Measuring stationarity in long-memory processes

Kemal Sen, Philip Preuß, Holger Dette

Ruhr-Universität Bochum

Fakultät für Mathematik

44780 Bochum

Germany

email: kemal.sen@ruhr-uni-bochum.de

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

email: holger.dette@ruhr-uni-bochum.de

March 14, 2013

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 will 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.

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, under the assumption of stationarity many real world phenomena can only be described by complicated and less intuitive models. A typical example
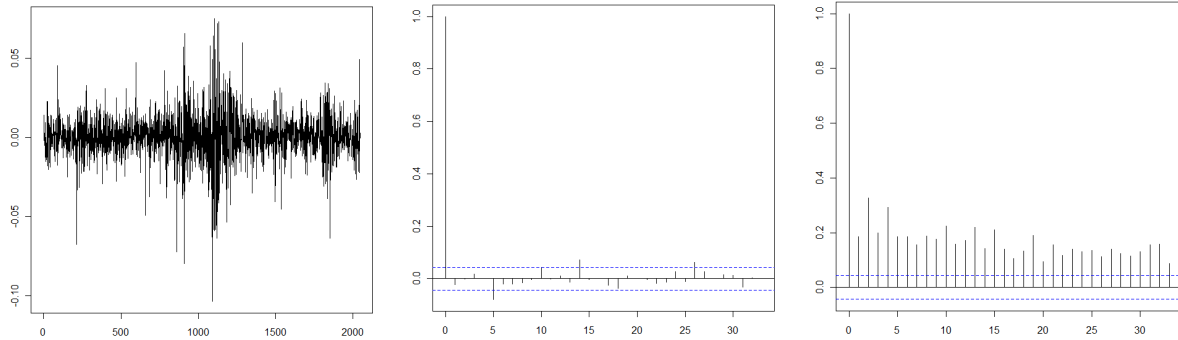
Figure 1: *Left panel: log-returns of the IBM stock between June 9th 2004 and July 24th 2012, middle panel: ACF of the log-returns X_t, right panel: ACF of the squared returns X_t^2*

can be found in the left panel of Figure 1 which shows 2048 log-returns of the IBM stock between June 9th 2004 and July 24th 2012. We observe that the autocovariance function (ACF) $\gamma(k) = \text{Cov}(X_0, X_k)$ of the log-returns X_t is converging to zero very "fast" as $k \to \infty$, while this is not the case for the ACF of the squared returns X_t^2 [see the middle and right panel in Figure 1]. The latter effect serves as the usual motivation to employ stationary long-memory models in the analysis of stock volatilities [see Breidt et al. (1998)]. This means that stationary processes satisfy

$$\gamma(k) \sim Ck^{2d-1}, \quad k \to \infty \tag{1.1}$$

for some $d \in (0, 0.5)$, which is called the long-memory parameter. Examples which fit into this framework are the well-known FARIMA(p, d, q)-models introduced by Granger and Joyeux (1980) and Hosking (1981). However, these kinds of processes are not very intuitive and it was suggested by several authors that one should use simple but non-stationary "short-memory" models instead [see for example Mikosch and Starica (2004), Starica and Granger (2005), Fryzlewicz et al. (2006) or Chen et al. (2010) among others]. Therefore an important question of interest in this context is, if the data should be analyzed by a stationary long-range dependent model or by a non-stationary "short-memory" model.

In the present paper 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. Since the assumption of stationarity is crucial in the application of various statistical tools, there exist several procedures to validate this condition in the context of short-memory processes. A first test for stationarity in locally stationary processes [as introduced by Dahlhaus (1997)] was proposed by von Sachs and Neumann (2000) and is based on the estimation of wavelet coefficients by a localized version of the periodogram. Paparoditis (2009, 2010) suggested an L_2-distance between the estimated spectral densities under the assumptions of stationarity and of local stationarity, and Dwivedi and Subba Rao (2010) developed a Portmanteau-type test statistic to detect deviations from stationarity. Besides the choice of a window width for the localized periodogram, which is inherent in essentially any statistical inference for locally stationary processes, all these methods require the choice of at least one additional 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 et al. (2011a) developed tests for stationarity in the framework of locally stationary processes which only require the choice of one regularization parameter, namely the window length for the localized periodogram in the estimation procedure. These authors considered the L_2-distance

$$D^2 := \int_0^1 \int_{-\pi}^{\pi} \left(f(u, \lambda) - \int_0^1 f(v, \lambda) dv \right)^2 d\lambda du \tag{1.2}$$

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 et al. (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 et al. (2012) proposed an alternative measure for deviations from stationarity which is based on the Kolmogorov-Smirnov distance

$$D_{KS} := \sup_{(v, \omega) \in [0,1]^2} \frac{1}{\pi} \left| \int_0^v \int_{-\pi\omega}^{\pi\omega} f(u, \lambda) d\lambda du - v \int_{-\pi\omega}^{\pi\omega} \int_0^1 f(u, \lambda) du d\lambda \right| \tag{1.3}$$

[see also Dahlhaus (2009)]. Both approaches have their pros and cons. In particular tests based on the distance (1.3) 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 et al. (2011a) and Preuß et al. (2012) 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 employed numerously 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 et al. (2011) consider CUSUM and Wilcoxon type tests to discriminate between long-range dependence and one change in mean. However 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 (2012) developed a test for stationarity which includes the long-range dependent case and is based on the distance (1.3). As mentioned in the previous paragraph there

exist several situations where this approach is not the best and for this reason we consider in this paper an alternative test which is based on the measure defined in (1.2). For this purpose, we estimate the integrals in the distance D^2 in (1.2) by Riemann sums where the unknown spectral densities are replaced by periodograms. For the resulting statistic we will 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.

Although that the proof of asymptotic normality seems to be more of theoretical nature, because the bootstrap procedure derived in the second part of the paper can in principle also be applied in the case $d(u) < 1/8$, these results are of interest from several perspectives. Firstly, several arguments used in the proof of asymptotic normality are also required in the proof of consistency of the bootstrap procedure and easier illustrated in the unconditional case. Secondly, and most important, 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 $\int_0^1 \int_{-\pi}^{\pi} f^2(u, \lambda)\, d\lambda\, du$, respectively. For this purpose we use Riemann sums of the squared periodogram instead of not 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 a 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 remaining part of this 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 $\hat{D}_T^2$ and prove its consistency. The finite sample properties are investigated in Section 5. Finally, 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 which is flexible enough to cover stationary long-memory processes and a reasonable time-varying extension of it as well. For this reason, we consider the following 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)].

Assumption 2.1. *Let $\{X_{t,T}\}_{t=1,\ldots,T}$ denote a sequence of stochastic processes which have a MA(∞) representation of the form*

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

such that

$$\sup_{T \in \mathbb{N}} \sup_{t \in \{1,\ldots,T\}} \sum_{l=0}^{\infty} \psi_{t,T,l}^2 < \infty, \tag{2.2}$$

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

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

$$\sup_{t=1,\ldots,T} \left| \psi_{t,T,l} - \psi_l(t/T) \right| \leq \frac{C}{TI(l)^{1-d_\infty}} \quad \forall l \in \mathbb{N} \tag{2.3}$$

and

$$\psi_l(u) = \frac{a(u)}{I(l)^{1-d(u)}} + O\left(\frac{1}{I(l)^{2-d_\infty}}\right) \tag{2.4}$$

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

2) *The time varying spectral density*

$$f(u, \lambda) := \frac{1}{2\pi} \left| \sum_{l=0}^{\infty} \psi_l(u) \exp(-i\lambda l) \right|^2 \tag{2.5}$$

is twice continuously differentiable 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 $C \in \mathbb{R}^+$, which is independent of u and λ, such that for $l \neq 0$*

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

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

In addition, we assume

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

$$\sup_{u \in (0,1)} \left| \frac{\partial^2}{\partial u^2} f(u, \lambda) \right| \leq \frac{C \log^2(\lambda)}{\lambda^{2d_\infty}}. \tag{2.9}$$

For the sake of a transparent notation, we will use $C \in \mathbb{R}^+$ as a universal constant throughout this paper. Note that the process is stationary if $\psi_{l,t,T} = \psi_l$ for all $l, t, T \in \mathbb{N}$. Condition (2.2) ensures that the infinite sum in (2.1) exists in the L_2 sense, and (2.3) means that the process $X_{t,T}$ can be approximated by a stationary model on a small time interval. It is also worthwhile to mention that the assumption of Gaussianity is only imposed here to simplify technical arguments [since they are quite involved even in this case]; see Remark 3.5 for more details. Next, we consider the process

$$X_t(u) := \sum_{l=0}^{\infty} \psi_l(u) Z_{t-l} \tag{2.10}$$

in order to visualize some properties of a locally stationary long-memory process. Firstly, $X_t(u)$ is stationary for every fixed $u \in [0, 1]$ and analogously to the stationary case, the condition (2.4) implies the existence of bounded functions $y_i : [0, 1] \to \mathbb{R}$ $(i = 1, 2)$ such that

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

and

$$f(u, \lambda) \sim \frac{y_2(u)}{\lambda^{2d(u)}} \quad \text{as } \lambda \to 0; \tag{2.11}$$

[see Palma and Olea (2010)]. Consequently, the autocovariance function $\gamma(u, k) = \mathrm{Cov}(X_0(u), X_k(u))$ 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 which fits in this theoretical set-up we consider the time varying FARIMA(p, d, q) model defined by the equation

$$a(t/T, B)(1 - B)^{d(t/T)} X_{t,T} = b(t/T, B) Z_t, \quad t = 1, ..., T, \tag{2.12}$$

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 $a_j, b_j : [0, 1] \to \mathbb{R}$, and $d : [0, 1] \to (0, d_\infty]$ is twice continuously differentiable with $d_\infty < 1/2$. It is shown in Preuß and Vetter (2012) that under certain regularity conditions on the functions a_j, b_j, these equations have a locally stationary solution in the sense of Assumption 2.1. If the functions a_j, b_j and d do not depend on u, (2.12) 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 (1.2) we note that

$$D^2 = 2\pi F_1 - 4\pi F_2, \tag{2.13}$$

where

$$F_1 := \frac{1}{2\pi} \int_0^1 \int_{-\pi}^{\pi} f^2(u, \lambda) d\lambda du, \tag{2.14}$$

6

$$F_2 := \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\int_0^1 f(u,\lambda)du \right)^2 d\lambda. \tag{2.15}$$

Consequently, it follows from (2.11) that the distance D^2 is only well defined if $d_\infty < \frac{1}{4}$. We assume without loss of generality that the sample size T can be decomposed into M blocks with length N (i.e. $T = NM$) 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 $u \in [0,1]$ which is defined by

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

where $X_{j,T} = 0$ for $j \notin \{1,\ldots,T\}$ [see Dahlhaus (1997)]. This is the usual periodogram computed from the observations $X_{\lfloor uT \rfloor - N/2 + 1, T}, \ldots, X_{\lfloor uT \rfloor + N/2, T}$, and it can be shown that it is asymptotically unbiased for the time-varying spectral density $f(u,\lambda)$ if $N \to \infty$ 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 now construct empirical versions of (2.14) and (2.15) 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, we define the rescaled mid-points of the M blocks

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

and consider the statistics

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

$$\hat{F}_{2,T} := \frac{1}{N} \sum_{k=1}^{\lfloor \frac{N}{2} \rfloor} \left(\frac{1}{M} \sum_{j=1}^{M} I_N(u_j, \lambda_{k,N}) \right)^2, \tag{2.17}$$

where $\lambda_{k,N} = 2\pi k/N$ denote the usual Fourier frequencies. The empirical measure of stationarity (1.2) is finally given by

$$\hat{D}_T^2 := 2\pi \hat{F}_{1,T} - 4\pi \hat{F}_{2,T}. \tag{2.18}$$

We would like to point out here that 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 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 well 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 will 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

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

and obtain the following results.

Theorem 3.1. *Suppose Assumption 2.1 holds with* $\sup_{u\in[0,1]} d(u) < 1/4$ *and that the conditions*

$$N \to \infty, \quad \frac{N}{T} \to 0 \tag{3.1}$$

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 $T \to \infty$.

Theorem 3.2. *Suppose Assumption 2.1 holds with* $d_\infty = \sup_{u\in[0,1]} d(u) < 1/8$ *and that the conditions*

$$N \to \infty, \quad \frac{N}{T} \to 0, \quad \frac{\sqrt{T}}{N^{1-4d_\infty}} \to 0 \tag{3.2}$$

are satisfied. Then as $T \to \infty$ *we have*

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

where the covariance matrix $\boldsymbol{\Sigma}$ *and the constant* $d_{N,T}$ *are given by*

$$\boldsymbol{\Sigma} = \begin{pmatrix} \frac{5}{\pi} \int_{-\pi}^{\pi} \int_0^1 f^4(u, \lambda) \, du \, d\lambda & \frac{2}{\pi} \int_{-\pi}^{\pi} \Big(\int_0^1 f(u, \lambda) du \int_0^1 f^3(u, \lambda) du\Big) d\lambda \\ \frac{2}{\pi} \int_{-\pi}^{\pi} \Big(\int_0^1 f(u, \lambda) du \int_0^1 f^3(u, \lambda) du\Big) d\lambda & \frac{1}{\pi} \int_{-\pi}^{\pi} \Big(\big(\int_0^1 f(u, \lambda) du\big)^2 \int_0^1 f^2(u, \lambda) du\Big) d\lambda \end{pmatrix} \tag{3.3}$$

$$d_{N,T} = \frac{1}{4\pi M^2} \sum_{j=1}^{M} \int_{-\pi}^{\pi} f^2(u_j, \lambda) \, d\lambda, \tag{3.4}$$

respectively, and the vector $\boldsymbol{C_T} \in \mathbb{R}^2$ *is of order* $O\big(N^2/T^2 + \log(N)/(MN^{1-4d_\infty})\big)$. *In particular, this term vanishes if the functions* $\psi_l(u)$ *are independent of* u *for all* $l \in \mathbb{Z}$ *[i.e. the spectral density* $f(u, \lambda)$ *of the underlying process* $X_{t,T}$ *is independent of* u].

A similar result for the short-memory situation has been derived by Dette et al. (2011a). In contrast to their result, there appears an additional bias term $\boldsymbol{C_T}$ in Theorem 3.2. This term is negligible under the additional condition $N^2/T^{3/2} \to 0$ which holds under the stronger restriction $d_\infty < 1/12$ due to (3.2). On the other hand, under the null hypothesis of a time independent spectral density

$$H_0 : f(u, \lambda) \text{ is independent of } u, \tag{3.5}$$

we have that $C_T = 0$ (this follows from the proof of Theorem 3.2 in the Appendix). Since the covariance matrix (3.3) contains the integrated fourth power of the spectral density, we obtain from (2.11) that Theorem 3.2 is not valid whenever $d_\infty \geq 1/8$. Writing $(C_{1,T}, C_{2,T})^T := \boldsymbol{C_T}$, a straightforward application of the Delta-method yields the following result.

Corollary 3.3. *Under the assumptions of Theorem 3.2, it holds*

$$\sqrt{T}\left(\hat{D}_T^2 - D_T^2 + 4\pi d_{N,T} + 4\pi C_{2,T} - 2\pi C_{1,T}\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \tau^2), \tag{3.6}$$

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

$$\tau^2 := 20\pi \int_{-\pi}^{\pi} \int_0^1 f^4(u,\lambda)dud\lambda - 32\pi \int_{-\pi}^{\pi} \left(\int_0^1 f(u,\lambda)du \int_0^1 f^3(u,\lambda)du\right)d\lambda \tag{3.7}$$

$$+ 16\pi \int_{-\pi}^{\pi} \left(\left(\int_0^1 f(u,\lambda)du\right)^2 \int_0^1 f^2(u,\lambda)du\right)d\lambda.$$

Under the null hypothesis (3.5) we have $D_T^2 = C_{1,T} = C_{2,T} = 0$ and the asymptotic variance in (3.7) reduces to $\tau_{H_0}^2 := 4\pi \int_{-\pi}^{\pi} f^4(\lambda)d\lambda$. The asymptotic bias $4\pi d_{N,T} = \frac{2\pi N}{T} F_{1,T}$ can easily be estimated by the statistic $\hat{B}_T := \frac{2\pi N}{T}\hat{F}_{1,T}$ and we infer from Theorem 3.2

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

Thus Slutzky's Lemma together with (3.6) yields

$$\sqrt{T}\left(\hat{D}_T^2 + \hat{B}_T\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \tau_{H_0}^2) \tag{3.8}$$

under the null hypothesis. To construct an asymptotic level α-test for stationarity, it therefore remains to estimate the variance $\tau_{H_0}^2$ in (3.8), and an estimator for this quantity is given by $\hat{\tau}_{H_0}^2 := 4\pi^2 \hat{\tau}_1^2$ with

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

The consistency of this estimator follows from the next theorem.

Theorem 3.4. *If the assumptions of Theorem 3.2 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 (3.8) with Theorem 3.4 yields

$$\sqrt{T}\left(\hat{D}_T^2 + \hat{B}_T\right)\bigg/\sqrt{\hat{\tau}_{H_0}^2} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \tag{3.9}$$

and therefore an asymptotic level α-test is obtained by rejecting the null hypothesis (3.5) whenever

$$\sqrt{T}\left(\hat{D}_T^2 + \hat{B}_T\right)\bigg/\sqrt{\hat{\tau}_{H_0}^2} \geq u_{1-\alpha}, \tag{3.10}$$

where $u_{1-\alpha}$ denotes the $(1-\alpha)$-quantile of the standard normal distribution. It follows from Theorem 3.2 that this test is consistent, because the left hand side of (3.10) 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 3.5. If the innovation process $(Z_t)_{t\in\mathbb{Z}}$ in (2.1) is not Gaussian, it can be shown that Corollary 3.3 is still valid where the asymptotic variance τ^2 in (3.7) has to be replaced by

$$\tau_g^2 = \tau^2 + \frac{\kappa_4}{\kappa_2^2}\Bigg\{4\int_0^1\Bigg(\int_{-\pi}^{\pi}f^2(u,\lambda)\,d\lambda\Bigg)^2 du + 4\int_0^1\Bigg(\int_{-\pi}^{\pi}f(u,\lambda)\Bigg(\int_0^1 f(\nu,\lambda)\,d\nu\Bigg)d\lambda\Bigg)^2 du$$

$$-8\int_0^1\Bigg(\int_{-\pi}^{\pi}f^2(u,\lambda)\,d\lambda\int_{-\pi}^{\pi}f(u,\lambda)\Bigg(\int_0^1 f(\nu,\lambda)\,d\nu\Bigg)d\lambda\Bigg)du\Bigg\}$$

and κ_2 and κ_4 denote the second and fourth cumulants of the innovation process, respectively. 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 (3.10) are necessary to address non normal distributed innovations.

Remark 3.6. We note that 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 (2.16), are different. In fact it follows by similar arguments as given in the appendix that

$$\lim_{T\to\infty}T\,\mathrm{Var}(\tilde{F}_{1,T}) = \frac{14}{3\pi}\int_{-\pi}^{\pi}\int_0^1 f^4(u,\lambda)\,du\,d\lambda,$$

while

$$\lim_{T\to\infty}T\,\mathrm{Var}(\hat{F}_{1,T}) = \frac{5}{\pi}\int_{-\pi}^{\pi}\int_0^1 f^4(u,\lambda)\,du\,d\lambda$$

by Theorem 3.2. Moreover, similar arguments as given in the proof of this statement 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 $d_\infty < \frac{1}{4}$ as specified in Assumption 2.1. We will show that in this case a bootstrap procedure can be used to approximate the distribution of $\hat{D}_T^2$ under the null hypothesis (3.5). We employ the FARI(∞) bootstrap which was recently introduced by Preuß and Vetter (2012) and fits an FARIMA($p,\underline{d},0$)-model to the data, where $p = p(T)$ converges to infinity with increasing sample size T. To prove consistency of this procedure, we require the following technical assumptions.

Assumption 4.1. *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*

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

has an AR(∞)-representation of the form

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

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

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

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

Note that 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($\infty$) bootstrap incorporates the following steps: First we 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). Secondly we calculate an estimator of

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

by fitting an AR(p)-model to the data. In order to describe the main idea of our procedure in more detail, we introduce the "true" approximating process $Y_t^{AR}(p)$ by

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

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 $\sigma_p^2 = \mathbb{E}(Y_t - \sum_{j=1}^{p} a_{j,p}Y_{t-j})^2$. If $p = p(T) \to \infty$ 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 $\{X_{t,T}\}_{t=1,...,T}$. Hence, observing the data $X_{1,T}, ..., X_{T,T}$, the FARI(∞) bootstrap precisely works as follows:

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

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

11

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

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

is the usual periodogram, and

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

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

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

$$Y_{t,T}^* = Y_{t,T}; \quad t = 1, ..., p,$$

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

where Z_j^* denotes an independent sequence of standard normal distributed random variables.

3) Create the pseudo-series $X_{1,T}^*, ..., X_{T,T}^*$ from the equation

$$X_{i,T}^* = (1 - B)^{-\hat{d}} Y_{i,T}^* \tag{4.7}$$

and compute $\hat{D}_T^{2,*}$ in the same way as $\hat{D}_T^2$ where the original observations $X_{1,T}, ..., X_{T,T}$ are replaced by the bootstrap replicates $X_{1,T}^*, ..., X_{T,T}^*$.

Our main theorem in this section describes the theoretical properties of this procedure.

Theorem 4.2. *Assume that the null hypothesis (3.5) holds and let Assumption 2.1 and 4.1 be fulfilled. Furthermore, suppose that the conditions*

$$N \to \infty, \quad \frac{N}{T} \to 0, \quad \frac{T}{N^{1+\delta}} \to 0 \tag{4.8}$$

are satisfied for some $0 < \delta < 1/2$, and assume for the growth rate (rate of convergence) of $p = p(T)$ the following:

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

$$p_{max}^9(T) \log(T)^3 N^\delta T^{-1} = O(1), \tag{4.9}$$
$$\sqrt{T} p_{min}^{-9}(T) / \sqrt{\log(T)} = o(1). \tag{4.10}$$

ii) *The condition*

$$||\hat{\theta}_{T,p} - \theta_p||_\infty = O_P\left(\sqrt{\frac{\log(T)}{T}} \right) \tag{4.11}$$

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

12

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

$$(a) \quad \hat{D}^2_{T,a} \overset{\mathcal{D}}{=} \hat{D}^{2,*}_{T,a},$$

$$(b) \quad \mathrm{Var}(\hat{D}^2_T)^{-1/2}\left(\hat{D}^2_T - \hat{D}^2_{T,a}\right) = o_P(1),$$

$$(c) \quad \mathrm{Var}(\hat{D}^{2,*}_T)^{-1/2}\left(\hat{D}^{2,*}_T - \hat{D}^{2,*}_{T,a}\right) = o_P(1),$$

$$(d) \quad \mathbb{E}|\hat{D}^{2,*}_{T,a}| = O\left(\frac{N^{\max(4\underline{d}-1/2,0)}}{\sqrt{T}} + \frac{\sqrt{\log(N)}1_{\{\underline{d}=1/8\}}}{\sqrt{T}} + \frac{1}{N^{1-4\underline{d}}}\right).$$

The estimate in (d) also holds if the null hypothesis (3.5) is not satisfied.

Note that conditions like (4.9)-(4.11) are standard in the context of parametric bootstraps [see for example Berg et al. (2010) or Kreiß et al. (2011)] and a detailed discussion of them is given in Preuß and Vetter (2012). We now obtain an asymptotic level α-test based on $\hat{D}^2_T$ as follows: Calculate B bootstrap replicates $\hat{D}^{2,*}_T$, denote by $(\hat{D}^{2,*}_T)_{T,1}, ..., (\hat{D}^{2,*}_T)_{T,B}$ the resulting order statistic and reject the null hypothesis whenever

$$\hat{D}^2_T > (\hat{D}^{2,*}_T)_{T,\lfloor(1-\alpha)B\rfloor}. \tag{4.12}$$

Theorem 4.2 and the argumentation in Paparoditis (2010) indicate that this procedure is valid for obtaining an asymptotic level α-test. In order to prove this more formally, we follow Bickel and Freedman (1981) by considering the Mallow metric $d_2(F,G) = \inf \sqrt{\mathbb{E}(X-Y)^2}$ 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 4.2 then yields the following result which states that the test (4.12) has, in fact, asymptotic level α.

Theorem 4.3. *Suppose the null hypothesis (3.5) and the assumptions of Theorem 4.2 are satisfied. Then, as $T \to \infty$, the Mallow distance d_2 between the distributions of the random variables*

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

converges to zero in probability.

Consistency under the alternative follows since Theorem 4.2 d) yields that each bootstrap statistic $\hat{D}^{2,*}_T$ converges to zero while $\hat{D}^2_T$ exceeds some positive constant (for T sufficiently large) due to Theorem 3.1.

5 Finite sample properties

In this section we examine the finite sample properties of the proposed decision rule (4.12). 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)], which is defined by

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

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 (4.12) using the FARI(∞) bootstrap exhibits a remarkable robustness with respect to the choice of N. All reported results of this section are based on 1000 simulation runs and 200 bootstrap replications.

5.1 Size of the test

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

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

and the FARIMA$(0, d, 1)$ process

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

for different values of ϕ, θ and d where the random variables Z_t are independent standard normal distributed. The rejection probabilities for the bootstrap test (4.12) are displayed in Table 1–4 where $d \in \{0.1, 0.2\}$. We observe a very precise approximation of the nominal level in nearly all cases which is rather robust with respect to different choices of the parameter M and N.

In order to study the power of the test we consider the following alternatives

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

$$(1 - 0.6 \sin(4\pi t/T)B)\,(1 - B)^d X_{t,T} = Z_t, \tag{5.4}$$

$$(1 - B)^d X_{t,T} = \sqrt{\sin(\pi t/T)} Z_t, \tag{5.5}$$

where $d = 0.2$. 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 (4.12) are depicted in Figure 2–4 for different combinations of T and N. Additionally, the results for the Kolmogorov-Smirnov approach of Preuß and Vetter (2012) are presented. We observe that the new procedure clearly outperforms the test of Preuß and Vetter (2012) for the models (5.3) and (5.4) while the Kolmogorov-Smirnov test works 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.

14

	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	.126	.182	.072	.123	.036	.074	.073	.126	.084	.147
A2	128	8	16	.140	.200	.085	.132	.041	.090	.073	.128	.084	.118
B1	256	32	8	.065	.135	.064	.119	.042	.088	.062	.113	.075	.147
B2	256	16	16	.080	.132	.056	.108	.040	.082	.051	.109	.062	.109
B3	256	8	32	.068	.111	.045	.095	.046	.097	.072	.147	.049	.114
C1	512	64	8	.054	.109	.049	.106	.039	.089	.049	.114	.082	.134
C2	512	32	16	.038	.093	.043	.086	.039	.085	.059	.108	.065	.132
C3	512	16	32	.061	.095	.051	.102	.045	.081	.059	.109	.043	.104
C4	512	8	64	.060	.107	.053	.098	.045	.083	.060	.116	.042	.093
D1	1024	128	8	.039	.104	.042	.093	.042	.085	.035	.093	.079	.132
D2	1024	64	16	.053	.104	.058	.097	.050	.110	.057	.101	.068	.126
D3	1024	32	32	.033	.076	.058	.114	.046	.086	.070	.107	.062	.115
D4	1024	16	64	.046	.089	.036	.091	.044	.084	.054	.109	.044	.099
D5	1024	8	128	.037	.073	.041	.091	.041	.091	.061	.131	.045	.097

Table 1: *Rejection probabilities of the bootstrap test* (4.12) *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	.107	.164	.063	.114	.050	.108	.072	.121	.108	.166
A2	128	8	16	.106	.160	.064	.118	.041	.085	.073	.124	.078	.138
B1	256	32	8	.064	.123	.048	.104	.042	.094	.075	.131	.079	.137
B2	256	16	16	.058	.125	.051	.101	.040	.101	.065	.112	.055	.116
B3	256	8	32	.079	.124	.047	.089	.051	.091	.053	.106	.050	.105
C1	512	64	8	.050	.093	.048	.090	.051	.103	.047	.104	.075	.133
C2	512	32	16	.047	.104	.044	.087	.039	.085	.053	.109	.068	.124
C3	512	16	32	.042	.097	.044	.087	.057	.106	.046	.105	.060	.104
C4	512	8	64	.050	.102	.053	.101	.052	.088	.058	.121	.062	.114
D1	1024	128	8	.044	.090	.046	.102	.051	.107	.039	.092	.076	.140
D2	1024	64	16	.043	.082	.040	.088	.050	.098	.046	.098	.060	.106
D3	1024	32	32	.045	.089	.054	.097	.057	.103	.060	.104	.066	.115
D4	1024	16	64	.044	.087	.038	.087	.049	.094	.059	.106	.051	.101
D5	1024	8	128	.041	.082	.041	.089	.038	.086	.061	.103	.054	.103

Table 2: *Rejection probabilities of the bootstrap test* (4.12) *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.

				$\theta = -0.9$		$\theta = -0.5$		$\theta = 0$		$\theta = 0.5$		$\theta = 0.9$	
	T	N	M	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
A1	128	16	8	.072	.116	.054	.107	.041	.085	.044	.088	.077	.123
A2	128	8	16	.068	.122	.054	.112	.059	.125	.073	.117	.070	.133
B1	256	32	8	.045	.100	.060	.101	.041	.081	.042	.082	.036	.084
B2	256	16	16	.053	.096	.058	.104	.045	.094	.045	.102	.060	.104
B3	256	8	32	.064	.123	.057	.113	.049	.101	.042	.092	.061	.130
C1	512	64	8	.043	.089	.043	.095	.044	.086	.045	.088	.041	.095
C2	512	32	16	.046	.109	.067	.112	.052	.093	.051	.096	.043	.086
C3	512	16	32	.048	.099	.055	.095	.062	.114	.050	.102	.051	.098
C4	512	8	64	.038	.097	.055	.100	.047	.100	.046	.093	.042	.092
D1	1024	128	8	.053	.103	.060	.099	.051	.099	.071	.118	.044	.094
D2	1024	64	16	.044	.100	.062	.124	.048	.090	.068	.119	.042	.093
D3	1024	32	32	.053	.107	.064	.116	.044	.082	.045	.094	.043	.098
D4	1024	16	64	.044	.096	.038	.084	.042	.093	.045	.087	.042	.087
D5	1024	8	128	.049	.109	.042	.085	.054	.109	.048	.083	.042	.096

Table 3: *Rejection probabilities of the bootstrap test* (4.12) *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.*

				$\theta = -0.9$		$\theta = -0.5$		$\theta = 0$		$\theta = 0.5$		$\theta = 0.9$	
	T	N	M	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
A1	128	16	8	.068	.112	.060	.103	.030	.081	.060	.108	.053	.111
A2	128	8	16	.060	.117	.051	.103	.061	.110	.062	.114	.068	.117
B1	256	32	8	.059	.122	.048	.102	.045	.094	.038	.078	.040	.083
B2	256	16	16	.053	.109	.041	.095	.047	.093	.040	.080	.048	.091
B3	256	8	32	.060	.100	.048	.098	.057	.119	.050	.092	.061	.106
C1	512	64	8	.059	.110	.064	.122	.052	.099	.056	.099	.055	.101
C2	512	32	16	.060	.122	.044	.107	.041	.103	.043	.113	.046	.086
C3	512	16	32	.061	.116	.056	.122	.049	.089	.046	.088	.052	.099
C4	512	8	64	.056	.095	.057	.118	.057	.110	.047	.100	.055	.102
D1	1024	128	8	.063	.125	.054	.102	.039	.086	.044	.101	.051	.098
D2	1024	64	16	.051	.109	.061	.112	.047	.107	.056	.106	.047	.100
D3	1024	32	32	.055	.092	.057	.111	.048	.095	.057	.106	.047	.119
D4	1024	16	64	.065	.124	.061	.116	.043	.092	.048	.087	.049	.098
D5	1024	8	128	.059	.116	.052	.095	.049	.093	.035	.075	.050	.115

Table 4: *Rejection probabilities of the bootstrap test* (4.12) *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.*
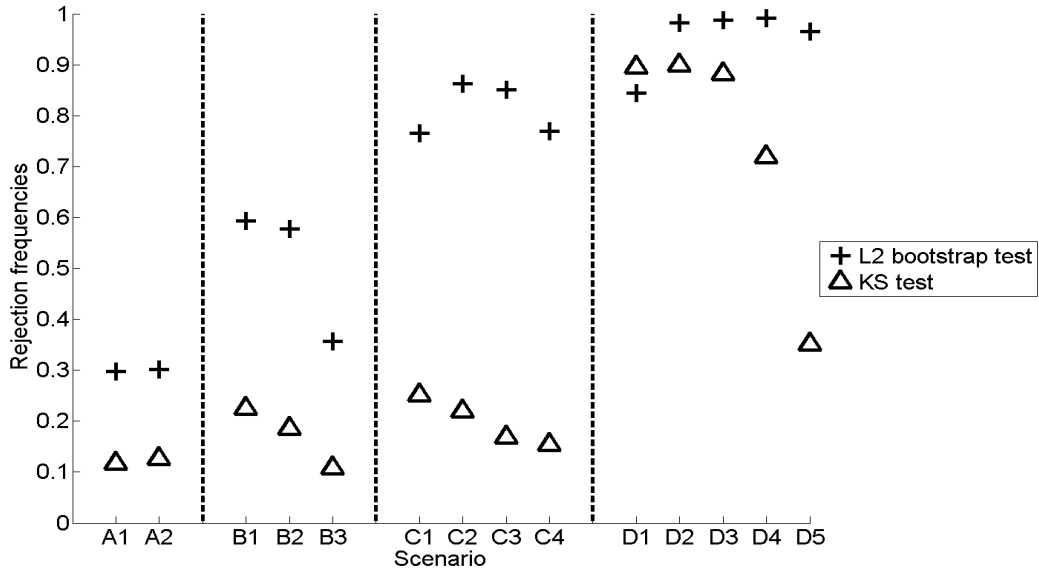
Figure 2: *Power of the test (4.12) and the Kolmogorov-Smirnov test of Preuß and Vetter (2012) at 5% level for the model (5.3) under the scenarios A1-D5 from Table 1.*
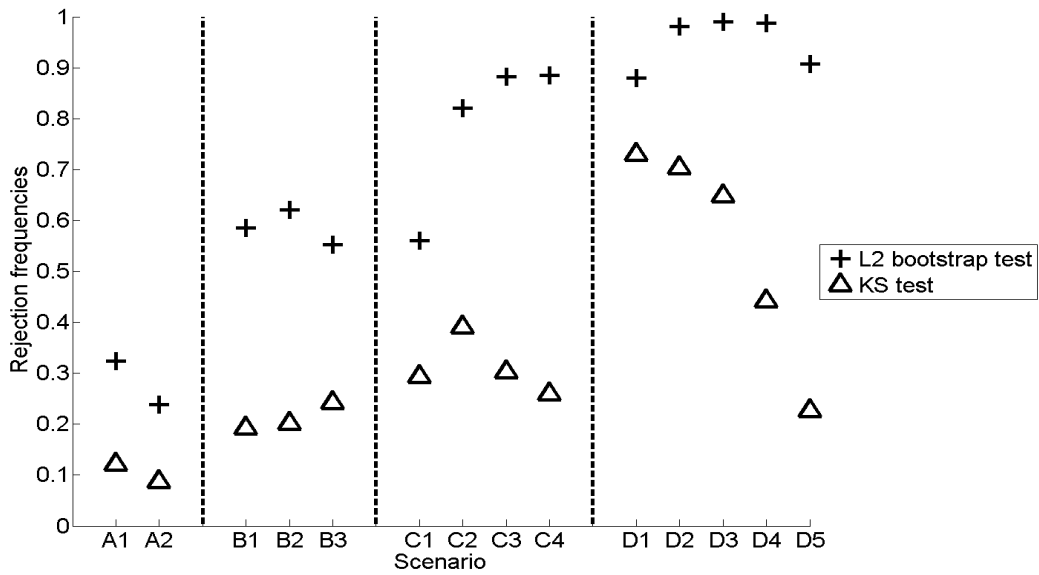

Figure 3: *Power of the test (4.12) and the Kolmogorov-Smirnov test of Preuß and Vetter (2012) at 5% level for the model (5.4) under the scenarios A1-D5 from Table 1.*
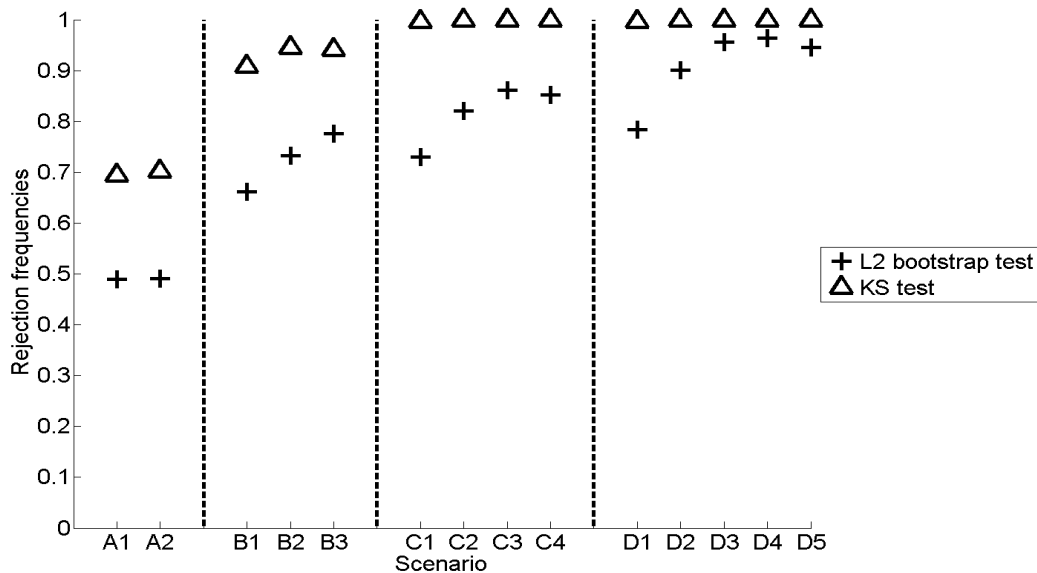
Figure 4: *Power of the test (4.12) and the Kolmogorov-Smirnov test of Preuß and Vetter (2012) at 5% level for the model (5.5) under the scenarios A1-D5 from Table 1.*

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. *Journal of Statistical Planning and Inference*, 139:900–915.

Berg, A., Paparoditis, E., and Politis, D. N. (2010). A bootstrap test for time series linearity. *Journal of Statistical Planning and Inference*, 140:3841–3857.

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

Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *Annals of Statistics*, 9:1196–1217.

Breidt, J., Crato, N., and De Lima, P. (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, 83:325–348.

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

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

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

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

Dehling, H., Rooch, A., and Taqqu, M. S. (2011). Nonparametric change-point tests for long-range dependent data. *to appear in Scandinavian Journal of Statistics.*

Deo, R. and Chen, W. (2000). On the integral of the squared periodogram. *Stochastic Process and their Applications*, 85:159–176.

Dette, H., Preuß, P., and Vetter, M. (2011a). A measure of stationarity in locally stationary processes with applications to testing. *Journal of the American Statistical Association*, 106(495):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.

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

Erdelyi, A. (1954a). *Tables of Integral Transforms, Vol.1.* McGraw-Hill, New York.

Erdelyi, A. (1954b). *Tables of Integral Transforms, Vol.2.* McGraw-Hill, New York.

Eubank, R. and LaRiccia, V. (1992). Asymptotic comparison of Cramer-von Mises and nonparametric function estimation techniques for testing goodness-of-fit. *Annals of Statistics*, 20(4):2071–2086.

Fan, J. (1996). Test of significance based on wavelet thresholding and Neyman's truncation. *Journal of the American Statistical Association*, 91(434):674–688.

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

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

Ghosh, B. and Huang, W.-M. (1991). The power and optimal kernel of the Bickel-Rosenblatt test for goodness of fit. *Annals of Statistics*, 19(2):999–1009.

Gradshteyn, I. and Ryzhik, I. (1980). *Table of Integrals, Series, and Products.* Academic Press, New York.

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

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

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

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

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

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

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

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

Preuß, P. and Vetter, M. (2012). On discriminating between long-range dependence and non stationarity. Technical report, Ruhr-Universität Bochum. http://www.ruhr-uni-bochum.de/imperia/md/content/mathematik3/publications/longmemory24092012.pdf.

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

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

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

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

Taniguchi, M. (1980). On estimation of the integrals of certain functions of spectral density. *Journal of Applied Probability*, 17:73–83.

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

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

6 Appendix: technical details

In the following, we will state two results which will be central for the proof of the statements in Sections 3 and 4.

Theorem 6.1. *If Assumption 2.1 is satisfied with $d_\infty < 1/4$, the following statements are correct.*

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

where the vector $\tilde{C}_T \in \mathbb{R}^2$ is of order $O\big(N^2/T^2 + \log(N)/(MN^{1-4d_\infty})\big)$. In particular, this term vanishes if the functions $\psi_l(u)$ are independent of u for all $l \in \mathbb{Z}$.

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

where

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

and $\tilde{\Sigma}_T$ is the same as the matrix Σ in (3.3) 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 $l_1 + l_2 \ge 3$, then we have*

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

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 (2012), that there exist parameters $\hat{\psi}_{l,p}$ such that the bootstrap process $X_{t,T}^*$ defined in (4.7) can be represented as

$$X_{t,T}^* = \sum_{l=0}^\infty \hat{\psi}_{l,p} Z_{t-l}^*, \tag{6.1}$$

where Z_t^* are the innovations from part 2) of the bootstrap description. We now assume that the null hypothesis (3.5) holds, and consider the process

$$X_{t,T,2}^* = \sum_{l=0}^\infty \psi_l Z_{t-l}^*, \tag{6.2}$$

where the coefficients $\psi_l = \psi_l(u)$ are the coefficients in (2.10). We then define $\hat{D}_{T,2}^{2,*}$ as $\hat{D}_T^2$ in (2.18) whereby the random variables $X_{t,T}$ are replaced by $X_{t,T,2}^*$. The next theorem shows that the random variable $\hat{D}_T^{2,*}$ can be approximated by $\hat{D}_{T,2}^{2,*}$.

Theorem 6.2. *Let $\alpha > 0$ be fixed and denote with $A_T(\alpha)$ the event where $|\hat{\underline{d}} - \underline{d}| \le \alpha/4$. If Assumption 2.1 and the inequality*

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

are satisfied, then

a) $\mathbb{E}\Big((\hat{D}_T^{2,*} - \hat{D}_{T,2}^{2,*})1_{A_T(\alpha)}\Big) = O\Big(p^4 \log(T)^{3/2}N^{4\underline{d}-1+\alpha}T^{-1/2}\Big),$

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

6.1 Proof of Theorem 6.1:

Proof of part a): We define $\tilde{t}_j := t_j - N/2 + 1$, $\tilde{\psi}_l(u_{j,p}) := \psi_l\left(\frac{\tilde{t}_j + p}{T}\right)$, $Z_{a,b} := Z_{a-N/2+1+b}$ and write similar to Dette et al. (2011b)

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

$$
= \frac{1}{2T}\sum_{j=1}^{M}\sum_{k=-\lfloor \frac{N-1}{2}\rfloor}^{\lfloor \frac{N}{2}\rfloor}\frac{1}{(2\pi N)^2}\sum_{p,q,r,s=0}^{N-1}\sum_{l,m,n,o=0}^{\infty} e^{-i(p-q+r-s)\lambda_{k,N}}
$$

$$
\psi_{\tilde{t}_j+p,T,l}\psi_{\tilde{t}_j+q,T,m}\psi_{\tilde{t}_j+r,T,n}\psi_{\tilde{t}_j+s,T,o}\mathbb{E}[Z_{t_j,p-l}Z_{t_j,q-m}Z_{t_j,r-n}Z_{t_j,s-o}]
$$

$$
= E_{N,T}^1 + E_{N,T}^2 + A_{N,T} + B_{N,T},
$$

where we use the notation

$$
E_{N,T}^1 = \frac{1}{2T}\sum_{j=1}^{M}\sum_{k=-\lfloor \frac{N-1}{2}\rfloor}^{\lfloor \frac{N}{2}\rfloor}\frac{1}{(2\pi N)^2}\sum_{p,q,r,s=0}^{N-1}\sum_{l,m,n,o=0}^{\infty}\psi_l(u_j)\psi_m(u_j)\psi_n(u_j)\psi_o(u_j)e^{-i(p-q+r-s)\lambda_{k,N}}
$$

$$
\left(\mathbb{E}[Z_{t_j,p-l}Z_{t_j,q-m}]\mathbb{E}[Z_{t_j,r-n}Z_{t_j,s-o}] + \mathbb{E}[Z_{t_j,p-l}Z_{t_j,s-o}]\mathbb{E}[Z_{t_j,q-m}Z_{t_j,r-n}]\right),
$$

$$
E_{N,T}^2 = \frac{1}{2T}\sum_{j=1}^{M}\sum_{k=-\lfloor \frac{N-1}{2}\rfloor}^{\lfloor \frac{N}{2}\rfloor}\frac{1}{(2\pi N)^2}\sum_{p,q,r,s=0}^{N-1}\sum_{l,m,n,o=0}^{\infty}\psi_l(u_j)\psi_m(u_j)\psi_n(u_j)\psi_o(u_j)
$$

$$
e^{-i(p-q+r-s)\lambda_{k,N}}\mathbb{E}[Z_{t_j,p-l}Z_{t_j,r-n}]\mathbb{E}[Z_{t_j,q-m}Z_{t_j,s-o}],
$$

$$
A_{N,T} = \frac{1}{2T}\sum_{j=1}^{M}\sum_{k=-\lfloor \frac{N-1}{2}\rfloor}^{\lfloor \frac{N}{2}\rfloor}\frac{1}{(2\pi N)^2}\sum_{p,q,r,s=0}^{N-1}\sum_{l,m,n,o=0}^{\infty} e^{-i(p-q+r-s)\lambda_{k,N}}\mathbb{E}[Z_{t_j,p-l}Z_{t_j,q-m}Z_{t_j,r-n}Z_{t_j,s-o}]
$$

$$
\left\{\left(\tilde{\psi}_l(u_{j,p}) - \psi_l(u_j)\right)\tilde{\psi}_m(u_{j,q})\tilde{\psi}_n(u_{j,r})\tilde{\psi}_o(u_{j,s}) + \psi_l(u_j)\left(\tilde{\psi}_m(u_{j,q}) - \psi_m(u_j)\right)\tilde{\psi}_n(u_{j,r})\tilde{\psi}_o(u_{j,s})\right.
$$

$$
\left. + \psi_l(u_j)\psi_m(u_j)\left(\tilde{\psi}_n(u_{j,r}) - \psi_n(u_j)\right)\tilde{\psi}_o(u_{j,s}) + \psi_l(u_j)\psi_m(u_j)\psi_n(u_j)\left(\tilde{\psi}_o(u_{j,s}) - \psi_o(u_j)\right)\right\},
$$

$$
B_{N,T} = \frac{1}{2T}\sum_{j=1}^{M}\sum_{k=-\lfloor \frac{N-1}{2}\rfloor}^{\lfloor \frac{N}{2}\rfloor}\frac{1}{(2\pi N)^2}\sum_{p,q,r,s=0}^{N-1}\sum_{l,m,n,o=0}^{\infty} e^{-i(p-q+r-s)\lambda_{k,N}}\mathbb{E}[Z_{t_j,p-l}Z_{t_j,q-m}Z_{t_j,r-n}Z_{t_j,s-o}]
$$

$$
\left\{\left(\psi_{\tilde{t}_j+p,T,l} - \tilde{\psi}_l(u_{j,p})\right)\tilde{\psi}_m(u_{j,q})\tilde{\psi}_n(u_{j,r})\tilde{\psi}_o(u_{j,s}) + \psi_{\tilde{t}_j+p,T,l}\left(\psi_{\tilde{t}_j+q,T,m} - \tilde{\psi}_m(u_{j,q})\right)\tilde{\psi}_n(u_{j,r})\tilde{\psi}_o(u_{j,s})\right.
$$

$$
\left. + \psi_{\tilde{t}_j+p,T,l}\psi_{\tilde{t}_j+q,T,m}\left(\psi_{\tilde{t}_j+r,T,n} - \tilde{\psi}_n(u_{j,r})\right)\tilde{\psi}_o(u_{j,s}) + \psi_{\tilde{t}_j+p,T,l}\psi_{\tilde{t}_j+q,T,m}\psi_{\tilde{t}_j+r,T,n}\left(\psi_{\tilde{t}_j+s,T,o} - \tilde{\psi}_o(u_{j,s})\right)\right\}.
$$

Note that $B_{N,T}$ corresponds to the error which 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 following four statements conclude the proof for $\mathbb{E}[\hat{F}_{1,T}]$

$$
E_{N,T}^1 = \frac{1}{2\pi M}\sum_{j=1}^{M}\int_{-\pi}^{\pi} f^2(u_j, \lambda)\, d\lambda + O\left(\frac{1}{N^{1-4d_\infty}}\right), \tag{6.4}
$$

$$E_{N,T}^2 = O\left(\frac{1}{N^{1-4d_\infty}}\right), \tag{6.5}$$

$$A_{N,T} = O\left(\frac{\log N}{MN^{1-4d_\infty}}\right) + O\left(\frac{N^2}{T^2}\right), \tag{6.6}$$

$$B_{N,T} = O\left(\frac{1}{T}\right). \tag{6.7}$$

Proof of (6.4): 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, we obtain that only those terms contribute to the sum where the conditions $0 \le p = q + l - m \le N - 1$ and $0 \le r = s + n - o \le N - 1$ are satisfied, which implies the inequality $\max\{|l - m|, |n - o|\} \le N - 1$. Thus, the first term in $E_{N,T}^1$ can be expressed as

$$\frac{1}{2T} \sum_{j=1}^{M} \sum_{k=-\lfloor\frac{N-1}{2}\rfloor}^{\lfloor\frac{N}{2}\rfloor} \frac{1}{(2\pi N)^2} \sum_{\substack{l,m,n,o=0 \\ |l-m|\le N-1 \\ |n-o|\le N-1}}^{\infty} \psi_l(u_j)\psi_m(u_j)\psi_n(u_j)\psi_o(u_j)e^{-i(l-m+n-o)\lambda_{k,N}}$$

$$\times (N - |l - m|)(N - |n - o|)$$

$$= \frac{1}{2M} \sum_{j=1}^{M} \sum_{h=-1}^{1} \frac{1}{(2\pi N)^2} \sum_{\substack{l,m,n,o=0 \\ |l-m|\le N-1 \\ |n-o|\le N-1 \\ l-m+n-o=hN}}^{\infty} \psi_l(u_j)\psi_m(u_j)\psi_n(u_j)\psi_o(u_j)(N - |l - m|)(N - |n - o|),$$

where we used the well known identity

$$\sum_{k=0}^{N-1} \exp(-i\lambda_{k,N}r) = \begin{cases} N & \text{if } r = Nh \text{ for some } h \in \mathbb{Z} \\ 0 & \text{else,} \end{cases} \tag{6.8}$$

[note that we only have to consider three possible values of h since $\max\{|l - m|, |n - o|\} \le N - 1$]. 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

$$E_{N,T,0}^1 = \frac{1}{2M} \sum_{j=1}^{M} \frac{1}{(2\pi)^2} \sum_{\substack{l,m,n,o=0 \\ l-m+n-o=0}}^{\infty} \psi_l(u_j)\psi_m(u_j)\psi_n(u_j)\psi_o(u_j) = \frac{1}{4\pi M} \sum_{j=1}^{M} \int_{-\pi}^{\pi} f^2(u_j, \lambda)\, d\lambda,$$

$$E_{N,T,1}^1 = \frac{1}{MN^2} \sum_{j=1}^{M} \sum_{h=-1}^{1} \sum_{\substack{l,m,n,o=0 \\ |l-m|\le N-1 \\ |n-o|\le N-1 \\ l-m+n-o=hN}}^{\infty} \psi_l(u_j)\psi_m(u_j)\psi_n(u_j)\psi_o(u_j) \tag{6.9}$$

$$\times (-N|l - m| - N|n - o| + |l - m||n - o|),$$

$$E_{N,T,2}^1 = \frac{1}{M} \sum_{j=1}^{M} \sum_{h\in\{-1,1\}} \sum_{\substack{l,m,n,o=0 \\ |l-m|\le N-1 \\ |n-o|\le N-1 \\ l-m+n-o=hN}}^{\infty} \psi_l(u_j)\psi_m(u_j)\psi_n(u_j)\psi_o(u_j), \tag{6.10}$$

$$E_{N,T,3}^1 \;=\; \frac{1}{M}\sum_{j=1}^{M}\sum_{\substack{l,m,n,o=0\\ N\le |l-m|\\ N\le |n-o|\\ l-m+n-o=0}}^{\infty}\psi_l(u_j)\psi_m(u_j)\psi_n(u_j)\psi_o(u_j). \tag{6.11}$$

In order to complete the proof of (6.4), it therefore suffices to demonstrate that the last three expressions are of order $O\big(\frac{1}{N^{1-4d_\infty}}\big)$. We commence with (6.9). Setting $o = l - m + n - hN \ge 0$ and using (2.4), it follows that there exists a constant $C \in \mathbb{R}$ such that

$$E_{N,T,1}^1 \;\le\; \frac{C}{MN}\sum_{j=1}^{M}\sum_{h=-1}^{1}\sum_{\substack{l,m,n=1\\ 1\le|l-m-hN|\le N-1\\ 1\le l-m+n-hN}}^{\infty}\frac{1}{l^{1-d_\infty}}\frac{1}{m^{1-d_\infty}}\frac{1}{n^{1-d_\infty}}\frac{|l-m-hN|}{(l-m+n-hN)^{1-d_\infty}} \tag{6.12}$$

(note that 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 will be employed continuously throughout this proof without mentioning it explicitly. Note that 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 (6.9), and that $C \in \mathbb{R}^+$ denotes a universal constant throughout the whole proof. Setting $z := l - m - hN$, we obtain $|z| = |l-m-hN| \le N-1$ and the expression on the right hand side of (6.12) can be written as

$$\frac{C}{N}\sum_{h=-1}^{1}\sum_{\substack{z\in\mathbb{Z}\\ 1\le|z|\le N-1}}\sum_{\substack{m,n=1\\ 1\le z+m+hN\\ 1\le n+z}}^{\infty}\frac{1}{(z+m+hN)^{1-d_\infty}}\frac{1}{m^{1-d_\infty}}\frac{1}{n^{1-d_\infty}}\frac{|z|}{(n+z)^{1-d_\infty}}$$

$$\overset{(7.2)}{\lesssim}\;\frac{1}{N}\sum_{h=-1}^{1}\sum_{\substack{z\in\mathbb{Z}\\ 1\le|z|\le N-1}}\frac{|z|^{2d_\infty}}{|z+hN|^{1-2d_\infty}}\;\lesssim\;\frac{1}{N^{1-2d_\infty}}\sum_{h=-1}^{1}\sum_{\substack{z\in\mathbb{Z}\\ 1\le|z|\le N-1}}\frac{1}{|z+hN|^{1-2d_\infty}}\;\lesssim\;\frac{1}{N^{1-4d_\infty}},$$

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 (2.4), (6.8), (7.2) and similar arguments, we obtain that (6.10) is bounded by

$$\sum_{h\in\{-1,1\}}\sum_{\substack{z\in\mathbb{Z}\\ 1\le|z|\le N-1}}\frac{1}{|z+hN|^{1-2d_\infty}}\frac{1}{|z|^{1-2d_\infty}}\overset{(7.1),(7.2)}{\lesssim}\frac{1}{N^{1-4d_\infty}},$$

and since (6.11) is shown analogously, we therefore conclude the proof of (6.4).

Proof of (6.5): The result follows by similar arguments as used in the treatment of (6.9)–(6.11).

Proof of (6.6): 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, i.e. we only consider

$$A_{N,T,1} := \frac{1}{2T} \sum_{j=1}^{M} \sum_{k=-\lfloor\frac{N-1}{2}\rfloor}^{\lfloor\frac{N}{2}\rfloor} \frac{1}{(2\pi N)^2} \sum_{p,q,r,s=0}^{N-1} \sum_{l,m,n,o=0}^{\infty} e^{-i(p-q+r-s)\lambda_{k,N}}$$

$$\mathbb{E}[Z_{t_j,p-l}Z_{t_j,q-m}]\mathbb{E}[Z_{t_j,r-n}Z_{t_j,s-o}]\Big(\tilde{\psi}_l(u_{j,p}) - \psi_l(u_j)\Big)\psi_m(u_j)\psi_n(u_j)\psi_o(u_j).$$

Using a Taylor expansion, we can write

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

with $\eta_{l,j,p} \in (u_j - N/(2T), u_j + N/(2T))$, and therefore $A_{N,T,1}$ splits into two terms which will be 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 (6.8) afterwards yields $0 \leq m = l + r - s - hN$ with $h \in \{-1, 0, 1\}$ and this combined with (2.4) and (2.6) implies

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

$$\bigg| \sum_{\substack{p=0\\0\leq p+r-s-hN\leq N-1}}^{N-1} \Big(\frac{-N/2+1+p}{T}\Big)\bigg|.$$

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

$$\bigg| \sum_{\substack{p=0\\0\leq p+q\leq N-1}}^{N-1} \Big(\frac{-N/2+1+p}{T}\Big)\bigg| \tag{6.13}$$

$$= \frac{N}{2T} \times 1_{\{q=0\}} + \min\Big(\frac{N|q|}{T}, \frac{(N-|q|)|q|}{T}\Big) O(1) \times 1_{\{1\leq|q|\leq N-1\}} \tag{6.14}$$

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

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

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

where we used Lemma 7.1c) and 7.1b) for the first and second inequality, respectively. Next, we show that $A_{N,T,1}^{(2)}$ is of order $O(N^2T^{-2})$, and for this reason we choose $\varepsilon > 0$ such that $1-4d_\infty-\varepsilon > 0$.

25

Using (2.4), (2.6), (2.7) and $h \in \{-1, 0, 1\}$, the claim then follows by a further application of Lemma 7.1b)

$$\frac{C}{N^2} \sum_{r,s=0}^{N-1} \sum_{\substack{l,o=1 \\ 1 \leq o+r-s \\ 1 \leq l+r-s-hN}}^{\infty} \frac{\log^2(l)}{l^{1-d_\infty}} \frac{1}{(l+r-s-hN)^{1-d_\infty}} \frac{1}{(o+r-s)^{1-d_\infty}} \frac{1}{o^{1-d_\infty}}$$

$$\times \sum_{\substack{p=0 \\ 0 \leq p+r-s-hN \leq N-1}}^{N-1} \left(\frac{-N/2+1+p}{T}\right)^2$$

$$\lesssim \frac{N}{T^2} \sum_{r,s=0}^{N-1} \sum_{\substack{l,o=1 \\ 1 \leq o+r-s \\ 1 \leq l+r-s-hN}}^{\infty} \frac{1}{l^{1-d_\infty-\varepsilon}} \frac{1}{(l+r-s-hN)^{1-d_\infty}} \frac{1}{(o+r-s)^{1-d_\infty}} \frac{1}{o^{1-d_\infty}}$$

$$\lesssim \frac{N}{T^2} \sum_{r,s=0}^{N-1} \frac{1}{|r-s-hN|^{1-2d_\infty-\varepsilon}} \frac{1}{|r-s|^{1-2d_\infty}} \lesssim \frac{N^2}{T^2}.$$

Proof of (6.7): The statement follows from (2.3) and similar arguments as given in the proofs of (6.4) and (6.6).

In order to proof the assertion for $\hat{F}_{2,T}$, one proceeds in the same way and the details are omitted for the sake of brevity. 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 will be denoted by $d_{N,T}$.

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

$$\text{Var}(\hat{F}_{1,T}) = \frac{1}{4}\text{Var}\left(\frac{1}{T}\sum_{j=1}^{M}\sum_{k=-\lfloor\frac{N-1}{2}\rfloor}^{\lfloor\frac{N}{2}\rfloor} I_N(u_j, \lambda_{k,N})^2\right) = \frac{5}{\pi}\frac{1}{TM}\sum_{j=1}^{M}\int_{-\pi}^{\pi} f^4(u_j, \lambda)\,d\lambda + O(T, d_\infty)$$

and recall the definition of the remainder

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

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 et al. (2011b), we obtain that

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

where

$$V^*(\nu') = \frac{1}{T^2}\sum_{j_1,j_2=1}^{M}\sum_{k_1,k_2=-\lfloor\frac{N-1}{2}\rfloor}^{\lfloor\frac{N}{2}\rfloor} \frac{1}{(2\pi N)^4} \sum_{p_1,q_1,r_1,s_1=0}^{N-1}\sum_{p_2,q_2,r_2,s_2=0}^{N-1}\sum_{v_1,w_1,x_1,y_1=0}^{\infty}\sum_{v_2,w_2,x_2,y_2=0}^{\infty}$$

26

$$\psi_{v_1}(u_{j_1})\psi_{w_1}(u_{j_1})\psi_{x_1}(u_{j_1})\psi_{y_1}(u_{j_1})\psi_{v_2}(u_{j_2})\psi_{w_2}(u_{j_2})\psi_{x_2}(u_{j_2})\psi_{y_2}(u_{j_2})$$

$$e^{-i(p_1-q_1+r_1-s_1)\lambda_{k_1}}e^{-i(p_2-q_2+r_2-s_2)\lambda_{k_2}}$$

$$\mathbb{E}[Z_{t_{j_1},p_1-v_1}Z_{t_{j_1},q_1-w_1}]\mathbb{E}[Z_{t_{j_1},r_1-x_1}Z_{t_{j_2},p_2-v_2}]\mathbb{E}[Z_{t_{j_1},s_1-y_1}Z_{t_{j_2},q_2-w_2}]\mathbb{E}[Z_{t_{j_2},r_2-x_2}Z_{t_{j_2},s_2-y_2}]$$

and

$$V^*(\nu'') = \frac{1}{T^2}\sum_{j_1,j_2=1}^{M}\sum_{k_1,k_2=-\lfloor\frac{N-1}{2}\rfloor}^{\lfloor\frac{N}{2}\rfloor}\frac{1}{(2\pi N)^4}\sum_{p_1,q_1,r_1,s_1=0}^{N-1}\sum_{p_2,q_2,r_2,s_2=0}^{N-1}\sum_{v_1,w_1,x_1,y_1=0}^{\infty}\sum_{v_2,w_2,x_2,y_2=0}^{\infty}$$

$$\psi_{v_1}(u_{j_1})\psi_{w_1}(u_{j_1})\psi_{x_1}(u_{j_1})\psi_{y_1}(u_{j_1})\psi_{v_2}(u_{j_2})\psi_{w_2}(u_{j_2})\psi_{x_2}(u_{j_2})\psi_{y_2}(u_{j_2})$$

$$e^{-i(p_1-q_1+r_1-s_1)\lambda_{k_1}}e^{-i(p_2-q_2+r_2-s_2)\lambda_{k_2}}\mathbb{E}[Z_{t_{j_1},p_1-v_1}Z_{t_{j_2},p_2-v_2}]\mathbb{E}[Z_{t_{j_1},r_1-x_1}Z_{t_{j_2},r_2-x_2}]$$

$$\mathbb{E}[Z_{t_{j_1},q_1-w_1}Z_{t_{j_2},q_2-w_2}]\mathbb{E}[Z_{t_{j_1},s_1-y_1}Z_{t_{j_2},s_2-y_2}].$$

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 (6.8). Then, $V^*(\nu')$ can be written as

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

$$\psi_{v_1}(u_{j_1})\psi_{w_1}(u_{j_1})\psi_{x_1}(u_{j_1})\psi_{y_1}(u_{j_1})\psi_{v_2}(u_{j_2})\psi_{w_2}(u_{j_2})\psi_{x_2}(u_{j_2})\psi_{y_2}(u_{j_2})$$

$$e^{-i(v_1-w_1+r_1-s_1)\lambda_{k_1}}e^{-i(r_1+v_2-x_1-s_1-w_2+y_1-y_2+x_2)\lambda_{k_2}}$$

$$= \frac{1}{M^2}\sum_{j_1,j_2=1}^{M}\sum_{h_1,h_2=-1}^{1}\frac{1}{(2\pi N)^4}\sum_{r_1,s_1=0}^{N-1}\sum_{\substack{v_1,w_1,x_1,y_1=0 \\ v_1-w_1+r_1-s_1=h_1N \\ 0\le|v_1-w_1|\le N-1}}^{\infty}\sum_{\substack{v_2,w_2,x_2,y_2=0 \\ r_1+v_2-x_1-s_1-w_2+y_1-y_2+x_2=h_2N \\ 0\le r_1+v_2-x_1+(j_1-j_2)N\le N-1 \\ 0\le s_1+w_2-y_1+(j_1-j_2)N\le N-1 \\ 0\le|y_2-x_2|\le N-1}}^{\infty}$$

$$\psi_{v_1}(u_{j_1})\psi_{w_1}(u_{j_1})\psi_{x_1}(u_{j_1})\psi_{y_1}(u_{j_1})\psi_{v_2}(u_{j_2})\psi_{w_2}(u_{j_2})\psi_{x_2}(u_{j_2})\psi_{y_2}(u_{j_2})$$

$$(N^2 - N|y_2-x_2| - N|v_1-w_1| + |v_1-w_1||y_2-x_2|).$$

An application of (2.4) yields similar to the proof of part a) that the above expression is of order $O\big(\frac{1}{N^{1-8d_\infty}T}\big)$, if

(i) $h_1,h_2 \in \{-1,1\}$ [compare (6.10)],

(ii) $j_1 \neq j_2$ [we prove this claim in Lemma 7.2 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 (6.9)],

(iv) we drop $0 \leq |v_1-w_1| \leq N-1$ and $0 \leq |y_2-x_2| \leq N-1$ [compare (6.11)],

(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 (6.9)].

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

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

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

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

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

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

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

we deduce

$$
\begin{aligned}
V^*(\nu'') &= \left(\frac{2}{3} + \frac{1}{6} + \frac{1}{6}\right) \frac{1}{M^2 N} \sum_{j_1=1}^{M} \frac{1}{(2\pi)^4} \sum_{\substack{v_1,w_1,x_1,y_1=0}}^{\infty} \sum_{\substack{v_2,w_2,x_2,y_2=0 \\ v_1-v_2+w_2-w_1+x_1-x_2+y_2-y_1=0}}^{\infty} \\
&\quad \psi_{v_1}(u_{j_1})\psi_{w_1}(u_{j_1})\psi_{x_1}(u_{j_1})\psi_{y_1}(u_{j_1})\psi_{v_2}(u_{j_1})\psi_{w_2}(u_{j_1})\psi_{x_2}(u_{j_1})\psi_{y_2}(u_{j_1}) + O(T, d_\infty) \\
&= \frac{1}{TM}\frac{1}{2\pi} \sum_{j_1=1}^{M} \int_{-\pi}^{\pi} f^4(u_{j_1}, \lambda)\, d\lambda + O(T, d_\infty). \tag{6.15}
\end{aligned}
$$

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 et al. (2011a), it is sufficient to choose an arbitrary indecomposable partition

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

of the table

$$
\begin{array}{cccc}
Z_{t_{j_1},p_1-g_1} & Z_{t_{j_1},q_1-m_1} & Z_{t_{j_1},r_1-n_1} & Z_{t_{j_1},s_1-o_1} \\
\vdots & \vdots & \vdots & \vdots \\
Z_{t_{j_l},p_l-g_l} & Z_{t_{j_l},q_l-m_l} & Z_{t_{j_l},r_l-n_l} & Z_{t_{j_l},s_l-o_l}
\end{array}
\tag{6.16}
$$

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

$$
\frac{T^{\frac{l}{2}}}{T^l}\frac{1}{N^{2l}}\sum_{j_1,\ldots,j_l=1}^{M}\sum_{k_1,\ldots,k_l=-\lfloor\frac{N-1}{2}\rfloor}^{\lfloor\frac{N}{2}\rfloor}\sum_{p_1,\ldots,s_l=0}^{N-1}\sum_{v_1,w_1,x_1,y_1=0}^{\infty}\cdots\sum_{v_l,w_l,x_l,y_l=0}^{\infty}
$$

$$
\tilde{\psi}_{v_1}(u_{j_1,p_1})\tilde{\psi}_{w_1}(u_{j_1,q_1})\tilde{\psi}_{x_1}(u_{j_1,r_1})\tilde{\psi}_{y_1}(u_{j_1,s_1})\tilde{\psi}_{v_2}(u_{j_2,p_2})\tilde{\psi}_{w_2}(u_{j_2,q_2})\tilde{\psi}_{x_2}(u_{j_2,r_2})\tilde{\psi}_{y_2}(u_{j_2,s_2})
$$

$$
\cdots
$$

$$
\tilde{\psi}_{v_l}(u_{j_l,p_l})\tilde{\psi}_{w_l}(u_{j_l,q_l})\tilde{\psi}_{x_l}(u_{j_l,r_l})\tilde{\psi}_{y_l}(u_{j_l,s_l})e^{-i(p_1-q_1+r_1-s_1)\lambda_{k_1}}\cdots e^{-i(p_l-q_l+r_l-s_l)\lambda_{k_l}}
$$

$$
\mathbb{E}[Z_{t_{i_1},a_1-v_1}Z_{t_{i_2},a_2-w_1}]\mathbb{E}[Z_{t_{i_3},a_3-x_1}Z_{t_{i_4},a_4-y_1}]\mathbb{E}[Z_{t_{i_5},a_5-v_2}Z_{t_{i_6},a_6-w_2}]\mathbb{E}[Z_{t_{i_7},a_7-x_2}Z_{t_{i_8},a_8-y_2}]
$$

$$
\cdots
$$

$$
\mathbb{E}[Z_{t_{i_{4l-3}},a_{4l-3}-v_l}Z_{t_{i_{4l-2}},a_{4l-2}-w_l}]\mathbb{E}[Z_{t_{i_{4l-1}},a_{4l-1}-x_l}Z_{t_{i_{4l}},a_{4l}-y_l}]
$$

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

$$
a_{4m+1}-a_{4m+2}+w_{m+1}-v_{m+1}+(i_{4m+1}-i_{4m+2})N=0 \text{ for } m=0,\ldots,l-1, \tag{6.17}
$$

$$
a_{4m+3}-a_{4m+4}+y_{m+1}-x_{m+1}+(i_{4m+3}-i_{4m+4})N=0 \text{ for } m=0,\ldots,l-1, \tag{6.18}
$$

$$
p_i-q_i+r_i-s_i=h_iN \text{ for } i=1,2,\ldots,l \text{ and } h_i\in\{-1,0,1\} \tag{6.19}
$$

follow. Rearranging the equations in (6.19) for a variable and plugging them into the l equations (6.17) (where in every equation only <u>one</u> variable is replaced) yields, due to the indecomposability of the partition and $v_{m+1},x_{m+1}\geq 0$, that the conditions

$$
\begin{array}{ll}
(1) & 0\leq v_1=\tilde{a}_1-\tilde{a}_2+\tilde{a}_3-\tilde{a}_4+w_1+(i_1-i_2+h_1)N, \\
(2) & 0\leq v_2=\tilde{a}_7-\tilde{a}_8+\tilde{a}_9-\tilde{a}_{10}+w_2+(i_5-i_6+h_2)N, \\
& \vdots \\
(l) & 0\leq v_l=\tilde{a}_{6l-5}-\tilde{a}_{6l-4}+\tilde{a}_{6l-3}-\tilde{a}_{6l-2}+w_l+(i_{4l-3}-i_{4l-2}+h_l)N, \\
(l+1) & 0\leq x_1=\tilde{a}_5-\tilde{a}_6+y_1+(i_3-i_4)N, \\
(l+2) & 0\leq x_2=\tilde{a}_{11}-\tilde{a}_{12}+y_2+(i_7-i_8)N, \\
& \vdots \\
(2l) & 0\leq x_l=\tilde{a}_{6l-1}-\tilde{a}_{6l}+y_l+(i_{4l-1}-i_{4l})N
\end{array}
\tag{6.20}
$$

must hold, where $\{\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_{6l}\} \in \{p_1, \ldots, s_l\}$ and $|\{\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_{6l}\}| = 3l$. By employing (2.4), we can bound the above expression up to a constant by

$$\frac{1}{M^l} \frac{T^{\frac{l}{2}}}{N^{2l}} \sum_{j_1,\ldots,j_l=1}^{M} \sum_{h_1,\ldots,h_l=-1}^{1} \sum_{\tilde{a}_1,\ldots,\tilde{a}_{6l}=0}^{N-1}$$

$$\sum_{\substack{w_1,y_1=1 \\ \tilde{a}_1-\tilde{a}_2+\tilde{a}_3-\tilde{a}_4+w_1+(i_1-i_2+h_1)N\geq 1 \\ \tilde{a}_5-\tilde{a}_6+y_1+(i_3-i_4)N\geq 1}}^{\infty} \cdots \sum_{\substack{w_l,y_l=1 \\ \tilde{a}_{6l-5}-\tilde{a}_{6l-4}+\tilde{a}_{6l-3}-\tilde{a}_{6l-2}+w_l+(i_{4l-3}-i_{4l-2}+h_l)N\geq 1 \\ \tilde{a}_{6l-1}-\tilde{a}_{6l}+y_l+(i_{4l-1}-i_{4l})N\geq 1}}^{\infty}$$

$$\frac{1}{(\tilde{a}_1-\tilde{a}_2+\tilde{a}_3-\tilde{a}_4+w_1+(i_1-i_2+h_1)N)^{1-d_\infty}} \frac{1}{w_1^{1-d_\infty}} \frac{1}{(\tilde{a}_5-\tilde{a}_6+y_1+(i_3-i_4)N)^{1-d_\infty}} \frac{1}{y_1^{1-d_\infty}}$$

$$\frac{1}{(\tilde{a}_7-\tilde{a}_8+\tilde{a}_9-\tilde{a}_{10}+w_2+(i_5-i_6+h_2)N)^{1-d_\infty}} \frac{1}{w_2^{1-d_\infty}} \frac{1}{(\tilde{a}_{11}-\tilde{a}_{12}+y_2+(i_7-i_8)N)^{1-d_\infty}} \frac{1}{y_2^{1-d_\infty}}$$

$$\cdots$$

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

$$\frac{1}{(\tilde{a}_{6l-1}-\tilde{a}_{6l}+y_l+(i_{4l-1}-i_{4l})N)^{1-d_\infty}} \frac{1}{y_l^{1-d_\infty}}.$$

Using Lemma 7.1b) in the appendix, this term can be (up to a constant) bounded by

$$\frac{1}{M^l} \frac{T^{\frac{l}{2}}}{N^{2l}} \sum_{j_1,\ldots,j_l=1}^{M} \sum_{h_1,\ldots,h_l=-1}^{1} \sum_{\substack{\tilde{a}_1,\tilde{a}_2,\ldots,\tilde{a}_6=0 \\ |\tilde{a}_1-\tilde{a}_2+\tilde{a}_3-\tilde{a}_4+(i_1-i_2+h_1)N|\geq 1 \\ |\tilde{a}_5-\tilde{a}_6+(i_3-i_4)N|\geq 1}}^{N-1} \cdots \sum_{\substack{\tilde{a}_{6l-5},\tilde{a}_{6l-4},\ldots,\tilde{a}_{6l}=0 \\ |\tilde{a}_{6l-5}-\tilde{a}_{6l-4}+\tilde{a}_{6l-3}-\tilde{a}_{6l-2}+(i_{4l-3}-i_{4l-2}+h_l)N|\geq 1 \\ |\tilde{a}_{6l-1}-\tilde{a}_{6l}+(i_{4l-1}-i_{4l})N|\geq 1}}^{N-1}$$

$$\frac{1}{|\tilde{a}_1-\tilde{a}_2+\tilde{a}_3-\tilde{a}_4+(i_1-i_2+h_1)N|^{1-2d_\infty}} \frac{1}{|\tilde{a}_5-\tilde{a}_6+(i_3-i_4)N|^{1-2d_\infty}}$$

$$\frac{1}{|\tilde{a}_7-\tilde{a}_8+\tilde{a}_9-\tilde{a}_{10}+(i_5-i_6+h_2)N|^{1-2d_\infty}} \frac{1}{|\tilde{a}_{11}-\tilde{a}_{12}+(i_7-i_8)N|^{1-2d_\infty}}$$

$$\cdots$$

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

We now assume without loss of generality that

$$\tilde{a}_{6m+1}-\tilde{a}_{6m+2}+\tilde{a}_{6m+3}-\tilde{a}_{6m+4}+(i_{4m+1}-i_{4m+2}+h_{m+1})N \geq 1,$$
$$\tilde{a}_{6m+5}-\tilde{a}_{6m+6}+(i_{4m+3}-i_{4m+4})N \geq 1 \tag{6.21}$$

holds for $m = 0, 1, 2, \ldots, l-1$ (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 et al. (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, $k = 2, \ldots, l$. Thus it suffices to consider the following sum

$$\frac{1}{M^{l-1}} \frac{T^{\frac{l}{2}}}{N^{2l}} \sum_{\substack{\tilde{a}_1,\tilde{a}_2,\ldots,\tilde{a}_6=0 \\ \tilde{a}_1-\tilde{a}_2+\tilde{a}_3-\tilde{a}_4+C_1N\geq 1 \\ \tilde{a}_5-\tilde{a}_6+C_{l+1}N\geq 1}}^{N-1} \cdots \sum_{\substack{\tilde{a}_{6l-5},\tilde{a}_{6l-4},\ldots,\tilde{a}_{6l}=0 \\ \tilde{a}_{6l-5}-\tilde{a}_{6l-4}+\tilde{a}_{6l-3}-\tilde{a}_{6l-2}+C_lN\geq 1 \\ \tilde{a}_{6l-1}-\tilde{a}_{6l}+C_{2l}N\geq 1}}^{N-1}$$

$$\frac{1}{(\tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + C_1 N)^{1-2d_\infty}} \frac{1}{(\tilde{a}_5 - \tilde{a}_6 + C_{l+1}N)^{1-2d_\infty}}$$
$$\frac{1}{(\tilde{a}_7 - \tilde{a}_8 + \tilde{a}_9 - \tilde{a}_{10} + C_2 N)^{1-2d_\infty}} \frac{1}{(\tilde{a}_{11} - \tilde{a}_{12} + C_{l+2}N)^{1-2d_\infty}}$$
$$\cdots$$
$$\frac{1}{(\tilde{a}_{6l-5} - \tilde{a}_{6l-4} + \tilde{a}_{6l-3} - \tilde{a}_{6l-2} + C_l N)^{1-2d_\infty}} \frac{1}{(\tilde{a}_{6l-1} - \tilde{a}_{6l} + C_{2l}N)^{1-2d_\infty}}$$

with $C_1, C_2, \ldots, C_l \in \{-1, 0 \ldots, M\}$ and $C_{l+1}, C_{l+2}, \ldots, C_{2l} \in \{0, 1 \ldots, M-1\}$ (because of (6.21) and $\tilde{a}_i \in \{0, 1, 2, \ldots, N-1\}$ 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 will 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 7.1a) that

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

$$\lesssim \frac{1}{(-\tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + \tilde{a}_5 + (C_1 + C_{l+1})N)^{1-4d_\infty}} \qquad (6.22)$$

$$\lesssim \frac{T^{2d_\infty}}{(-\tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + \tilde{a}_5 + (C_1 + C_{l+1})N)^{1-2d_\infty}}. \qquad (6.23)$$

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 $\tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + C_1 N \geq 1$ 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 7.1b) that

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

$$\lesssim \frac{1}{|-\tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 - \tilde{a}_5 + (C_1 - C_{l+1})N|^{1-4d_\infty}} \qquad (6.24)$$

$$\lesssim \frac{T^{2d_\infty}}{|-\tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 - \tilde{a}_5 + (C_1 - C_{l+1})N|^{1-2d_\infty}}. \qquad (6.25)$$

In both cases, it is possible that variables cancel out, for example if $\tilde{a}_4 = \tilde{a}_5$ and $\tilde{a}_3 = \tilde{a}_5$ in the first and second case, respectively. We apply (6.22)–(6.25) in total $2l-2$-times. In the first $2l-4$-applications, we use (6.23) and (6.25) (depending on the algebraic sign of the variable which appears in both fractions) and in the $(2l-3)$th and $(2l-2)$th application we employ (6.22) and (6.24). We furthermore assume that h variables cancel out while utilizing these inequalities. Afterwards

$3l - (2l - 2) - h = l + 2 - h$ variables remain with $0 \leq h \leq l$, namely $\tilde{a}_{6l-1}$, $\tilde{a}_{6l}$ and $l - h$ other variables with values in $\{0, 1, 2, \ldots, N - 1\}$. Denoting these $l - h$ variables with $b_1, b_2, \ldots, b_{l-h}$ we obtain

$$\frac{1}{M^{l-1}} \frac{T^{\frac{l}{2}}}{N^{2l}} \sum_{\substack{\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_6 = 0 \\ \tilde{a}_1 - \tilde{a}_2 + \tilde{a}_3 - \tilde{a}_4 + C_1 N \geq 1 \\ \tilde{a}_5 - \tilde{a}_6 + C_{l+1} N \geq 1}}^{N-1} \cdots \sum_{\substack{\tilde{a}_{6l-5}, \tilde{a}_{6l-4}, \ldots, \tilde{a}_{6l} = 0 \\ \tilde{a}_{6l-5} - \tilde{a}_{6l-4} + \tilde{a}_{6l-3} - \tilde{a}_{6l-2} + C_l N \geq 1 \\ \tilde{a}_{6l-1} - \tilde{a}_{6l} + C_{2l} N \geq 1}}^{N-1}$$

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

$$\frac{1}{(\tilde{a}_7 - \tilde{a}_8 + \tilde{a}_9 - \tilde{a}_{10} + C_2 N)^{1 - 2d_\infty}} \frac{1}{(\tilde{a}_{11} - \tilde{a}_{12} + C_{l+2} N)^{1 - 2d_\infty}}$$

$$\cdots$$

$$\frac{1}{(\tilde{a}_{6l-5} - \tilde{a}_{6l-4} + \tilde{a}_{6l-3} - \tilde{a}_{6l-2} + C_l N)^{1 - 2d_\infty}} \frac{1}{(\tilde{a}_{6l-1} - \tilde{a}_{6l} + C_{2l} N)^{1 - 2d_\infty}}$$

$$\lesssim \frac{1}{M^{l-1}} \frac{T^{\frac{l}{2}}}{N^{2l}} \sum_{\substack{\tilde{a}_{6l-1}, \tilde{a}_{6l} = 0 \\ \tilde{a}_{6l-1} - \tilde{a}_{6l} + C_{2l} N \geq 1}}^{N-1} \sum_{\substack{b_1, b_2, \ldots, b_{l-h} = 0 \\ 1 \leq |\tilde{a}_{6l-1} - \tilde{a}_{6l} + \sum_{j=1}^{l-h} (-1)^{k_j} 2b_j + \sum_{j=1}^{2l-1} (-1)^{k_j} C_j N|}}^{N-1}$$

$$\frac{N^h T^{(2l-4)2d_\infty}}{|\tilde{a}_{6l-1} - \tilde{a}_{6l} + \sum_{j=1}^{l-h} (-1)^{k_j} 2b_j + \sum_{j=1}^{2l-1} (-1)^{k_j} C_j N|^{1 - 6d_\infty}} \frac{1}{(\tilde{a}_{6l-1} - \tilde{a}_{6l} + C_{2l} N)^{1 - 2d_\infty}}$$

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 the above term equals

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

If $|\sum_{j=1}^{2l-1} (-1)^{k_j} C_l N| \geq 1$ or $C_{2l} = 1$, we apply Lemma 7.1a) and b) in order to obtain the same upper bound (it can be shown that, in this case, there appears an additional factor $N^{1 - 8d_\infty}$ in the denominator, so the corresponding term is, in fact, of smaller order). The same upper bound holds for $h \leq l - 1$. □

6.2 Proof of Remark 3.6:

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 as in the proof of Theorem 6.1b) except that the term, where the variable h_1 in $V^*(v'')$ equals -1 or 1, does not occur, because for the integrated version one can use

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(-i\lambda r) d\lambda = \begin{cases} 1 & \text{if } r = 0, \\ 0 & \text{else,} \end{cases} \tag{6.26}$$

(for $r \in \mathbb{Z}$) instead of (6.8). Therefore, in the integrated case, we obtain

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

$$\psi_{v_1}(u_{j_1})\psi_{w_1}(u_{j_1})\psi_{x_1}(u_{j_1})\psi_{y_1}(u_{j_1})\psi_{v_2}(u_{j_1})\psi_{w_2}(u_{j_1})\psi_{x_2}(u_{j_1})\psi_{y_2}(u_{j_1}) + O(T, d_\infty)$$

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

instead of (6.15) and we recall that the order $O(T, d_\infty)$ is defined in (6.1). This yields that the asymptotic variance of $\sqrt{T}\tilde{F}_{1,T}$ is $\frac{14}{3\pi}\int_{-\pi}^{\pi}\int_0^1 f^4(u, \lambda)dud\lambda$ and does not coincide with the asymptotic variance of $\sqrt{T}\hat{F}_{1,T}$.

6.3 Proof of Theorem 6.2:

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 (6.1) and writing $I_N^*(u, \lambda)$ for the bootstrap analogue of $I_N(u, \lambda)$, we then get

$$\mathbb{E}((\hat{F}_{1,T}^* - \hat{F}_{1,T,2}^*)1_{A_T(\alpha)}|X_{1,T}, ..., X_{T,T}) = \frac{1}{2T}\sum_{j=1}^{M}\sum_{k=-\lfloor\frac{N-1}{2}\rfloor}^{\lfloor\frac{N}{2}\rfloor}\frac{1}{(2\pi N)^2}\sum_{p,q,r,s=0}^{N-1}\sum_{l,m,n,o=0}^{\infty}e^{-i(p-q+r-s)\lambda_{k,N}}$$

$$\hat{\psi}_{l,m,n,o,p}1_{A_T(\alpha)}\mathbb{E}[Z_{t_j,p-l}^*Z_{t_j,q-m}^*Z_{t_j,r-n}^*Z_{t_j,s-o}^*]$$

[compare the first set of equalities in the proof of Theorem 6.1 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

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

the above expression splits into four terms and for the sake of brevity we only consider the first one. The other cases are treated similarly. As in the proof of Theorem 6.1 a) we then 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 $\psi_l(u_j)\psi_m(u_j)\psi_m(u_j)\psi_o(u_j)$ 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 (6.3) and combine it with the fact that $|\hat{\underline{d}} - \underline{d}| < \alpha/4$ on the set $A_T(\alpha)$, we get

$$|\hat{\psi}_{l,p} - \psi_l| \leq Cp^4\log(T)^{3/2}T^{-1/2}|l|^{\alpha/4+\underline{d}-1} \quad \forall l \in I\!N, \tag{6.27}$$

which together with (2.4) and the assumptions of the theorem implies

$$|\hat{\psi}_{l,p}| \leq C|l|^{\alpha/4+\underline{d}-1} \quad \forall l, p \in I\!N. \tag{6.28}$$

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

$$\sup_u |\psi_l(u)| \leq C|l|^{\underline{d}-1} \tag{6.29}$$

in the proof of Theorem 6.1 a) to obtain appropriate bounds for the error. By using (6.27) and (6.28) instead of (6.29) and similar arguments as given in the proof of Theorem 6.1 a) it follows that

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

where $F_{1,T}^{*;-}$ 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 $F_{1,T}^{*;-}$ 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 6.1 b) employing (6.27) and (6.28) instead of (6.29) as above. The details are omitted for the sake of brevity. □

6.4   Proofs of the results in Section 3 and 4:

Proof of Theorem 3.2: The claim follows by employing the Cramér-Wold device in combination with Theorem 6.1. □

Proof of Theorem 3.4: Similarly to the proof of Theorem 6.1, the two equations

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

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

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

Proof of Theorem 4.2: 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 (2.10). 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 distributed. Part b) follows from the proof of Theorem 6.1, so we focus on part c) and d). Note that (2.11) and Theorem 6.1 a), b) imply

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

which directly yields part d). If we have

$$P(A_T(\alpha)) \to 1 \quad \text{as } T \to \infty \tag{6.30}$$

for every $\alpha > 0$, Part c) follows from Theorem 6.2, (4.8), the conditions on the rate of $p(T)$ if α is chosen sufficiently small. Finally, (6.30) is a consequence of Lemma 4.3 in Preuß and Vetter (2012). $\square$

Proof of Theorem 4.3: By employing the triangle inequality we can bound the Mallow 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

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

[where $\hat{D}_{T,a}^2$ and $\hat{D}_{T,a}^{2,*}$ are the random variables from Theorem 4.2 specified in the proof of which]. It follows from the proof of Theorem 6.1 that the first summand converges to zero and the second summand equals zero because of Theorem 4.2 a). So it suffices to treat the third summand which is bounded by

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

We obtain from Theorem 6.1 a) and b) that a constant $L > 0$ exists such that

$$\text{Var}(\hat{D}_{T,a}^{2,*}) \geq L N^{8d_\infty - 1} T^{-1} \tag{6.31}$$

[note the we are under the null hypothesis and that we therefore have $d_\infty = \underline{d}$]. This combined with Theorem 6.2, (6.30) 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

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

If we follow the proof of Theorem 6.1 a), b) and employ (6.28) and (6.30), we obtain that

$$\mathbb{E}((\hat{D}_T^{2,*})^2) = O(\log(N)N^{\max(8\underline{d}-1,0)+2\alpha}T^{-1} + N^{8\underline{d}+2\alpha-2}) \tag{6.32}$$

holds for every fixed $\alpha > 0$. By employing (6.27) and similar arguments as in the proof of Theorem 6.2 we obtain thereafter

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

The assertion then follows with (6.30)–(6.32) and the assumptions on the growth rate of $p = p(T)$. $\square$

7 Appendix: Auxiliary Lemmas

Finally we show some lemmas which were employed in the above proofs.

Lemma 7.1. *Suppose $\mu, \nu, a, b \in \mathbb{R}$. Then there exists a constant $C \in \mathbb{R}$ such that the following holds:*

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

$$\sum_{\substack{k=0 \\ k-a \geq 1 \\ -k+b \geq 1}}^{N-1} \frac{1}{(k-a)^{1-\mu}} \frac{1}{(b-k)^{1-\nu}} \leq \sum_{k=1+a}^{b-1} \frac{1}{(k-a)^{1-\mu}} \frac{1}{(b-k)^{1-\nu}} \leq \frac{C}{(b-a)^{1-\mu-\nu}}. \tag{7.1}$$

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

$$\sum_{\substack{k=1 \\ k+b \geq 1 \\ k-a \geq 1}}^{N-1} \frac{1}{(k+b)^{1-\mu}} \frac{1}{(k-a)^{1-\nu}} \leq \sum_{\substack{k=1 \\ k+b \geq 1 \\ k-a \geq 1}}^{\infty} \frac{1}{(k+b)^{1-\mu}} \frac{1}{(k-a)^{1-\nu}} \leq \frac{C}{|a+b|^{1-\mu-\nu}}. \tag{7.2}$$

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

$$\sum_{k=1+y}^{\infty} \frac{\log(k)}{k^{1-\mu}} \frac{1}{(k-y)^{1-\nu}} \leq C \frac{\log(y)}{y^{1-\mu-\nu}}, \tag{7.3}$$

$$\sum_{k=1}^{\infty} \frac{\log(k+z)}{(k+z)^{1-\mu}} \frac{1}{k^{1-\nu}} \leq C \frac{\log(z)}{z^{1-\mu-\nu}}. \tag{7.4}$$

Proof: a) Using equation 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 (7.2) by

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

$$\leq \int_{a}^{\infty} \frac{1}{(x+b)^{1-\mu}} \frac{1}{(x-a)^{1-\nu}} \, dx \lesssim \frac{1}{(a+b)^{1-\mu-\nu}}.$$

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

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

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

36

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

c) We start with (7.3). Using equation 13.2(18) in Erdelyi (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 (7.4) we use equation 6.4(23) in Erdelyi (1954a) which implies

$$\sum_{k=1}^{\infty} \frac{\log(k+z)}{(k+z)^{1-\mu}} \frac{1}{k^{1-\nu}} \leq \int_{0}^{\infty} \frac{\log(x+z)}{(x+z)^{1-\mu}} \frac{1}{x^{1-\nu}} \, dx \lesssim \frac{\log(z)}{z^{1-\mu-\nu}}.$$

$\square$

Lemma 7.2. *If* $0 < d_\infty < \frac{1}{4}$, *then*

$$\frac{1}{M^2 N^4} \sum_{\substack{j_1,j_2=1 \\ j_1 \neq j_2}}^{M} \sum_{r_1,s_1=0}^{N-1} \sum_{\substack{v_1,w_1,x_1,y_1=0 \\ v_1-w_1+r_1-s_1=0 \\ 0 \leq |v_1-w_1| \leq N-1}}^{\infty} \sum_{\substack{v_2,w_2,x_2,y_2=0 \\ r_1+v_2-x_1-s_1-w_2+y_1-y_2+x_2=0 \\ 0 \leq r_1+v_2-x_1+(j_1-j_2)N \leq N-1 \\ 0 \leq s_1+w_2-y_1+(j_1-j_2)N \leq N-1 \\ 0 \leq |y_2-x_2| \leq N-1}}^{\infty}$$

$$\psi_{v_1}(u_{j_1})\psi_{w_1}(u_{j_1})\psi_{x_1}(u_{j_1})\psi_{y_1}(u_{j_1})\psi_{v_2}(u_{j_2})\psi_{w_2}(u_{j_2})\psi_{x_2}(u_{j_2})\psi_{y_2}(u_{j_2})$$

$$(N^2 - N|y_2-x_2| - N|v_1-w_1| + |v_1-w_1||y_2-x_2|) = O\Big(\frac{\log(N)}{N^{1-8d_\infty T}}\Big).$$

Proof: Firstly, 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 $p \in \{0, 1, 2 \ldots, N-1\}$, 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 $1 \leq |C_1| \leq M-1$

$$\frac{1}{MN^4} \sum_{r_1,s_1=0}^{N-1} \sum_{\substack{v_1,w_1,x_1,y_1=0 \\ v_1-w_1+r_1-s_1=0 \\ 0 \leq |v_1-w_1| \leq N-1}}^{\infty} \sum_{\substack{v_2,w_2,x_2,y_2=0 \\ r_1+v_2-x_1-s_1-w_2+y_1-y_2+x_2=0 \\ 0 \leq r_1+v_2-x_1+C_1 N \leq N-1 \\ 0 \leq s_1+w_2-y_1+C_1 N \leq N-1 \\ 0 \leq |y_2-x_2| \leq N-1}}^{\infty}$$

$$\psi_{v_1}(u_{j_1})\psi_{w_1}(u_{j_1})\psi_{x_1}(u_{j_1})\psi_{y_1}(u_{j_1})\psi_{v_2}(u_{j_2})\psi_{w_2}(u_{j_2})\psi_{x_2}(u_{j_2})\psi_{y_2}(u_{j_2})$$

$$(N^2 - N|y_2-x_2| - N|v_1-w_1| + |v_1-w_1||y_2-x_2|)$$

$$\overset{(2.4)}{\lesssim} \frac{1}{MN^2} \sum_{p,r_1,s_1=0}^{N-1} \sum_{\substack{v_1,x_1,y_1=1 \\ 1 \leq r_1-s_1+v_1}}^{\infty} \sum_{\substack{v_2,w_2,x_2,y_2=1 \\ 1 \leq r_1-p+v_2+C_1 N \\ 1 \leq s_1-p+w_2+y_2-y_1+C_1 N}}^{\infty} \frac{1}{v_1^{1-d_\infty}} \frac{1}{y_1^{1-d_\infty}} \frac{1}{v_2^{1-d_\infty}} \frac{1}{w_2^{1-d_\infty}} \frac{1}{y_2^{1-d_\infty}}$$

$$\frac{1}{(r_1-s_1+v_1)^{1-d_\infty}} \frac{1}{(r_1-p+v_2+C_1 N)^{1-d_\infty}} \frac{1}{(s_1-p+w_2+y_2-y_1+C_1 N)^{1-d_\infty}}$$

$$\overset{(7.1)}{\lesssim} \frac{1}{MN^2} \sum_{p,r_1=0}^{N-1} \sum_{v_1,x_1,y_1=1}^{\infty} \sum_{\substack{v_2,w_2,x_2,y_2=1 \\ 1 \leq r_1-p+v_2+C_1 N \\ 1 \leq r_1-p+w_2-y_1+y_2+v_1+C_1 N}}^{\infty} \frac{1}{v_1^{1-d_\infty}} \frac{1}{y_1^{1-d_\infty}} \frac{1}{v_2^{1-d_\infty}} \frac{1}{w_2^{1-d_\infty}} \frac{1}{y_2^{1-d_\infty}}$$

37

$$\frac{1}{(r_1 - p + v_2 + C_1 N)^{1-d_\infty}} \frac{1}{(r_1 - p + w_2 - y_1 + y_2 + v_1 + C_1 N)^{1-2d_\infty}}$$

$$\overset{(7.1),(7.2)}{\lesssim} \frac{1}{MN^2} \sum_{p,r_1=0}^{N-1} \frac{1}{|r_1 - p + C_1 N|^{2-8d_\infty}} \lesssim \frac{1}{MN^2} \sum_{p,r_1=0}^{N-1} \frac{1}{(r_1 - p + N)^{2-8d_\infty}}$$

$$\lesssim \frac{\log(N)}{MN^{2-8d_\infty}} = \frac{\log(N)}{N^{1-8d_\infty}T}.$$

$\square$