arguments as given in the proof of Theorem A.1(a), it follows that

$$E((\hat{F}_{1,T}^* - \hat{F}_{1,T,2}^*)1_{A_T(\alpha)} | X_{1,T}, \dots, X_{T,T}) = F_{1,T}^{*,\bar{-}} + O(N^{4\underline{d}-1+\alpha} p^4 \log(T)^{3/2} T^{-1/2}),$$

where $F_{1,T}^{*,\bar{-}}$ is defined as $F_{1,T}$ but with $f(u, \lambda)$ replaced by

$$\frac{\sigma_p^2}{2\pi} \sum_{l,m,n,o=-\infty}^{\infty} \hat{\psi}_{l,m,n,o,p} \exp(-i\lambda(l - m + n - o)) \times 1_{A_T(\alpha)}.$$

Since $\hat{F}_{2,T}^*$ and $\hat{F}_{2,T,2}^*$ are treated analogously, the claim follows (note that $F_{1,T}^{*,\bar{-}}$ cancels out since the coefficients do not possess any time dependence).

Proof of part (b). The assertion follows by similar arguments as given in the proof of Theorem A.1(b) employing (A.26) and (A.27) instead of (A.28) as above. The details are omitted.

A.4. Proofs of the results in Section 3 and 4.

Proof of Theorem 2. The claim follows by employing the Cramér-Wold device in combination with Theorem A.1.

Proof of Theorem 3. Similarly to the proof of Theorem A.1, the equations

$$E(\hat{\tau}_1^2) = \frac{1}{\pi M} \sum_{j=1}^M \int_{-\pi}^{\pi} f^4(u_j, \lambda) + O\left(\frac{1}{N^{1-8d_\infty}}\right),$$

$$\text{Var}(\hat{\tau}_1^2) = O\left(\frac{1}{MN^{1-8d_\infty}}\right)$$

can be established. By Markov’s inequality the assertion of the theorem follows.

Proof of Theorem 4. We define $\hat{D}_{T,a}^{2,*}$ as $\hat{D}_{T,2}^{2,*}$ and $\hat{D}_{T,a}^2$ as \hat{D}_T^2 , but with $X_{t,T}$ replaced by $X_t(t/T)$ from (2.9). Then part a) is obvious, because we have $\psi_l = \psi_l(u)$ for all $u \in [0, 1]$ under the null hypothesis, and Z_t and Z_t^* are both independent and standard normal. Part b) follows from the proof of Theorem A.1, so we focus on part (c) and (d). Here (2.10) and Theorem A.1 (a), (b) imply

$$\frac{C_1 N^{\max(8\underline{d}-1,0)}}{T} \leq \frac{\text{Var}(\hat{D}_{T,a}^*)}{T} \leq \frac{C_2(N^{\max(8\underline{d}-1,0)} + \log(N)1_{\{\underline{d}=1/8\}})}{T}$$

which directly yields part (d). If we have

$$P(A_T(\alpha)) \rightarrow 1 \quad \text{as } T \rightarrow \infty \tag{A.29}$$

for every $\alpha > 0$, Part (c) follows from Theorem A.2, (4.7), the conditions on the rate of $p(T)$ if α is chosen sufficiently small. Finally, (A.29) is a consequence of Lemma 4.3 in Preuß and Vetter (2013).

Proof of Theorem 5. By employing the triangle inequality we can bound the Mallows metric between $\hat{D}_T^2/\sqrt{\text{Var}(\hat{D}_T^2)}$ and $\hat{D}_T^{2,*}/\sqrt{\text{Var}(\hat{D}_T^{2,*})}$ by

$$d_2\left(\frac{\hat{D}_T^2}{\sqrt{\text{Var}(\hat{D}_T^2)}, \frac{\hat{D}_{T,a}^2}{\sqrt{\text{Var}(\hat{D}_{T,a}^2)}}\right) + d_2\left(\frac{\hat{D}_{T,a}^2}{\sqrt{\text{Var}(\hat{D}_{T,a}^2)}, \frac{\hat{D}_{T,a}^{2,*}}{\sqrt{\text{Var}(\hat{D}_{T,a}^{2,*})}}\right) + d_2\left(\frac{\hat{D}_{T,a}^{2,*}}{\sqrt{\text{Var}(\hat{D}_{T,a}^{2,*})}}, \frac{\hat{D}_T^{2,*}}{\sqrt{\text{Var}(\hat{D}_T^{2,*})}}\right),$$

where $\hat{D}_{T,a}^2$ and $\hat{D}_{T,a}^{2,*}$ are the random variables in the proof of Theorem 4. It follows from the proof of Theorem A.1 that the first summand converges to zero and the second summand is zero because of Theorem 4(a). So it suffices to treat the third summand which is bounded by

$$2E\left(\frac{\hat{D}_{T,a}^{2,*}}{\sqrt{\text{Var}(\hat{D}_{T,a}^{2,*})}} - \frac{\hat{D}_T^{2,*}}{\sqrt{\text{Var}(\hat{D}_T^{2,*})}}\right)^2 + 2E\left(\frac{\hat{D}_T^{2,*}}{\sqrt{\text{Var}(\hat{D}_T^{2,*})}} - \frac{\hat{D}_{T,a}^{2,*}}{\sqrt{\text{Var}(\hat{D}_{T,a}^{2,*})}}\right)^2.$$

We obtain from Theorem A.1 (a) and (b) that a constant $L > 0$ exists such that

$$\text{Var}(\hat{D}_{T,a}^{2,*}) \geq LN^{8d_\infty-1}T^{-1} \tag{A.30}$$

(We are under the null hypothesis go $d_\infty = \underline{d}$). This combined with Theorem A.2, (A.29) and the conditions on the growth rate on $p = p(T)$ yields that we can restrict ourselves to the second term, which is (up to the constant 2) bounded by

$$\frac{E((\hat{D}_T^{2,*})^2)}{\text{Var}(\hat{D}_{T,a}^{2,*})\text{Var}(\hat{D}_T^{2,*})} \left(\sqrt{\text{Var}(\hat{D}_{T,a}^{2,*})} - \sqrt{\text{Var}(\hat{D}_T^{2,*})}\right)^2 \leq \frac{E((\hat{D}_T^{2,*})^2)}{\text{Var}(\hat{D}_{T,a}^{2,*})\text{Var}(\hat{D}_T^{2,*})} \left|\text{Var}(\hat{D}_{T,a}^{2,*}) - \text{Var}(\hat{D}_T^{2,*})\right|.$$

If we follow the proof of Theorem A.1(a), (b) and employ (A.27) and (A.29), we obtain that

$$E((\hat{D}_T^{2,*})^2) = O(\log(N)N^{\max(8\underline{d}-1,0)+2\alpha}T^{-1} + N^{8\underline{d}+2\alpha-2}) \tag{A.31}$$

holds for every fixed $\alpha > 0$. By employing (A.26) and similar arguments as in the proof of Theorem A.2 we obtain thereafter

$$\begin{aligned} & \left| \text{Var}(\hat{D}_{T,a}^{2,*} 1_{A_T(\alpha)}) - \text{Var}(\hat{D}_T^{2,*} 1_{A_T(\alpha)}) \right| \\ &= O(p^8 \log(T)^3 \log(N)^2 (N^{\max(8d-1,0)+2\alpha} T^{-2} + N^{8d-2+2\alpha} T^{-1})). \end{aligned}$$

The assertion then follows with (A.29)–(A.31) and the assumptions on the growth rate of $p = p(T)$.

Appendix B: Auxiliary Lemmas

Finally we show some lemmas that have been employed here.

Lemma B.1. Suppose $\mu, \nu, a, b \in \mathbb{R}$. Then there exists a constant $C \in \mathbb{R}$ such that the following hold:

(a) If $\mu, \nu > 0$ and $b > a$, then

$$\sum_{\substack{k=0 \\ k-a \geq 1 \\ -k+b \geq 1}}^{N-1} \frac{1}{(k-a)^{1-\mu}} \frac{1}{(b-k)^{1-\nu}} \leq \sum_{k=1+a}^{b-1} \frac{1}{(k-a)^{1-\mu}} \frac{1}{(b-k)^{1-\nu}} \leq \frac{C}{(b-a)^{1-\mu-\nu}}. \tag{B.1}$$

(b) If $0 < \mu, \nu$ and $0 < 1 - \mu - \nu$, then it follows for $|a + b| > 0$

$$\sum_{\substack{k=1 \\ k+b \geq 1 \\ k-a \geq 1}}^{N-1} \frac{1}{(k+b)^{1-\mu}} \frac{1}{(k-a)^{1-\nu}} \leq \sum_{\substack{k=1 \\ k+b \geq 1 \\ k-a \geq 1}}^{\infty} \frac{1}{(k+b)^{1-\mu}} \frac{1}{(k-a)^{1-\nu}} \leq \frac{C}{|a+b|^{1-\mu-\nu}}. \tag{B.2}$$

(c) If $0 < \nu < 1 - \mu$ and $y, z \geq 1$, then

$$\sum_{k=1+y}^{\infty} \frac{\log(k)}{k^{1-\mu}} \frac{1}{(k-y)^{1-\nu}} \leq C \frac{\log(y)}{y^{1-\mu-\nu}}, \tag{B.3}$$

$$\sum_{k=1}^{\infty} \frac{\log(k+z)}{(k+z)^{1-\mu}} \frac{1}{k^{1-\nu}} \leq C \frac{\log(z)}{z^{1-\mu-\nu}}. \tag{B.4}$$

Proof. (a) Using 3.196(3) in Gradshteyn and Ryzhik (1980), it follows that

$$\sum_{k=1+a}^{b-1} \frac{1}{(k-a)^{1-\mu}} \frac{1}{(b-k)^{1-\nu}} \leq \int_a^b \frac{1}{(x-a)^{1-\mu}} \frac{1}{(b-x)^{1-\nu}} dx \lesssim \frac{1}{(b-a)^{1-\mu-\nu}}.$$

(b) If $a + b > 0$ we can bound the middle sum in (B.2) by

$$\sum_{k=\max\{1, 1-b, 1+a\}}^{\infty} \frac{1}{(k+b)^{1-\mu}} \frac{1}{(k-a)^{1-\nu}} \leq \sum_{k=1+a}^{\infty} \frac{1}{(k+b)^{1-\mu}} \frac{1}{(k-a)^{1-\nu}}$$

$$\begin{aligned} &\leq \int_a^\infty \frac{1}{(x+b)^{1-\mu}} \frac{1}{(x-a)^{1-\nu}} dx \\ &\lesssim \frac{1}{(a+b)^{1-\mu-\nu}}. \end{aligned}$$

The last inequality follows from the equations 3.196(2) and 3.191(2) (for choosing $b = 0$) in Gradshteyn and Ryzhik (1980). On the other hand, if $a + b < 0$ we have

$$\begin{aligned} &\sum_{k=\max\{1,1-b,1+a\}}^\infty \frac{1}{(k+b)^{1-\mu}} \frac{1}{(k-a)^{1-\nu}} \leq \sum_{k=1-b}^\infty \frac{1}{(k+b)^{1-\mu}} \frac{1}{(k-a)^{1-\nu}} \\ &= \sum_{k=1+(-b)}^\infty \frac{1}{(k-(-b))^{1-\mu}} \frac{1}{(k+(-a))^{1-\nu}} \lesssim \frac{1}{(-a-b)^{1-\mu-\nu}}. \end{aligned}$$

The last inequality follows with Gradshteyn and Ryzhik (1980) as above.

(c) We start with (B.3). Using 13.2(18) in Erdélyi (1954b) yields

$$\sum_{k=1+y}^\infty \frac{\log(k)}{k^{1-\mu}} \frac{1}{(k-y)^{1-\nu}} \leq \int_y^\infty \frac{\log(x)}{x^{1-\mu}} \frac{1}{(x-y)^{1-\nu}} dx \lesssim \frac{\log(y)}{y^{1-\mu-\nu}}.$$

Concerning (B.4) we use 6.4(23) in Erdélyi (1954a) that implies

$$\sum_{k=1}^\infty \frac{\log(k+z)}{(k+z)^{1-\mu}} \frac{1}{k^{1-\nu}} \leq \int_0^\infty \frac{\log(x+z)}{(x+z)^{1-\mu}} \frac{1}{x^{1-\nu}} dx \lesssim \frac{\log(z)}{z^{1-\mu-\nu}}.$$

Lemma B.2. If $0 < d_\infty < 1/4$, then

$$\begin{aligned} &\frac{1}{M^2 N^4} \sum_{\substack{j_1, j_2=1 \\ j_1 \neq j_2}}^M \sum_{r_1, s_1=0}^{N-1} \sum_{\substack{v_1, w_1, x_1, y_1=0 \\ v_1-w_1+r_1-s_1=0 \\ 0 \leq |v_1-w_1| \leq N-1}}^\infty \sum_{\substack{v_2, w_2, x_2, y_2=0 \\ r_1+v_2-x_1-s_1-w_2+y_1-y_2+x_2=0 \\ 0 \leq r_1+v_2-x_1+(j_1-j_2)N \leq N-1 \\ 0 \leq s_1+w_2-y_1+(j_1-j_2)N \leq N-1 \\ 0 \leq |y_2-x_2| \leq N-1}}^\infty \\ &\psi_{v_1}(u_{j_1}) \psi_{w_1}(u_{j_1}) \psi_{x_1}(u_{j_1}) \psi_{y_1}(u_{j_1}) \psi_{v_2}(u_{j_2}) \psi_{w_2}(u_{j_2}) \psi_{x_2}(u_{j_2}) \psi_{y_2}(u_{j_2}) \\ &(N^2 - N|y_2 - x_2| - N|v_1 - w_1| + |v_1 - w_1|y_2 - x_2|) = O\left(\frac{\log(N)}{N^{1-8d_\infty T}}\right). \end{aligned}$$

Proof. We set $0 \leq w_1 = r_1 - s_1 + v_1$ and $0 \leq x_2 = s_1 - r_1 - v_2 + x_1 + w_2 - y_1 + y_2$. Then we define $p := r_1 + v_2 - x_1 + (j_1 - j_2)N$ and rearrange to $0 \leq x_1 = r_1 - p + v_2 + (j_1 - j_2)N$. Since $p \in \{0, 1, 2, \dots, N-1\}$, it follows that if p, r_1, v_2, x_1, j_1 are fixed, there are at most two possible values for j_2 . Hence it is enough to

consider the following expression with $1 \leq |C_1| \leq M - 1$

$$\begin{aligned}
 & \frac{1}{MN^4} \sum_{r_1, s_1=0}^{N-1} \sum_{\substack{v_1, w_1, x_1, y_1=0 \\ v_1-w_1+r_1-s_1=0 \\ 0 \leq |v_1-w_1| \leq N-1}}^{\infty} \sum_{\substack{v_2, w_2, x_2, y_2=0 \\ r_1+v_2-x_1-s_1-w_2+y_1-y_2+x_2=0 \\ 0 \leq r_1+v_2-x_1+C_1N \leq N-1 \\ 0 \leq s_1+w_2-y_1+C_1N \leq N-1 \\ 0 \leq |y_2-x_2| \leq N-1}}^{\infty} \\
 & \psi_{v_1}(u_{j_1}) \psi_{w_1}(u_{j_1}) \psi_{x_1}(u_{j_1}) \psi_{y_1}(u_{j_1}) \psi_{v_2}(u_{j_2}) \psi_{w_2}(u_{j_2}) \psi_{x_2}(u_{j_2}) \psi_{y_2}(u_{j_2}) \\
 & (N^2 - N|y_2 - x_2| - N|v_1 - w_1| + |v_1 - w_1||y_2 - x_2|) \\
 (2.3) \quad & \lesssim \frac{1}{MN^2} \sum_{p, r_1, s_1=0}^{N-1} \sum_{\substack{v_1, x_1, y_1=1 \\ 1 \leq r_1-s_1+v_1}}^{\infty} \sum_{\substack{v_2, w_2, x_2, y_2=1 \\ 1 \leq r_1-p+v_2+C_1N \\ 1 \leq s_1-p+w_2+y_2-y_1+C_1N}}^{\infty} \frac{1}{v_1^{1-d_\infty}} \frac{1}{y_1^{1-d_\infty}} \frac{1}{v_2^{1-d_\infty}} \\
 & \frac{1}{w_2^{1-d_\infty}} \frac{1}{y_2^{1-d_\infty}} \frac{1}{(r_1 - s_1 + v_1)^{1-d_\infty}} \frac{1}{(r_1 - p + v_2 + C_1N)^{1-d_\infty}} \\
 & \frac{1}{(s_1 - p + w_2 + y_2 - y_1 + C_1N)^{1-d_\infty}} \\
 (B.1) \quad & \lesssim \frac{1}{MN^2} \sum_{p, r_1=0}^{N-1} \sum_{v_1, x_1, y_1=1}^{\infty} \sum_{\substack{v_2, w_2, x_2, y_2=1 \\ 1 \leq r_1-p+v_2+C_1N \\ 1 \leq r_1-p+w_2-y_1+y_2+v_1+C_1N}}^{\infty} \frac{1}{v_1^{1-d_\infty}} \frac{1}{y_1^{1-d_\infty}} \\
 & \frac{1}{w_2^{1-d_\infty}} \frac{1}{y_2^{1-d_\infty}} \frac{1}{(r_1 - p + v_2 + C_1N)^{1-d_\infty}} \\
 & \frac{1}{(r_1 - p + w_2 - y_1 + y_2 + v_1 + C_1N)^{1-2d_\infty}} \\
 (B.1), (B.2) \quad & \lesssim \frac{1}{MN^2} \sum_{p, r_1=0}^{N-1} \frac{1}{|r_1 - p + C_1N|^{2-8d_\infty}} \lesssim \frac{1}{MN^2} \sum_{p, r_1=0}^{N-1} \frac{1}{(r_1 - p + N)^{2-8d_\infty}} \\
 & \lesssim \frac{\log(N)}{MN^{2-8d_\infty}} = \frac{\log(N)}{N^{1-8d_\infty} T}.
 \end{aligned}$$

The factor $\log(N)$ is due to the possible case $D = 0$.

References

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. *J. Statist. Plann. Inference* **139**, 900-915.

Beran, J., Feng, Y., Ghosh, S. and Kulik, R. (2013). *Long-Memory Processes: Probabilistic Properties and Statistical Methods*. Springer, Heidelberg.

- Berg, A., Paparoditis, E. and Politis, D. N. (2010). A bootstrap test for time series linearity. *J. Statist. Plann. Inference* **140**, 3481-3857.
- Berkes, I., Horváth, L., Kokoszka, P. and Shao, Q. M. (2006). On discriminating between long-range dependence and changes in mean. *Ann. Statist.* **34**, 1140-1165.
- Bickel, P. and Freedman, D. (1981). Some asymptotic theory for bootstrap. *Ann. Statist.* **9**, 1196-1217.
- Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory*. McGraw Hill, New York.
- Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25**, 1-37.
- Dahlhaus, R. (2000). A likelihood approximation for locally stationary processes. *Ann. Statist.* **28**, 1762-1794.
- Dahlhaus, R. (2009). Local inference for locally stationary time series based on the empirical spectral measure. *J. Econometrics* **151**, 101-112.
- Dahlhaus, R. and Polonik, W. (2006). Nonparametric quasi maximum likelihood estimation for Gaussian locally stationary processes. *Ann. Statist.* **34**, 2790-2824.
- Dehling, H., Rooch, A. and Taqqu, M. S. (2013). Nonparametric change-point tests for long-range dependent data. *Scand. J. Statist.* **40**, 153-173.
- Deo, R. S. and Chen, W. W. (2000). On the integral of the squared periodogram. *Stochastic Process. Appl.* **85**, 159-176.
- Dette, H., Preuß, P. and Vetter, M. (2011a). A measure of stationarity in locally stationary processes with applications to testing. *J. Amer. Statist. Assoc.* **106**, 1113-1124.
- Dette, H., Preuß, P. and Vetter, M. (2011b). Supplemental material to "A measure of stationarity in locally stationary processes with applications to testing". <http://dx.doi.org/10.1198/jasa.2011.tm10811>, Ruhr-Universität Bochum.
- Dwivedi, Y. and Subba Rao, S. (2010). A test for second order stationarity of a time series based on the Discrete Fourier Transform. *J. Time Series Anal.* **32**, 68-91.
- Erdélyi, A. (1954a). *Tables of Integral Transforms, Vol.1*. McGraw Hill, New York.
- Erdélyi, A. (1954b). *Tables of Integral Transforms, Vol.2*. McGraw Hill, New York.
- Eubank, R. L. and LaRiccia, V. N. (1992). Asymptotic comparison of Cramer-von Mises and nonparametric function estimation techniques for testing goodness-of-fit. *Ann. Statist.* **20**, 2071-2086.
- Fan, J. (1996). Test of significance based on wavelet thresholding and Neyman's truncation. *J. Amer. Statist. Assoc.* **91**, 674-688.
- Fox, R. and Taqqu, M. S. (1987). Central limit theorems for quadratic forms in random variables having long-range dependence. *Probab. Theory Related Fields* **74**, 213-240.
- Ghosh, B. K. and Huang, W.-M. (1991). The power and optimal kernel of the Bickel-Rosenblatt test for goodness of fit. *Ann. Statist.* **19**, 999-1009.
- Gradshteyn, I. S. and Ryzhik, I. M. (1980). *Table of Integrals, Series, and Products*. Academic Press, New York.
- Jentsch, C. and Subba Rao, S. (2015). A test for second order stationarity of a multivariate time series. *J. Econom.* **185**, 124-161.
- Jin, L., Wang, S. and Wang, H. (2015). A new nonparametric stationarity test of time series in time domain. *J. Roy. Statist. Soc. Ser. B* **77**, 893-922.
- Kreiß, J.-P., Paparoditis, E. and Politis, D. N. (2011). On the range of the validity of the autoregressive sieve bootstrap. *Ann. Statist.* **39**, 1827-2242.
- Lahiri, S. N. (2003). A necessary and sufficient condition for asymptotic independence of discrete Fourier transforms under short-and long-range dependence. *Ann. Statist.* **31**, 613-641.

- Nason, G. (2013). A test for second-order stationarity and approximate confidence intervals for localized autocovariances for locally stationary time series. *J. Roy. Statist. Soc. Ser. B* **75**, 879-904.
- Nason, G. P., von Sachs, R. and Kroisandt, G. (2000). Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. *J. Roy. Statist. Soc. Ser. B* **62**, 271-292.
- Neumann, M. H. and von Sachs, R. (1997) Wavelet thresholding in anisotropic function classes and applications to adaptive estimation of evolutionary spectra. *Ann. Statist.* **25**, 38-76.
- Palma, W. (2007). *Long-Memory Time Series: Theory and Methods*. Wiley Series in Probability and Statistics.
- Palma, W. (2010) On the sample mean of locally stationary long-memory processes. *J. Statist. Plann. Inference* **140**, 3764-3774.
- Palma, W. and Olea, R. (2010) An efficient estimator for locally stationary Gaussian Long-Memory processes. *Ann. Statist.* **38**, 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. *J. Amer. Statist. Assoc.* **105**, 839-851.
- Preuß, P. and Vetter, M. (2013) On discriminating between long-range dependence and non stationarity. *Electronic J. Statist.* **7**, 2241-2297.
- Preuß, P., Vetter, M. and Dette, H. (2013) A test for stationarity based on empirical processes. *Bernoulli* **19**, 2715-2749.
- Priestley, M. B. and Subba Rao, T. (1968). A test for non-stationary of time series. *J. Roy. Statist. Soc. Ser. B* **31**, 140-149.
- Priestley, M. B. (1965). Evolutionary spectra and non-stationary processes. *J. Roy. Statist. Soc. Ser. B* **62**, 204-237.
- Roueff, F. and von Sachs, R. (2011). Locally stationary long memory estimation. *Stochastic Process. Appl.* **121**, 813-844.
- Sibbertsen, P. and Kruse, R. (2009). Testing for a change in persistence under long-range dependencies. *J. Time Series Anal.* **30**, 263-285.
- Taniguchi, M. (1980). On estimation of the integrals of certain functions of spectral density. *J. Appl. Probab.* **17**, 73-83.
- von Sachs, R. and Neumann, M. H.. (2000). A wavelet-based test for stationarity. *J. Time Series Anal.* **21**, 597-613.
- Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis*. Almqvist and Wiksell, Uppsala.
- Wu, W. and Zhou, Z. (2014). Simultaneous quantile inference for non-stationary long memory time series. Preprint, University of Toronto.
- Zhou, Z. and Wu, W. B. (2009). Local linear quantile estimation for non-stationary time series. *Ann. Statist.* **37**, 2696-2729.
- Zhou, Z. and Wu, W. B. (2010). Simultaneous inference of linear models with time varying coefficients. *J. Roy. Statist. Soc. Ser. B* **72**, 513-531.

Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany.
E-mail: kemal.sen@ruhr-uni-bochum.de
Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany.
E-mail: philip.preuss@ruhr-uni-bochum.de
Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany.
E-mail: holger.dette@ruhr-uni-bochum.de

(Received February 2014; accepted February 2015)