Estimation and Testing in a Perturbed Multivariate Long Memory Framework

Vivien Less[a] Philipp Sibbertsen[a*]

[a]Wirtschaftswissenschaftliche Fakultät, Leibniz University Hannover, Hannover, Germany

December 12, 2022

Abstract

We propose a semiparametric multivariate estimator and a multivariate score-type testing procedure under a perturbed multivariate fractional process. The estimator is based on the periodogram and uses a local Whittle criterion function which is generalised by an additional constant to capture the perturbation given in the long memory process. Explicitly addressing the noise term when approximating the spectral density near the origin results in a bias reduction, but at the cost of an increase in the asymptotic variance of the estimator. Further, we introduce a multivariate testing procedure to detect spurious long memory under a perturbed fractional framework. The test statistic is based on the weighted sum of the partial derivatives of the multivariate local Whittle with noise estimator. We show consistency of the test against the alternatives of smooth trend and random level shift processes. In addition, we prove consistency and asymptotic normality of the local Whittle estimator and we derive the limiting distribution of the test. An empirical example on the squared returns and the realised volatilities from the BEL 20, S&P BSE SENSEX, and the Spanish IBEX is conducted, and shows the usefulness of the procedures.

Keywords: Signal-plus-noise · Multivariate local Whittle · Perturbation · Spurious long memory · Semiparametric estimation · Stochastic volatility.

JEL classification:C12, C13, C32

1 Introduction

Analysing the behaviour of the volatility of financial returns is among the most relevant topics in the literature when dealing with financial time series. When using log squared returns as an approximation for the underlying volatility process of the time series they show a strong persistence, which could be an indicator that the volatility process possesses long-range behaviour. In order to estimate the long memory parameter d the literature suggests using semiparametric

*Corresponding Author. Leibniz University Hannover, School of Economics and Management, Institute of Statistics, Königsworther Platz 1, D-30167 Hannover, Germany. E-Mail: sibbertsen@statistik.uni-hannover.de. Phone: +49-511-762-3783

procedures such as the local Whittle estimator of Robinson (1995a) or the GPH estimator of Geweke and Porter-Hudak (1983). However, recent results find that the usual procedures suffer from a severe downward bias when applied to squared returns (e.g. Deo and Hurvich (2001) and Arteche (2004)). The reason for this might be an inappropriate approximation of the spectral density near the zero frequency. For example, Hurvich and Ray (2003), Hurvich et al. (2005), and Frederiksen et al. (2012) assume that the squared returns might be better represented by a perturbed fractional process instead of the usual plain fractional process. The idea comes from the theoretical perspective that the squared returns are an unbiased but inconsistent estimator for the latent volatility process. This has an influence on the asymptotic behaviour of the spectrum. They suggest directly accounting for the perturbation, while approximating the spectral density near the origin. The inclusion of the perturbation leads to a decrease in the bias, especially when the memory and the noise parameter are high. Typically, this is the case in empirical applications. In general, a couple of univariate noise robust semiparametric estimators have been developed. For example, Sun and Phillips (2003), Hurvich and Ray (2003), Hurvich et al. (2005), Arteche (2006), and Frederiksen et al. (2012) have proposed estimators which can handle perturbed fractional processes. To our knowledge, there exist no multivariate estimator which takes the perturbation into account.

However, the same patterns which indicate a long memory process could also be induced by so-called low frequency contaminations, for example structural breaks, leading to spurious long-memory behaviour. As it is crucial to know what kind of a nature a process is following testing for spurious long memory is an important topic. There already exist multiple testing procedures in order to distinguish true long memory from spurious long memory processes in an univariate framework. Some examples are given by Dolado et al. (2005), Shimotsu (2006), Ohanissian et al. (2008), Perron and Qu (2010), Qu (2011), Haldrup and Kruse (2014), and Davidson and Rambaccussing (2015). Unfortunately, the literature regarding multivariate testing procedures is rather sparse. Sibbertsen et al. (2018) introduce a multivariate test based on the score of the multivariate Gaussian semiparametric estimator (GSE) of Shimotsu (2007).

A typical empirical example when testing for spurious long memory is the log squared returns of stock market indices. There exists a huge literature in favour of a spurious long memory process when investigating these squared returns (e.g. Granger and Ding (1996), Granger and Hyung (2004), Dolado et al. (2005), Lu and Perron (2010), Xu and Perron (2014), and Varneskov and Perron (2018)). However, when applying the same testing procedure to the realised variance the null hypothesis of a true long memory process cannot be rejected (see Qu (2011) and Sibbertsen et al. (2018)). Since both should be an unbiased estimator for the underlying volatility process, these contradictory results are of importance. We assume that those contradictions might arise from an imprecise approximation of the spectral density of the squared returns process near the origin of the spectrum. This presumption is coming from the fact that the properties of semiparametric estimators heavily depend on an accurate approximation of the periodogram to the local spectrum. The testing procedures of Qu (2011) and Sibbertsen et al. (2018) depend on the weighted sum of the partial derivatives of the local Whittle log likelihood and, therefore, on the appropriate approximation of the spectrum. As a consequence, this can lead to a false rejection of the null as the testing procedures heavily depend on the accuracy of the approximation of the periodogram.

All in all, this gives us the motivation to propose, on the one hand, a multivariate local Whittle estimator which explicitly accounts for the additional source of noise given in the log squared returns. It is a generalisation of the GSE of Shimotsu (2007) and a multivariate extension of the local Whittle with noise (LWN) estimator of Hurvich and Ray (2003) and Hurvich et al. (2005). On the other hand, we generalise the multivariate score-type test introduced by Sibbertsen et al. (2018) to make it robust against perturbation which might be given, for example, in the daily log squared returns.

Based on recent findings, we consider the estimation of the long memory parameter in a perturbed multivariate fractional process $Z_t = (z_{1t}, ..., z_{qt})'$, i.e. a multivariate signal-plus-noise process of the form

$$Z_t = Y_t + W_t. \tag{1}$$

The signal process $Y_t = (y_{1t}, ..., y_{qt})'$ is assumed to be a real-valued covariance stationary q-vector process which is able to capture long-range behaviour. The signal process is perturbed by a short-memory q-vector noise process $W_t = (w_{1t}, ..., w_{qt})'$. We assume independence of the signal and the noise processes.

Another motivation for considering this signal-plus-noise process is the special case of a long memory stochastic volatility model (LMSV) used for financial returns. A multivariate extension of the LMSV model of Breidt et al. (1998) is given by So and Kwok (2006), where the returns $r_t = (r_{1t}, ..., r_{qt})'$ are modelled as

$$r_{it} = \kappa_i e^{y_{it}/2} u_{it}, \ i = 1, ..., q,$$

such that taking a logarithmic transformation of the squared return time series, with $\log r_{it}^2 = y_{it} + \log \kappa_i^2 + \log u_{it}^2$ leads to the signal-plus-noise time series with y_{it} being the multivariate long memory component of the volatility process with memory parameters $0 < d_1, ..., d_q < 1/2$, and the noise is given by $w_{it} = \log \kappa_i^2 + \log u_{it}^2$.

We aim to analyse the behaviour of Eq. (1) in the frequency domain so that we have the spectral density of the signal process as $f_y(\lambda)$ and let $f_w(\lambda)$ represent the spectral density of the noise process. Under the assumption of independence between the signal and the noise process, the spectrum of Z_t is given by $f_z(\lambda) = f_y(\lambda) + f_w(\lambda)$. This independence assumption excludes the possibility of allowing for a leverage effect in the return series. However, it can be easily generalised towards it.

Assume that the spectral density of a multivariate short memory process u satisfies the local condition $f_u(\lambda) \sim G$, as $\lambda \to 0$, with G being a real, symmetric, finite, and positive definite matrix; Shimotsu (2007) suggests locally approximating the spectral density for non-perturbed multivariate fractional processes by

$$f_y(\lambda) \sim \mathrm{diag}(\lambda^{-d_a} e^{i(\pi-\lambda)d_a/2}) G \mathrm{diag}(\lambda^{-d_a} e^{-i(\pi-\lambda)d_a/2})[1 + O(\lambda^{\min\{\beta_y, 2\}})], \ \lambda \to 0. \tag{2}$$

In contrast, we suggest the following generalisation of the local approximation

$$f_z(\lambda) \sim \text{diag}(\lambda^{-d_a} e^{i(\pi-\lambda)d_a/2}(1+\theta_a\lambda^{2d_a})^{1/2}) G \text{diag}(\lambda^{-d_a} e^{-i(\pi-\lambda)d_a/2}(1+\theta_a\lambda^{2d_a})^{1/2}) \times [1 + O(\lambda^{\min\{\beta_y,2\}}) + \lambda^{2d_a} O(\lambda^{\min\{\beta_w,2\}})], \ \lambda \to 0, \quad (3)$$

in order to capture the behaviour of a perturbed fractional process. We propose to model the noise component in Z_t by an additional constant $\theta = (\theta_a, ..., \theta_q)'$. Note that the assumption on G rules out the possibility of cointegration.

The rest of the paper is structured as follows: Section 2 introduces the multivariate local Whittle with noise estimator and its asymptotic properties. In Section 3, we propose a modified score-type test. In Section 4, we investigate the finite sample performance in an extensive Monte Carlo study of the estimator and the test compared to the procedures of Shimotsu (2007) and Sibbertsen et al. (2018). An empirical application is presented in Section 5. Section 6 concludes. The proofs of our theorems can be found in the Appendix.

2 Local Whittle estimation under perturbation

In this section we introduce the MLWN estimator and present the necessary assumptions in order to establish asymptotic normality and consistency under a perturbed fractional process. We present the main results in two theorems.

2.1 MLWN estimator

We are interested in the semiparametric estimation of the memory parameter $d = (d_1, ..., d_q)'$ in a perturbed fractional setup as given in Eq. (1) by using only the Fourier frequencies in the vicinity of the origin. This results in an estimator which is non-parametric with respect to the short-run dynamics of the time series. Hence, the local Whittle estimator enjoys robustness towards them.

The multivariate GSE of Shimotsu (2007) generalises the univariate estimation method of Robinson (1995a). The log likelihood is formulated with respect to the assumption of the behaviour of the spectral density given in Eq. (2), yielding the following objective function

$$R(d) = \log \det \hat{G}(d) - \frac{2}{m}\sum_{a=1}^{q} d_a \sum_{j=1}^{m} \log \lambda_j,$$

$$\hat{G}(d) = \frac{1}{m}\sum_{j=1}^{m} \text{Re}\left[\Lambda_j(d)^{-1} I_z(\lambda_j) \Lambda_j^*(d)^{-1}\right],$$

where $m = m(n)$ equals the bandwidth with $m(n) \to \infty$ and $m/n \to 0$, $I_z(\lambda) = w_z(\lambda) w_z^*(\lambda)$ is the periodogram of Z_t and $w_z(\lambda) = (2\pi n)^{-1/2} \sum_{t=1}^{n} Z_t e^{it\lambda}$ is defined as the discrete Fourier transform of the Fourier frequencies $\lambda_j = 2\pi j/n$. The conjugate transpose is defined by x^*. Under this notation, we get for Eq. (2) $f_y(\lambda) \sim \Lambda_j(d) G \Lambda_j^*(d)$, $\Lambda_j(d) = \text{diag}(\Lambda_{ja}(d))$, so that $\Lambda_{ja}(d) = \lambda_j^{-d_a} e^{i(\pi-\lambda_j)d_a/2}$, as $\lambda_j \to 0$.

The estimator is given as the minimiser of the local Whittle log likelihood contrast function

$$\hat{d} = \arg\min_{d} R(d).$$

Shimotsu (2007) shows asymptotic normality and consistency of the estimator over the space of the true parameter values $d^0 \in [\Delta_1, \Delta_2]^q$, with $-1/2 < \Delta_1 < \Delta_2 < 1/2$. Nielsen (2011) generalises the results of Shimotsu (2007), and shows that the estimator is still consistent and asymptotically normal for $d \in (-1/2, \infty)$ with

$$m^{1/2}(\hat{d}_{\mathrm{GSE}} - d^0) \xrightarrow{d} \mathrm{N}(0, \Omega^{-1}), \qquad \Omega = 2\left[G^0 \odot (G^0)^{-1} + I_q + \frac{\pi^2}{4}(G^0 \odot (G^0)^{-1} - I_q)\right],$$
$$\hat{G}(\hat{d}) \xrightarrow{p} G^0,$$

with $\odot$ denoting the Hadamard product. Under a perturbed fractional process, we assume that the spectral density is better described by Eq. (3) rather than Eq. (2). Therefore, we propose to redefine $\Lambda_j(d) = \mathrm{diag}(\Lambda_{ja}(d))$ as follows

$$\Lambda_{ja}(d, \theta) = (\lambda_j^{-d_a} e^{i(\pi - \lambda_j)d_a/2}(1 + \theta_a \lambda_j^{2d_a})^{1/2}),$$

where $\theta_a = f_{a,w}(0)/f_{a,y}(0)$ is defined as the long-run noise-to-signal ratio, capturing the perturbation near the zero frequency. We assume that $\theta_a > 0$ for $a = 1, ..., q$.
It follows that the Gaussian log likelihood function near the origin is given by

$$Q(G, d, \theta) = \frac{1}{m}\sum_{j=1}^{m}\left\{\log\det\Lambda_j(d,\theta)G\Lambda_j^*(d,\theta) + \mathrm{tr}\left[(\Lambda_j(d,\theta)G\Lambda_j^*(d,\theta))^{-1}I(\lambda_j)\right]\right\}$$
$$= \frac{1}{m}\sum_{j=1}^{m}\left\{\log\det\Lambda_j(d,\theta)G\Lambda_j^*(d,\theta) + \mathrm{tr}\left[G^{-1}\mathrm{Re}\left[\Lambda_j(d,\theta)^{-1}I_z(\lambda_j)\Lambda_j^*(d,\theta)^{-1}\right]\right]\right\},$$

the second line follows from $Q(G, d, \theta)$ and G both being real. By the same argument as in Shimotsu (2007), we get

$$G = \frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\left[\Lambda_j(d,\theta)^{-1}I_z(\lambda_j)\Lambda_j^*(d,\theta)^{-1}\right].$$

Substituting G into $Q(G, d, \theta)$ together with

$$\log\det\Lambda_j(d,\theta) + \log\det\Lambda_j^*(d,\theta) = \log\det\Lambda_j(d,\theta)\Lambda_j^*(d,\theta) = -2\sum_{a=1}^{q} d_a \log\lambda_j + \sum_{a=1}^{q}\log(1 + \theta_a\lambda_j^{2d_a}),$$

yields the following objective function

$$R(d,\theta) = \log\det\hat{G}(d,\theta) - \frac{2}{m}\sum_{a=1}^{q} d_a \sum_{j=1}^{m}\log\lambda_j + \sum_{a=1}^{q}\sum_{j=1}^{m}\log(1 + \theta_a\lambda_j^{2d_a}),$$
$$\hat{G}(d,\theta) = \frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\left[\Lambda_j(d,\theta)^{-1}I_z(\lambda_j)\Lambda_j^*(d,\theta)^{-1}\right].$$

In the following, we will denote the true parameter values by a zero in the superscript. Our MLWN estimator is defined as

$$(\hat{d},\hat{\theta}) = \underset{(d,\theta)\in D\times\Theta}{\arg\min}\ R(d,\theta),$$

over the parameter space $D = [\Delta_1, \Delta_2]^q$, with $0 < \Delta_1 < \Delta_2 < 1/2$ and Θ is a compact and convex set in $\mathbb{R}^q$.

2.2 Asymptotic properties

Following Frederiksen et al. (2012) for the sake of conciseness, we give only one set of assumptions for the proof of consistency and asymptotic normality for the MLWN estimator. For the proof of the MLWNS test we need the same set of assumptions as stated here. However, the assumptions for consistency could be further relaxed (see Hurvich et al. (2005) and Shimotsu (2007)). In the following, $f_{ab}(\lambda)$ and G^0_{ab} represent the (a,b)-th element of $F(\lambda)$ and G^0.

Assumption 1. *The noise process $W_t = (w_{1t},...,w_{qt})'$ and the signal process $Y_t = (y_{1t},...,y_{qt})'$ are independent.*

Assumption 2. *The spectral density of $Z_t = (z_{1t},...,z_{qt})'$ satisfies the smoothness condition $f_{z,ab}(\lambda) - e^{i(\pi-\lambda)(d^0_a-d^0_b)/2}\lambda^{-d^0_a-d^0_b}(1+\theta^0_a\lambda_j^{2d^0_a})^{1/2}(1+\theta^0_b\lambda_j^{2d^0_b})^{1/2}G^0_{ab} = O(\lambda^{-d^0_a-d^0_b+\beta}) + O(\lambda^{-d^0_a+d^0_b})$ for $\beta\in(0,2]$ and $a,b = 1,...,q$.*

Assumption 3. *The signal has a linear representation $Y_t - EY_t = A(L)\epsilon_t = \sum_{j=0}^{\infty} A_j\epsilon_{t-j}$, where $\sum_{j=0}^{\infty}||A_j||^2 < \infty$, with $||\cdot||$ denoting the supremum norm and ϵ_t satisfies for $t = 0, \pm1, ...,$ $E(\epsilon_t|F_{t-1}) = 0$, $E(\epsilon_t\epsilon_t'|F_{t-1}) = I_q$ a.s. and for $a,b,c,d = 1,2$, $E(\epsilon_{at}\epsilon_{bt}\epsilon_{ct}|F_{t-1}) = \mu_{abc} < \infty$ a.s., $E(\epsilon_{at}\epsilon_{bt}\epsilon_{ct}\epsilon_{dt}|F_{t-1}) = \mu_{abcd} < \infty$ a.s., where F_t is the σ-field generated by $\epsilon_s,\ s\le t$.*
There exists a scalar random variable ϵ with $E(\epsilon^2) < \infty$ such that $\forall\tau > 0$ and some $K > 0$, $P(||\epsilon_t||^2 > \tau) \le KP(\epsilon^2 > \tau)$.
For $A(\lambda) = \sum_{j=0}^{\infty} A_j e^{ij\lambda}$, the derivative in the neighbourhood $(0,\delta)$ of the origin satisfies $\frac{\partial}{\partial\lambda}A_a(\lambda) = O(\lambda^{-1}||A_a(\lambda)||)$ as $\lambda \longrightarrow 0+$, where $A_a(\lambda)$ is the a-th row of $A(\lambda)$.

Assumption 4. *The noise term has a linear representation $W_t - EW_t = B(L)\xi_t = \sum_{j=0}^{\infty} B_j\xi_{t-j}$, where $\sum_{j=0}^{\infty}||B_j||^2 < \infty$. Further ξ_t satisfies for $t = 0,\pm1,...,$ $E(\xi_t|F_{t-1}) = 0$, $E(\xi_t\xi_t'|F_{t-1}) = I_q$ a.s. and for $a,b,c,d = 1,2$, $E(\xi_{at}\xi_{bt}\xi_{ct}|F_{t-1}) = \mu_{abc} < \infty$ a.s., $E(\xi_{at}\xi_{bt}\xi_{ct}\xi_{dt}|F_{t-1}) = \mu_{abcd} < \infty$ a.s., where F_t is the σ-field generated by $\xi_s,\ s\le t$.*
There exists a scalar random variable ξ with $E(\xi^2) < \infty$ such that $\forall\tau > 0$ and some $K > 0$, $P(||\xi_t||^2 > \tau) \le KP(\xi^2 > \tau)$.
For $A(\lambda) = \sum_{j=0}^{\infty} A_j e^{ij\lambda}$, the derivative in the neighbourhood $(0,\delta)$ of the origin satisfies $\frac{\partial}{\partial\lambda}A_a(\lambda) = O(\lambda^{-1}||A_a(\lambda)||)$ as $\lambda \longrightarrow 0+$, where $B_a(\lambda)$ is the a-th row of $B(\lambda)$.

Assumption 5. *The bandwidth m fulfils $\lim_{n\to\infty}(m^{-4max(d)-1}n^{4max(d)} + n^{-2\beta}m^{2\beta+1}log^2(m) + log(n)m^{-\gamma}) = 0$ for any $\gamma > 0$.*

Assumption 6. *There exists a finite real matrix L for which $\Lambda_j(d^0,\theta^0)^{-1}A(\lambda_j) = L + o(1)$, as $\lambda_j \longrightarrow 0$.*

Assumption 1 makes it possible to represent the spectral density of a signal-plus-noise process as the sum of the individual noise and signal components of the spectral density. Assumption 2 introduces a smoothness condition and is similar to the ones introduced by Robinson (1995b) and Hurvich et al. (2005). Assumptions 3 and 4 are multivariate extensions of the assumptions made in Hurvich et al. (2005) and Shimotsu (2007). The assumption on the bandwidth is given in Assumption 5 and is slightly stronger than the assumption given in Hurvich et al. (2005). The additional term is necessary in order to guarantee the convergence of the Hessian. Our last assumption is the same as in Shimotsu (2007).
Based on these assumptions, we are ready to establish the consistency of $\hat{d}$.

Theorem 1. *Let Assumptions 1–6 hold. Then* $\hat{d} \xrightarrow{p} d^0$ *as* $n \longrightarrow \infty$.

Note that our theorem proves consistency for the estimator of the memory parameter only. We do not prove the consistency for the estimator of the nsr. Our proof orientates on the method of proof used by Andrews and Sun (2004) and Frederiksen et al. (2012) which circumvents a separate consistency proof of $\hat{\theta}$ and rather includes it in the proof of asymptotic normality. Hence, we can proceed with a joint asymptotic normality and consistency result for $\hat{d}$ and $\hat{\theta}$.

Theorem 2. *Let Assumptions 1–6 hold. Then, for* d^0 *in the interior of* $D = [\Delta_1, \Delta_2]^q$, *where* $0 < \Delta_1 < \Delta_2 < 1/2$, $\hat{d}$ *and* $\hat{\theta}$ *are both consistent and*

$$B_n \begin{pmatrix} \hat{d} - d^0 \\ \hat{\theta} - \theta^0 \end{pmatrix} \xrightarrow{d} N(0, \mathbf{\Omega}^{-1}),$$

$$\mathbf{\Omega} = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix},$$

where $B_n = diag(m^{1/2}, m^{1/2}\lambda_m^{2d_a})$,

$$\Omega_{11} = 2\left[G^0 \odot (G^0)^{-1} + I_q + \frac{\pi^2}{4}(G^0 \odot (G^0)^{-1} - I_q)\right],$$

$$\Omega_{12} = \Omega_{21} = \omega_1 \left[G^0 \odot (G^0)^{-1} + I_q\right] \omega_1 + 2\left[\frac{\pi^2}{4}(G^0 \odot (G^0)^{-1} - I_q)\right],$$

$$\Omega_{22} = 2\left(\omega_2 \left[G^0 \odot (G^0)^{-1} + I_q\right] \omega_2\right),$$

with ω *being some* $q \times q$ *diagonal matrix depending on* d^0, *where the a-th element is given for* $\omega_{1,a}$ *by* $\frac{-d_a^0}{1+2d_a^0}$ *and for* $\omega_{2,a}$ *by* $\frac{(d_a^0)^2}{((1+2d_a^0)(1+4d_a^0))^{1/2}}$.

The first thing to notice is that the submatrix Ω_{11} is exactly the same as the asymptotic variance of the GSE from Shimotsu (2007). Secondly, the asymptotic variance of the estimation procedure does not depend on the unknown nsr values. However, introducing an additional constant in the estimation procedure results in a bias reduction but inflates the asymptotic variance. Further, the asymptotic variance depends on the true memory parameter d^0 and decreases as d^0 increases as our signal gets stronger and easier to estimate.
We show asymptotic normality and consistency of the MLWN estimator in the stationary case only. However, we expect the estimator to be still consistent and asymptotic normal in the non-stationary region by the same arguments as in Nielsen (2011) and Frederiksen et al. (2012).

3 A testing procedure under perturbation

In this section we propose a perturbation robust score-type test against spurious long memory. We derive the limiting distribution and show consistency of the procedure. The main results are stated in two theorems.

We propose testing the null of a true perturbed long memory process or, equivalently, we assume under the null that the spectral density follows the local approximation defined in Eq. (3). The alternative states that the process contains low frequency contaminations. The hypotheses can be stated as follows

$$H_0 : f_z(\lambda_j) \sim \Lambda_j(d,\theta)G\Lambda_j^*(d,\theta) \qquad \text{vs.} \qquad H_1 : f_z(\lambda_j) \nsim \Lambda_j(d,\theta)G\Lambda_j^*(d,\theta).$$

Our testing procedure is based on the score of the MLWN estimator presented already. Following Sibbertsen et al. (2018), the test statistic is based on an approximation of the first derivative of the local Whittle objective function. It is given by:

$$\begin{aligned} MLWNS =& \frac{1}{2} \sup_{r\in[\epsilon,1]} \Bigg|\Bigg| \frac{2}{\sqrt{m}} \sum_{a=1}^{q} \frac{1}{\sqrt{q}} \sum_{j=1}^{[mr]} \zeta_{ja} \left(g(\hat{d},\hat{\theta})_a \left\{ \mathrm{Re}\left[\Lambda(\hat{d},\hat{\theta})^{-1} I_{zj} \Lambda^*(\hat{d},\hat{\theta})^{-1} \right]_a \right\} - 1 \right) \zeta_{ja} \\ &+ \frac{1}{\sqrt{m}} \sum_{a=1}^{q} \frac{1}{\sqrt{q}} g(\hat{d},\hat{\theta})_a \sum_{j=1}^{[mr]} L_j \mathrm{Im}\left[\Lambda(\hat{d},\hat{\theta})^{-1} I_{zj} \Lambda^*(\hat{d},\hat{\theta})^{-1} \right]_a \Bigg|\Bigg|, \end{aligned} \tag{4}$$

where $\zeta_{ja} = (\widetilde{X}_{ja} - \frac{1}{m}\sum_{k=1}^{m} \widetilde{X}_{ka})$ with $\widetilde{X}_{ja} = (\log \lambda_j/(1+\theta_a \lambda_j^{2d_a})^{1/2}, -(j/m)^{d_a}/(2(1+\theta_a\lambda_j^{2d_a}))^{1/2})'$ and $L_j = ((\lambda_j - \pi)/2, 0)'$.

Our test statistic includes in the absence of noise as special cases the testing procedures of Sibbertsen et al. (2018) and in the univariate case the one proposed by Qu (2011). It is also possible to use the perturbation robust test we are presenting in a univariate context.

The limiting distribution of the test statistic in Eq. (4) can be derived by using the same arguments as in Sibbertsen et al. (2018), generalised for the case of perturbation.

Let $B(s)$ be a Brownian motion defined on $[0,1]$ and "$\Longrightarrow$" symbolises weak convergence in the Skorokhod space, then our limiting distribution is given by

Theorem 3. *Given our Assumptions 1–6 hold we have for $n \to \infty$*

$$MLWNS \Longrightarrow \sup_{r\in[\epsilon,1]} \left|\left| \int_0^r \mathbf{\Psi}^0(r)dB(s) - B(1)\int_0^r \mathbf{\Psi}^0(r)ds - F(r)\int_0^1 \mathbf{\Psi}^0(r)dB(s) \right|\right|,$$

with $F(r) = \int_0^r \mathbf{\Psi}^{2,0}(r)ds$ and $\mathbf{\Psi}^0(r) = (\Psi_1^0(r), \Psi_2^0(r))'$, where the single entries are given as

$$\Psi_{1,a}^0 = \int_0^r \frac{\log s}{1+\theta_a^0 s^{2d_a^0}} ds - r\int_0^1 \frac{\log s}{1+\theta_a^0 s^{2d_a^0}} ds,$$
$$\Psi_{2,a}^0 = \int_0^r \frac{s^{2d_a^0}}{1+\theta_a^0 s^{2d_a^0}} ds - r\int_0^1 \frac{s^{2d_a^0}}{1+\theta_a^0 s^{2d_a^0}} ds.$$

The limiting distribution of our test statistic is not pivotal because it depends on the unknown values of the memory parameter d^0 and the noise-to-signal ratio θ^0. The critical values of the limiting distribution are obtained numerically for different values of the long memory parameter and the noise-to-signal ratio.

Under the alternative of low frequency contaminations, we orientate for the data generating processes (DGPs) on the processes selected by Sibbertsen et al. (2018). The first one is a multivariate random level shift process defined as

$$X_t = \mu_t + \kappa_t \quad \text{with} \qquad (5)$$
$$\mu_t = (I_q - \phi\Pi_t)\mu_{t-1} + \Pi_t e_t.$$

We assume mutual independence between κ_t, $\Pi_t = \text{diag}(\pi_{1t}, ..., \pi_{qt})$ and e_t. Further, we define the correlation matrix Σ_π in order to allow the Bernoulli variables π_{it} and π_{jt} for $i, j = 1, ..., q$ to be correlated over the different q-dimensions of the process X_t. The probability of the occurrence of a shift in the time series is regulated by $p = \tilde{p}/n$, with $\tilde{p}$ being the expectation of the number of shifts that appear in the sample. The magnitude of those shifts is determined by $e_t \sim N(0, \Sigma_e)$, which is a q-column vector. Furthermore, we define the pairwise correlation coefficients of π_{it} and π_{jt}, e_{it} and e_{jt}, and u_{it} and u_{jt} by $\rho_{\pi,ij}$, $\rho_{e,ij}$, and $\rho_{u,ij}$, $\forall i, j = 1, ..., q$. The persistence of the process is controlled by the autoregressive coefficient ϕ so that we are able to investigate non-stationary as well as stationary processes.

Another possible DGP under the alternative is given by the smooth trend model as

$$X_t = H\left(\frac{t}{n}\right) + \kappa_t, \qquad (6)$$

where $H(t/n) = h_a(t/n)$ is a q-dimensional column vector with $a = 1, ..., q$ and $h_a(t/n)$ being a Lipschitz continuous function defined on $[0, 1]$. The noise κ_t is the same as in Eq. (5).

Based on the findings of Perron and Qu (2010) and McCloskey and Perron (2013) for the univariate and Sibbertsen et al. (2018) for the multivariate case, we know that the level shifts in the processes described in Eq. (5) and (6) only affect the periodogram up to $j = O(n^{1/2})$. These stochastic orders hold for the described processes, as well as for deterministic level shifts and fractional trends. The orders are exact under level shifts and approximate for slowly varying trends. The next theorem establishes consistency of the MLWNS test.

Theorem 4. *Assume that the DGP of X_t is either given by Eq. (5) or (6), further let $n \to \infty$ which gives $\frac{m}{n^{1/2}} \to \infty$, $P(\hat{d}_a - d_a^0 \geq 0) \to 1$ $\forall a \in \{1, ..., q\}$, with $\hat{G}(\hat{d}, \hat{\theta})$ being positive definite and Assumptions 1–6 hold will yield $MLWNS \xrightarrow{p} \infty$ as $n \to \infty$.*

Note that Theorem 4 does not require low frequency contamination in every single trajectory since the DGPs presented in Eq. (5) and Eq. (6) allow for low frequency contamination in the subvectors of X_t as well.

4 Finite sample results

In this section, we provide some simulation results in order to examine the finite sample performance of the MLWN estimator compared to the GSE estimator of Shimotsu (2007) in a multivariate perturbed fractional setup. Further, we compare the finite sample performance of

the MLWS test of Sibbertsen et al. (2018) with our noise robust alternative.
For the simulations, we use the following DGP in order to have a model in line with Eq. (1), e.g.:

$$Z_t = Y_t + W_t, \tag{7}$$

where we select the order of $q = 2$ so that $Z_t = (z_{1t}, z_{2t})'$ is a bivariate system. We model the multivariate signal process by using a VARFIMA(0,d,0) representation, where the long range behaviour of the time series is modelled by $(1 - L)^d$, where d gives the order of fractional integration. The relationship between the persistence of the shocks and the memory parameter can be expressed by the binomial expansion of $(1 - L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)} L^j$, with Γ being the gamma function and L given as the Backshift operator. The noise process is modelled by a multivariate white noise and has short memory. The noise-to-signal ratio is defined as $\text{nsr}_a = \frac{\sigma^2_{\kappa,a}}{\sigma^2_{u,a}}$. The Monte Carlo simulation conducts 10,000 replications of time series with 2,500 observations each, where we select for the nsr $\in \{5, 15\}$, which is chosen to reflect empirical findings of the volatility in the squared returns series. It should be noted that we used smaller and larger sample sizes as well, but we did not receive different results.
We want to analyse the behaviour of the estimator and the score-type test for different kinds of correlations ρ between the single signal processes, and choose $\rho \in \{0, 0.4, 0.8\}$. It follows that we have no, medium, and high dependence between the single elements of the time series.

	nsr=5, $\rho = 0$				nsr=15, $\rho = 0$			
d^0	MLWN		GSE		MLWN		GSE	
0.2	0.2240	(0.1488)	0.0637	(0.0341)	0.2142	(0.1728)	0.0251	(0.0342)
	0.2254	(0.1505)	0.0629	(0.0341)	0.2175	(0.1784)	0.0254	(0.0341)
0.3	0.3164	(0.1378)	0.1304	(0.0345)	0.3164	(0.1872)	0.0649	(0.0345)
	0.3151	(0.1342)	0.1307	(0.0350)	0.2997	(0.1831)	0.0652	(0.0347)
0.4	0.4092	(0.1170)	0.2185	(0.0356)	0.4078	(0.1693)	0.1291	(0.0363)
	0.4074	(0.1169)	0.2185	(0.0355)	0.3908	(0.1632)	0.1300	(0.0359)

Table 1: Simulation results applied to a multivariate perturbed ARFIMA process. The single trajectories have the same degree of persistence. The standard deviation is given in parentheses

These parameter settings are the same for all of the different kinds of simulations if not stated otherwise.
To investigate the performance of the estimators and the tests, we select $d_a \in \{0.2, 0.3, 0.4\}$ and allow for the same degree of memory in the time series. To examine the performance when the memories are different, we select $d = (0.4, 0.2)'$, as it is done by Shimotsu (2007). The bandwidth is set to be $m = n^{0.7}$ to guarantee comparability between the two estimators. However, it is possible to select a higher bandwidth for the MLWN estimator, e.g. $m = n^{0.8}$, reasoned by accounting for the additional noise term and under sufficient smoothness of the spectrum of the noise process near the origin. In unreported simulation, we find that selecting $m = n^{0.8}$ leads to a significant improvement of the MLWN estimator and we recommend using this in practice.
The memory estimates of the GSE and the MLWN estimator for a bivariate uncorrelated time series where each component obtains the same memory are stored in Table 1. Focusing first on

$d^0 = (0.4, 0.2)'$						
	nsr=5			nsr=15		
	$\rho = 0$	$\rho = 0.4$	$\rho = 0.8$	$\rho = 0$	$\rho = 0.4$	$\rho = 0.8$
GSE	0.1964	0.2164	0.2076	0.1126	0.1293	0.1261
	(0.0293)	(0.0350)	(0.0334)	(0.0299)	(0.0352)	(0.0348)
GSE	0.0601	0.0595	0.0462	0.0243	0.0241	0.0195
	(0.0276)	(0.0336)	(0.0319)	(0.0280)	(0.0338)	(0.0338)
MLWN	0.4042	0.4066	0.4037	0.4087	0.4091	0.4066
	(0.1096)	(0.1222)	(0.1091)	(0.1621)	(0.1703)	(0.1629)
MLWN	0.2257	0.2187	0.1844	0.1775	0.1944	0.1820
	(0.1489)	(0.1450)	(0.1291)	(0.1731)	(0.1755)	(0.1775)

Table 2: Simulation results of the GSE and the MLWN estimator for different amounts of correlation and different degrees of persistence for each of the perturbed ARFIMA processes. The standard deviation is given in parentheses.

the results of the GSE, we immediately see the heavily downward biased estimates of the memory parameter. In the low perturbation setup we already receive a downward bias of around 50%. If we increase the disturbance, the bias gets even more severe. Since in practice, the nsr found in financial time series is typically high, this is a huge drawback when using the GSE.

Focusing now on the performance of the MLWN estimator, we see that adding a constant in order to explicitly account for the additional source of noise obtained in the series significantly increases the precision of the estimator. Only when the nsr and the memory parameter are low, do we have a slightly upward biased estimate of the memory parameter. However, this is of minor interest since financial time series are typically rather persistent, with the memory parameter being in the higher stationary region up to the boundary case of 0.5. Another explanation is that we use a non-optimal bandwidth in this setting. We find that the MLWN estimator improves if the memory parameter, the bandwidth, the nsr, and/or the sample size increases.

The results given in Table 2 show that our estimator achieves consistent estimates when the time series have differing memory parameters for the single components over different kinds of nsr. We also allow for different kinds of correlations between the individual components. This results in an improvement of the estimators reasoned by the additional coherence information. However, the GSE still suffers from a severe bias which cannot be eliminated by the additional information caused by the correlation. When we allow for the same degree of persistence the results do not change. We get more precise estimates as the correlation increases. Therefore, the respective results are omitted. These results confirm the already mentioned problems in the univariate setting for the multivariate case. Even the additional information that we receive through the cross-periodogram is not sufficient to offset the influence of the perturbation on the estimation of the memory parameter.

Next, we consider the finite sample properties of the MLWNS test compared to the MLWS test. In order to scrutinise the size of the different testing procedures under a true long memory with noise process we select the same values for d, the nsr, and ρ as previously. In addition, we set $d^0 = (0.49, 0.49)'$ in order to analyse the performance of our testing procedure for the boundary

nsr=5						
d^0	$\rho = 0$		$\rho = 0.4$		$\rho = 0.8$	
	MLWS	MLWNS	MLWS	MLWNS	MLWS	MLWNS
0.2	0.074	0.083	0.0823	0.0844	0.0812	0.0954
0.3	0.2335	0.1229	0.2323	0.1182	0.237	0.089
0.4	0.5253	0.078	0.5309	0.0816	0.5582	0.073
0.49	0.7564	0.0354	0.7676	0.0346	0.8107	0.0406
nsr=15						
0.2	0.0569	0.0636	0.059	0.0667	0.054	0.0742
0.3	0.1565	0.0742	0.1521	0.0688	0.1607	0.0752
0.4	0.495	0.0809	0.4894	0.0697	0.4765	0.0612
0.49	0.8358	0.0316	0.8385	0.028	0.8144	0.0241

Table 3: Size comparison of the MLWS and MLWNS tests for a perturbed bivariate ARFIMA process. The degree of persistence is the same for all trajectories. The significance level is selected to be $\alpha = 5\%$.

	size			power					
nsr=5									
	$\rho = 0$	$\rho = 0.4$	$\rho = 0.8$	$\rho = 0$		$\rho = 0.4$		$\rho = 0.8$	
				$\phi = 1$	$\phi = 0$	$\phi = 1$	$\phi = 0$	$\phi = 1$	$\phi = 0$
MLWS	0.2542	0.251	0.2604	0.9774	0.975	0.9776	0.9727	0.9753	0.9754
MLWNS	0.0206	0.024	0.0405	0.8337	0.8115	0.8533	0.8769	0.8312	0.876
nsr=15									
MLWS	0.2082	0.2187	0.2069	0.9836	0.9808	0.9797	0.9826	0.9815	0.9833
MLWNS	0.0567	0.06	0.0748	0.8363	0.7896	0.8416	0.8426	0.7979	0.8084

Table 4: Size and power comparison of the MLWS and MLWNS tests for a perturbed bivariate ARFIMA process with $D = (0.4, 0.2)'$. The significance level is selected to be $\alpha = 5\%$.

case. For the significance level we select 5%. It should be noted that the bandwidth is again set to $m = n^{0.7}$ to ensure comparability between both approaches.

Table 3 contains the size results of the MLWS and the MLWNS tests for different degrees of correlation given in the signal process and the same degree of memory in the bivariate system. As expected the MLWS test does not hold its nominal size any longer if the persistence is high. The only exception is given for a rather low memory parameter of 0.2. Hence, the testing procedure suffers from heavy size distortions when the time series do have long memory but are (heavily) perturbed.

However, for our testing procedure, the size improves with an increasing nsr and with increasing correlation. Only for the boundary case where $d^0 = (0.49, 0.49)'$ we see that the MLWNS is slightly undersized. When the nsr is low, our testing procedure tends to be slightly liberal.

When we allow for different persistences upon the individual elements, the results do not change (cf. Table 4). The MLWS test is not able to hold the size, while for the MLWNS test the size improves with increasing nsr and correlation again. However, for some reason our testing

	nsr=5				nsr=15			
d^0	$\phi = 1$		$\phi = 0$		$\phi = 1$		$\phi = 0$	
	MLWS	MLWNS	MLWS	MLWNS	MLWS	MLWNS	MLWS	MLWNS
0.2	0.9748	0.9108	0.9747	0.9142	0.9776	0.93	0.9803	0.9363
0.3	0.9724	0.8931	0.9717	0.893	0.9788	0.9271	0.9789	0.9293
0.4	0.9754	0.8228	0.9789	0.8352	0.9852	0.9207	0.987	0.9204
0.49	0.9788	0.687	0.9771	0.7176	0.9944	0.9084	0.9939	0.9137

Table 5: Power comparison of the MLWS and MLWNS tests for a stationary (left) and a non-stationary perturbed long memory level shift process for the same degree of persistence among the time series. The shifts occur simultaneously. The significance level is selected to be $\alpha = 5\%$.

procedure appears to be conservative under a low nsr.

The power properties of the testing procedures are analysed under a perturbed long memory random level shift process and a multivariate smooth trend model. For the random level shift process we use the DGP for the size simulation to get the long memory-plus-noise component $Z_t = (z_{1t}, ..., z_{qt})'$, while the multivariate random level shift process is given by μ_t. This gives us the following DGP for our power analysis

$$X_t = \mu_t + Z_t, \tag{8}$$

$$\mu_t = (I_q - \phi\Pi_t)\mu_{t-1} + \Pi_t e_t, \tag{9}$$

where we assume mutual independence between the single noise processes included in Z_t, $\Pi_t = diag(\pi_{1t}, ..., \pi_{qt})$ and e_t. The process μ_t is defined as in Eq. (5). Hence, we allow for correlation between two signal processes in our DGP, but we do not allow for correlation in our noise component or correlation between the signal and noise components.

In order to scrutinise the power properties of our testing procedure for stationary as well as for non-stationary random level shift processes, we allow the autoregressive coefficient to be in the range of $(0, 1)$. As a result, ϕ represents the persistence of our multivariate system.

In our simulation study, we inspect the performance of our testing procedure with a shift probability fixed at $p = 5/n$, which gives five shifts in expectations with a standard deviation of $\sigma_e = 1$. The multivariate level shift process described in Eq. (8) and (9) is examined for $\phi = 1$, which represents a stationary and $\phi = 0$ a non-stationary system, respectively. In order to address the possibility of different kinds of information regarding the coherence reasoned by distinct behaviours of the breaks, we analyse a variety of values for ρ_π and ρ_e. For brevity, we only focus on $\rho_\pi = \rho_e$. As a result, we have the cases where the shifts occur independently in the single components of the system for $\rho_\pi = \rho_e = 0$, while for $\rho_\pi = \rho_e = 1$ the occurrence of the shifts will match in time and size.

For the multivariate smooth trend process we select, similar to Qu (2011), a non-monotonic deterministic trend for $H(t/n)$ and a white noise process for κ_t. All in all we get a multivariate perturbed long memory non-monotonic deterministic trend model of the form

$$X_t = Z_t + \sin(4\pi t/n) + \epsilon_t, \; \epsilon_t \overset{\text{iid}}{\sim} N(0, 1). \tag{10}$$

		nsr=5				nsr=15			
					$\rho_\pi = \rho_e = 1$				
					$\phi = 1$				
ρ	d^0	0.2	0.3	0.4	0.49	0.2	0.3	0.4	0.49
0	MLWS	0.9696	0.9659	0.9688	0.9689	0.9742	0.9780	0.9840	0.9922
0	MLWNS	0.9676	0.9545	0.9421	0.9457	0.9750	0.9719	0.9643	0.9701
0.4	MLWS	0.9681	0.9658	0.9659	0.9688	0.9733	0.9753	0.9820	0.9911
0.4	MLWNS	0.9697	0.9610	0.9546	0.9597	0.9750	0.9719	0.9697	0.9728
0.8	MLWS	0.9666	0.9664	0.9668	0.9742	0.9772	0.9737	0.9824	0.9916
0.8	MLWNS	0.9781	0.9740	0.9748	0.9760	0.9802	0.9768	0.9784	0.9814
					$\phi = 0$				
0	MLWS	0.9693	0.9662	0.9712	0.9725	0.9734	0.9761	0.9829	0.9929
0	MLWNS	0.9620	0.9508	0.9448	0.9513	0.9745	0.9695	0.9647	0.9678
0.4	MLWS	0.9685	0.9657	0.9708	0.9709	0.9724	0.9739	0.9826	0.9902
0.4	MLWNS	0.9709	0.9618	0.9626	0.9638	0.9743	0.9730	0.9707	0.9714
0.8	MLWS	0.9675	0.9678	0.9649	0.9737	0.9726	0.9733	0.9820	0.9903
0.8	MLWNS	0.9774	0.9733	0.9738	0.9758	0.9762	0.9752	0.9782	0.9842
					$\rho_\pi = \rho_e = 0$				
					$\phi = 1$				
		nsr=5				nsr=15			
ρ	d^0	0.2	0.3	0.4	0.49	0.2	0.3	0.4	0.49
0	MLWS	0.9964	0.9939	0.9915	0.9856	0.9973	0.9985	0.9988	0.9985
0	MLWNS	0.9576	0.9339	0.7383	0.9209	0.9670	0.9473	0.7672	0.9641
0.4	MLWS	0.9954	0.9949	0.9920	0.9872	0.9979	0.9980	0.9980	0.9981
0.4	MLWNS	0.9539	0.9287	0.7703	0.9244	0.9647	0.9465	0.7768	0.9663
0.8	MLWS	0.9958	0.9945	0.9946	0.9947	0.9968	0.9976	0.9985	0.9988
0.8	MLWNS	0.9653	0.9531	0.8583	0.9477	0.9680	0.9521	0.7963	0.9683
					$\phi = 0$				
0	MLWS	0.9963	0.9956	0.9933	0.9878	0.9973	0.9975	0.9979	0.9987
0	MLWNS	0.9713	0.9657	0.9546	0.9295	0.9789	0.9784	0.9814	0.9664
0.4	MLWS	0.9972	0.9947	0.9920	0.9899	0.9968	0.9982	0.9986	0.9993
0.4	MLWNS	0.9730	0.9652	0.9576	0.9324	0.9783	0.9753	0.9791	0.9685
0.8	MLWS	0.9976	0.9964	0.9966	0.9959	0.9969	0.9975	0.9982	0.9991
0.8	MLWNS	0.9732	0.9660	0.9578	0.9524	0.9765	0.9760	0.9773	0.9705

Table 6: Size-adjusted power comparison of the MLWS and MLWNS tests for a stationary and a non-stationary perturbed long memory level shift process for the same degree of persistence among the time series. The correlation ρ is given in the signal processes. The significance level is selected to be $\alpha = 5\%$.

Table 5 includes the results of the power simulation for the random level shift process where both trajectories share the same degree of persistence. We would already expect that the power of the MLWS test is higher compared to the MLWNS test. In general, we can see that the power in-

	nsr=5				nsr=15			
ρ \ d^0	0.2	0.3	0.4	0.49	0.2	0.3	0.4	0.49
0	0.9996	0.9991	0.9954	0.9921	0.9992	0.9996	0.9990	0.9965
0.4	0.9984	0.9986	0.9909	0.9792	0.9992	0.9996	0.9992	0.9983
0.8	0.9982	0.9953	0.9622	0.9423	0.9993	0.9997	0.9994	0.9975

Table 7: Size-adjusted power properties of the MLWNS test for a perturbed bivariate long memory deterministic smooth trend model. The memory parameter is the same for both trajectories. The significance level is selected to be $\alpha = 5\%$.

creases for increasing nsr and decreasing memory parameter values. This result is not surprising, since it gets harder to detect the level shift if the persistence of the process gets stronger. Also, we achieve better power results under a non-stationary process as well as for an increasing sample size. We find the same results when allowing for differing persistences in the single time series. In addition, we investigated the power properties of the testing procedures under a size-adjusted setting to ensure comparability among these. The results are given in Table 6. We are able to

	nsr=5				nsr=15			
ρ \ d^0	0.2	0.3	0.4	0.49	0.2	0.3	0.4	0.49
0	0.0755	0.2263	0.3921	0.4337	0.0566	0.1481	0.4339	0.5505
0.4	0.0800	0.2268	0.3988	0.4462	0.0580	0.1599	0.4307	0.5504
0.8	0.0846	0.2278	0.4166	0.5245	0.0582	0.1501	0.4203	0.5528

Table 8: Size results of the MLWS test after applying a pre-whitening procedure on the raw time series for different degrees of correlation. The memory parameter is the same among the single trajectories.

see a massive increase in the power results of the test. Now, the MLWNS test is very close to the performance of the MLWS test, giving a satisfactory power under the alternative. Apart from that, the table also consists of the size-adjusted power results for the case when $\rho_\pi = \rho_e = 0$, so that the shifts and the size of the shifts do not match. We observe that the power is slightly higher under a non-stationary random level shift process. The power increases if we allow for additional correlation in the signal processes, and hence increase the information given in the phase of the spectra. In addition, we observe a slightly higher power for processes with a lower memory parameter.

Further, we investigate in unreported simulations the performances of the tests when reducing the shift probability to three and to two shifts in expectation, as under a random level shift model the null can be true as well. Under this scenario we observe a decrease in the power of both testing procedures as the probability of observing the null gets higher with a decreasing shift probability.

The results of the size-adjusted power performance of the MLWNS test under a multivariate smooth trend model are given in Table 7. We observe a small power loss for an increase in the memory parameter and an increase in the correlation given in the long memory signal processes. Further, we see that the power is higher for the higher noise-to-signal ratios. These findings are due to the stronger signal compared to the low frequency contamination.

Next, we need to investigate the performance of the MLWS test in terms of size when applying a pre-whitening on the time series beforehand. Therefore, we follow the procedures of Qu (2011) and Sibbertsen et al. (2018) in order to account for the short run dynamics given in the time series of interest in finite samples. Instead of using a multivariate pre-whitening we are sticking to the univariate method as the multivariate fitting of a FIVARMA process suffers from convergence issues in the numerical optimization while doing the maximum likelihood estimation of the model in our setting, especially when the memory parameter and the correlation are high. As mentioned in Sibbertsen et al. (2018), the size is well controlled irrelevant of the used procedure. Hence, we can use the univariate method without any drawbacks. The simulation results are given in Table 8 for the MLWS test. We see an improvement in size especially for the higher nsr ratio. However, with an increasing persistence we still obtain a greatly size distorted testing procedure. As a result, the pre-whitening does not improve the performance of the MLWS test in such a way that it can be applied to a time series following a rather persistent perturbed fractional process.

5 Empirical example

We already know by the empirical applications of Qu (2011) and Sibbertsen et al. (2018), among others, that it is doubtful that the squared returns follow a true long memory process. In contrast, realised volatilities from the same stock or index seem to be generated by a true long memory process, as the null cannot be rejected. To shed some light on those contradictory results, we propose to re-investigate the daily log squared returns with a procedure that is able to accommodate for possible perturbation. Here, we are using the returns and realised volatilities of the BEL 20 (BFX), S&P BSE SENSEX (BSESN), and the Spanish IBEX (IBEX) starting in 2000 and ending in 2018 with a total of 4,687 observations. The data is obtained from the realised library of the Oxford-Man Institute of quantitative finance.
Our procedure, as it stands so far, is not able to handle fractionally cointegrated time series. As a consequence, our first step is to make sure that our analysed time series are not fractionally cointegrated. Therefore, we are using the rank estimation procedure of Nielsen and Shimotsu (2007) and the fractional cointegration test proposed by Robinson (2008). According to Leschinski et al. (2021) these procedures should be appropriate in our setup to detect possible fractional cointegrating relationships. Neither in the squared returns nor in the realised volatility are we able to detect any kind of fractional cointegrating relation. Hence, we can proceed with our multivariate procedures.
The results are given in Table 9. For all of the scenarios we use a bandwidth equal to $m = n^{0.7}$. We find that changing the bandwidth does not change the overall result. In general, we have slightly higher memory estimates in the realised volatility time series compared to the squared returns. They are close to the boundary case of 0.5 where the process gets non-stationary. However, the higher memory estimates are not surprising, as we have a closer approximation of the true underlying process when using tick data. When comparing the memory estimates of the GSE with the ones obtained by using the noise robust version, we see huge downward biased estimates for the squared return time series. This is what we expect if we face indeed a perturbed

	log squared returns			RV		
	BFX	BSESN	IBEX	BFX	BSESN	IBEX
MLWN	0.4989	0.4099	0.4249	0.5671	0.4237	0.4992
GSE	0.3948	0.3080	0.3311	0.5071	0.2779	0.4458
MLWNS	1.14628			1.8783		
MLWS	1.4458**			0.5448		

Table 9: Memory estimates of the GSE and the MLWN estimator and test statistic values of the MLWS and the MLWNS test. Two asteriks indicate statistically significant results at the 5% level.

fractional process. Turning to the realised volatilities, we see slightly smaller memory estimates when using the GSE in contrast to the MLWN estimator. This could be caused by microstructure noise in the realised volatilites. However, for the cases of the BFX and the IBEX the bias is not really severe and can be neglected, as it could be caused by the inflation of the variance of the MLWN estimator as well. In contrast, we observe a memory estimate of the GSE for the BSESN time series which is half as high as that of the MLWN estimator. This could be caused by an inappropriate sampling frequency for the RV, which might be perturbed by microstructure noise.

The results of the MLWS and the MLWNS tests are given in Table 9. The MLWS procedure rejects the null of a true long memory process in the squared returns at a significance level of 5% in favour of spurious long memory. In contrast to that the testing procedure is not able to reject the null when using the realised volatility as a proxy. If we use the perturbation robust testing procedure, we do not face the struggle of contradictory results regarding the nature of the process any longer. For the realised volatilities as well as for the squared returns we are not able to reject the null of a true long memory or a true long memory with noise process. From this perspective, the differing results regarding the nature of the volatility could be explained as arising from an inadequate approximation of the spectral density near the zero frequencies.

6 Conclusion

This paper contains a multivariate extension to the local Whittle with noise estimator of Hurvich and Ray (2003) and a multivariate testing procedure for detecting spurious long memory in a perturbed fractional framework based on the first derivative of the estimator. We suggest approximating the multivariate spectrum of a perturbed fractional process near the zero frequency by adding additional constants. This results in a bias reduction when estimating the memory parameter in a signal-plus-noise setting. However, adding an additional constant results in an increase of the asymptotic variance of the memory estimator $\hat{d}$ which depends on the true long memory parameter d^0. We show that the estimator is consistent and asymptotically normal.

Based on the first derivative of the MLWN estimator, we additionally propose a (multivariate) perturbation robust testing procedure to differentiate between true and spurious long memory using the same technique as Qu (2011) and Sibbertsen et al. (2018). Reasoned by the poor approximation of the periodogram to the spectral density in a perturbed fractional context, these testing procedures suffer from poor size properties. Our results show that accounting for the additional noise term in the log squared returns leads to a non-rejection of the null and, therefore, to the non-rejection of a true long memory process. This confirms the results of the testing procedures of Qu (2011) and Sibbertsen et al. (2018) when applied to the realised variance. We derive the limiting distribution and show consistency of the procedure against multivariate random level shift and smooth trend processes. However, the limiting distribution of the test statistic is not pivotal.

A comprehensive Monte Carlo Study shows that the MLWS estimator achieves considerable bias reduction compared to the GSE of Shimotsu (2007), especially when the memory parameter and the nsr are high. We show that our modified score type test has improved size properties while having appropriate power in a setting where the time series of interest are contaminated by low frequencies.

In our empirical application we utilised the procedures to the daily log squared returns and realised volatilities of the BEL 20, S&P BSE SENSEX, and the Spanish IBEX and were not able to reject the null of a true perturbed long memory process for the squared returns and a true long memory process (potentially perturbed by microstructure noise) for the realised volatilities. All in all, we conclude that contradictory results regarding the nature of the volatility process could arise from an inadequate approximation of the spectral density near the zero frequency.

Acknowledgements

The financial support of Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged. We are also grateful to Bent Jesper Christensen, Tomas del Barrio Castro, Morten Nielsen, the participants of the SNDE 2022, the DAGStat in Hamburg 2022, the NSF-NBER Time Series Conference in Boston 2022, the Econometrics Research Seminar at University of Rostock and the University of the Balearic Islands for helpful comments and discussions.

Appendix

Proof of Theorem 1

The proof of consistency mainly follows the proof of Shimotsu (2007). In order to prove consistency of $\hat{d}$ we need to prove that $\lim_{n\to\infty} P(\hat{d} \in D_1) = 0$ and $(\hat{d} - d^0)\mathbb{1}\{\hat{d} \in D_2\} \xrightarrow{p} 0$, with $D_1 = (-\infty, d^0 - 1/2 + \varepsilon)^q \cap D$, $D_2 = [d^0 - 1/2 + \varepsilon, +\infty)^q \cap D$ with $D = [\Delta_1, \Delta_2]^q$, with $0 < \Delta_1 < \Delta_2 < 1/2$, $0 < \varepsilon < 1/4$, Θ is a compact and convex set in $\mathcal{R}^q$, and $\mathbb{1}\{\}$ is defined as the indicator function of the respective event. In the case where $d_a^0 \in (0, 1/2)$, $\forall a = 1, ..., q$, and ε is chosen small enough, the parameter space D_1 is empty. If it is not empty, the proof is a straightforward adaption of the proof given by Shimotsu (2007). We start with the proof of $(\hat{d} - d^0)\mathbb{1}\{\hat{d} \in D_2\} \xrightarrow{p} 0$ and therefore define

$$
\begin{aligned}
E(d,\theta) &= R(d,\theta) - R(d^0,\theta^0) \\
&= \log\det \hat{G}(d,\theta) - \frac{2}{m}\sum_{a=1}^{q} d_a \sum_{j=1}^{m} \log\lambda_j + \sum_{a=1}^{q}\sum_{j=1}^{m}\log(1+\theta_a\lambda_j^{2d_a}) \\
&\quad - \left[\log\det \hat{G}(d^0,\theta^0) - \frac{2}{m}\sum_{a=1}^{q} d_a^0 \sum_{j=1}^{m} \log\lambda_j + \sum_{a=1}^{q}\sum_{j=1}^{m}\log(1+\theta_a^0\lambda_j^{2d_a^0})\right] \\
&= \log\det \hat{G}(d,\theta) - \log\det \hat{G}(d^0,\theta^0) - \frac{2}{m}\sum_{a=1}^{q}(d_a - d_a^0)\sum_{j=1}^{m}\log\lambda_j + \sum_{a=1}^{q}\sum_{j=1}^{m}\log\alpha_j(d_a,\theta_a) \\
&= \log\det \hat{G}(d,\theta) + \log\left(\frac{2\pi m}{n}\right)^{-2((d_1-d_1^0)+...+(d_q-d_q^0))} - \log\det \hat{G}(d^0,\theta^0) \\
&\quad - \frac{2}{m}\sum_{a=1}^{q}(d_a - d_a^0)\sum_{j=1}^{m}(\log j - \log m) + \sum_{a=1}^{q}\sum_{j=1}^{m}\log\alpha_j(d_a,\theta_a) \\
&= \log A(d,\theta) - \log B(d,\theta) - \log A(d^0,\theta^0) + \log B(d^0,\theta^0) + S_2(d) + S_\alpha(d,\theta),
\end{aligned}
$$

with

$$
A(d,\theta) = \left(\frac{2\pi m}{n}\right)^{-2((d_1-d_1^0)+...+(d_q-d_q^0))} \det \hat{G}(d,\theta),
$$

$$
B(d,\theta) = \prod_{a=1}^{q}(2(d_a - d_a^0)+1)^{-1}\det G^0,
$$

$$
S_2(d) = -2\sum_{a=1}^{q}(d_a - d_a^0)\left(\frac{1}{m}\sum_{j=1}^{m}\log j - \log m\right) - \sum_{a=1}^{q}\log(2(d_a - d_a^0)+1), \tag{11}
$$

$$
S_\alpha(d,\theta) = \sum_{a=1}^{q}\sum_{j=1}^{m}\log\alpha_j(d_a,\theta_a) = \sum_{a=1}^{q}\sum_{j=1}^{m}\log\left(\frac{1+\theta_a\lambda_j^{2d_a}}{1+\theta_a^0\lambda_j^{2d_a^0}}\right). \tag{12}
$$

The upper bound for Eq. (11) is directly given by the results from Shimotsu (2007). Further, note that there exist constants $C > 0$ and c such that for each $a = 1, ..., q$ we obtain for Eq. (12)

$$\sup_{(d_a,\theta_a)\in D\times\Theta, k=1,\dots,m} |\alpha_k(d_a,\theta_a) - 1| \leq C \log(n/m)^2 e^{-\sqrt{c \log(n/m)}} = o(1)$$

as $n \to \infty$.

For the remaining parts $A(d,\theta)$ and $B(d,\theta)$, we need to show that there exists a non-random matrix $\Xi(d,\theta)$ such that we get

$$\sup_{D_2\times\Theta} |A(d,\theta) - \Xi(d,\theta)| = o_p(1), \tag{13}$$

$$\Xi(d,\theta) \geq B(d,\theta), \tag{14}$$

$$\Xi(d^0,\theta^0) = B(d^0,\theta^0). \tag{15}$$

In order to show that $(\hat{d} - d^0)\mathbb{1}\{\hat{d} \in D_2\} \xrightarrow{p} 0$ as $n \to \infty$ it needs to hold that

$$\begin{aligned}
\log A(d,\theta) - \log B(d,\theta) &\geq \log A(d,\theta) - \log \Xi(d,\theta) \\
&= \log\left([\Xi(d,\theta) + o_p(1)]/\Xi(d,\theta)\right) = o_p(1), \\
\log A(d^0,\theta^0) - \log B(d^0,\theta^0) &= \log\left(\left[\Xi(d^0,\theta^0) + o_p(1)\right]/\Xi(d^0,\theta^0)\right) = o_p(1),
\end{aligned}$$

uniformly in D_2.

We begin with proving Eq. (13). We start the proof by re-expressing $\Lambda_j(d,\theta)^{-1}$ by

$$\begin{aligned}
\Lambda_j(d,\theta)^{-1} &= \Lambda_j(d-d^0,\theta-\theta^0)^{-1}\Lambda_j(d^0,\theta^0)^{-1} \\
&= \text{diag}(\lambda_j^{(d_a-d_a^0)} e^{i(\lambda_j-\pi)(d_a-d_a^0)/2}\alpha_j(d_a,\theta_a)^{-1/2})\text{diag}(\lambda_j^{d_a^0} e^{i(\lambda_j-\pi)d_a^0/2}(1+\theta_a^0\lambda_j^{2d_a^0})^{-1/2}) \\
&= \Lambda_j(\bar{d},\bar{\theta})^{-1}\Lambda_j(d^0,\theta^0)^{-1}.
\end{aligned}$$

Using this, we have

$$\begin{aligned}
A(d,\theta) &= \left(\frac{2\pi m}{n}\right)^{-2((d_1-d_1^0)+\ldots+(d_q-d_q^0))} \\
&\quad \times \det\left\{\frac{1}{m}\sum_{j=1}^{m} \text{Re}\left[\Lambda_j(\bar{d},\bar{\theta})^{-1}\Lambda_j(d^0,\theta^0)^{-1}I_z(\lambda_j)\Lambda_j^*(d^0,\theta^0)^{-1}\Lambda_j^*(\bar{d},\bar{\theta})^{-1}\right]\right\} \\
&= \det\left\{\frac{1}{m}\sum_{j=1}^{m} \text{Re}\left[M_j(\bar{d},\bar{\theta})\Lambda_j(d^0,\theta^0)^{-1}I_z(\lambda_j)\Lambda_j^*(d^{\cdot}\theta^0)^{-1}M_j^*(\bar{d},\bar{\theta})\right]\right\},
\end{aligned} \tag{16}$$

with $M_j(\bar{d},\bar{\theta}) = \text{diag}(e^{i(\lambda_j-\pi)(d_a-d_a^0)/2}(j/m)^{d_a-d_a^0}\alpha_j(d_a,\theta_a)^{-1/2})$. In the following, let $I_{zj} = I_z(\lambda_j)$ and $w_{aj} = w_a(\lambda_j)$ denote the a-th element of $w(\lambda_j)$ in order to ease the notation. Taking a closer look at the (ab)-th element of Eq. (16) yields

$$\frac{1}{m}\sum_{j=1}^{m} \text{Re}\left[e^{i(\lambda_j-\pi)(\bar{d}_a-\bar{d}_b)/2}\left(\frac{j}{m}\right)^{\bar{d}_a+\bar{d}_b}\alpha_j(d_a,\theta_a)^{-1/2}\alpha_j(d_b,\theta_b)^{-1/2}\frac{w_{aj}w_{bj}^*}{\Lambda_{ja}(d^0,\theta^0)\Lambda_{jb}^*(d^0,\theta^0)}\right].$$

By Lemma 2 and summation by parts we have uniformly in (a,b)

$$
\begin{aligned}
&\sup_{D_2\times\Theta}\left|\frac{1}{m}\sum_{j=1}^{m} e^{i(\lambda_j-\pi)(\bar d_a-\bar d_b)/2}\left(\frac{j}{m}\right)^{\bar d_a+\bar d_b}\alpha_j(d_a,\theta_a)^{-1/2}\alpha_j(d_b,\theta_b)^{-1/2}\left(\frac{w_{aj}w_{bj}^*}{\Lambda_{ja}(d^0,\theta^0)\Lambda_{jb}^*(d^0,\theta^0)}-G_{ab}^0\right)\right|\\
&\le \frac{1}{m}\sum_{r=1}^{m-1}\sup_{D_2\times\Theta}\left|\left(\frac{r}{m}\right)^{\bar d_a+\bar d_b} e^{i(\lambda_r-\pi)(\bar d_a-\bar d_b)/2}\alpha_r(d_a,\theta_a)^{-1/2}\alpha_r(d_b,\theta_b)^{-1/2}\right.\\
&\quad\left.-\left(\frac{r+1}{m}\right)^{\bar d_a+\bar d_b} e^{i(\lambda_{r+1}-\pi)(\bar d_a-\bar d_b)/2}\alpha_{r+1}(d_a,\theta_a)^{-1/2}\alpha_{r+1}(d_b,\theta_b)^{-1/2}\right|\\
&\quad\times\left|\sum_{j=1}^{r}\left(\frac{w_{aj}w_{bj}^*}{\Lambda_{ja}(d^0,\theta^0)\Lambda_{jb}^*(d^0,\theta^0)}-G_{ab}^0\right)\right|+\left|\frac{1}{m}\sum_{j=1}^{m}\left(\frac{w_{aj}w_{jb}^*}{\Lambda_{ja}(d^0,\theta^0)\Lambda_{jb}^*(d^0,\theta^0)}-G_{ab}^0\right)\right|\\
&\le c\sum_{r=1}^{m-1}\left(\frac{r}{m}\right)^{2\epsilon}\frac{1}{r^2}\left|\sum_{j=1}^{r}\left(\frac{w_{aj}w_{bj}^*}{\Lambda_{ja}(d^0,\theta^0)\Lambda_{jb}^*(d^0,\theta^0)}-G_{ab}^0\right)\right|\\
&\quad+\frac{1}{m}\left|\sum_{j=1}^{m}\left(\frac{w_{aj}w_{bj}^*}{\Lambda_{ja}(d^0,\theta^0)\Lambda_{jb}^*(d^0,\theta^0)}-G_{ab}^0\right)\right| \qquad (17)\\
&= o_p(1).
\end{aligned}
$$

Uniformly over $D_2\times\Theta$ we have

$$
\frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\left[M_j(\bar d,\bar\theta)\Lambda_j(d^0,\theta^0)^{-1}I_{zj}\Lambda_j^*(d^0,\theta^0)^{-1}M_j^*(\bar d,\bar\theta)\right]=\frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\left[M_j(\bar d,\bar\theta)G^0M_j^*(\bar d,\bar\theta)\right]+o_p(1).
$$

In order to derive an approximation of the right hand side we use the results from Robinson (1995a) and Shimotsu (2007) that $\sup_{c\ge\gamma\ge\Delta}\left|\gamma m^{-1}\sum_{j=1}^{m}(j/m)^{\gamma-1}-1\right|=O(m^{-\Delta})$ for $0<\Delta<c<\infty$ and $e^{i(\lambda-\pi)(\bar d_a-\bar d_b)/2}=e^{-i\pi(\bar d_a-\bar d_b)/2}+O(\lambda)$. Based on that we define $e^{-i\pi(\bar d_a-\bar d_b)/2}$ to be the (a,b)-th element of the matrix $\mathcal{E}(\bar d)$ and $(1+\bar d_a+\bar d_b)^{-1}\alpha(d_a,\theta_a)^{-1/2}\alpha(d_b,\theta_b)^{-1/2}=\int_0^1 x^{\bar d_a+\bar d_b}dx\,\alpha(d_a,\theta_a)^{-1/2}\alpha(d_b,\theta_b)^{-1/2}$ to be the (a,b)-th element of the matrix $M_\infty(\bar d,\bar\theta)$. Now we get

$$
\frac{1}{m}\sum_{j=1}^{m}\left[M_j(\bar d,\bar\theta)G^0M_j^*(\bar d,\bar\theta)\right]=\mathcal{E}(\bar d)\odot M_\infty(\bar d,\bar\theta)\odot G^0+O(mn)^{-1}+O(m^{-2\epsilon}),
$$

with $\odot$ symbolising the Hadamard product. By the same arguments as in Shimotsu (2007) Eq. (13) follows with

$$
\Xi(d,\theta)=\det(\mathrm{Re}[\mathcal{E}(\bar d)]\odot M_\infty(\bar d,\bar\theta)\odot G^0).
$$

The proof for Eq. (14) and Eq. (15) follows as in Shimotsu (2007) by using Oppenheim's inequality and the fact that $\alpha_j(d^0,\theta^0)=1$ for all entries in $M_\infty(0)$. This proves the first part.

Now consider the proof of $\lim_{n\to\infty} P(\hat{d} \in D_1) = 0$. We have

$$\begin{aligned}
&\log \det \hat{G}(d,\theta) - \log \det \hat{G}(d^0,\theta^0) - \frac{2}{m}\sum_{a=1}^{q}(d_a - d_a^0)\sum_{j=1}^{m}\log \lambda_j - \frac{1}{m}\sum_{a=1}^{q}\sum_{j=1}^{m}\log(\alpha_j(d_a,\theta_a)) \\
&= \log \det \frac{1}{m}\sum_{j=1}^{m}\text{Re}\left[\Lambda_j(d-d^0,\theta-\theta^0)^{-1}\Lambda_j(d^0,\theta^0)^{-1}I_{zj}\Lambda_j^*(d^0,\theta^0)^{-1}\Lambda_j^*(d-d^0,\theta-\theta^0)^{-1}\right] \\
&- \frac{2}{m}\sum_{a=1}^{q}(d_a - d_a^0)\sum_{j=1}^{m}\log \lambda_j - \log \det \hat{G}(d^0,\theta^0) - \frac{1}{m}\sum_{a=1}^{q}\sum_{j=1}^{m}\log(\alpha_j(d_a,\theta_a)) \\
&= \log \det \hat{D}(d,\theta) - \log \det \hat{D}(d^0,\theta^0) - \frac{1}{m}\sum_{a=1}^{q}\sum_{j=1}^{m}\log(\alpha_j(d_a,\theta_a))
\end{aligned}$$

with

$$\hat{D}(d,\theta) = \frac{1}{m}\sum_{j=1}^{m}\text{Re}\left[P_j(d-d^0,\theta-\theta^0)\Lambda_j(d^0,\theta^0)^{-1}I_{zj}\Lambda_j^*(d^0,\theta^0)^{-1}P_j^*(d-d^0,\theta-\theta^0)\right]$$

$$P_j(d-d^0,\theta-\theta^0) = \text{diag}(e^{i(\lambda_j-\pi)(d-d^0)/2}(j/p)^{d-d^0}\alpha_j(d,\theta)^{-1/2}), \quad p = (m!\,)^{1/m}.$$

Because of the fact that the logarithm is a monotone increasing function of its argument it is sufficient to show

$$P(\inf_{(d,\theta)\in D_1\times\Theta} \det \hat{D}(d,\theta) - \det \hat{D}(d^0,\theta^0) \le 0) \longrightarrow 0 \;\text{ as } n \longrightarrow \infty. \tag{18}$$

Lets define for each summand of $\hat{D}(d,\theta)$ a q-vector Υ_j as

$$\begin{aligned}
&\text{Re}\left[P_j(d-d^0,\theta-\theta^0)\Lambda_j(d^0,\theta^0)^{-1}I_{zj}\Lambda_j^*(d^0,\theta^0)^{-1}P_j^*(d-d^0,\theta-\theta^0)\right] \\
&= \text{Re}\left[\Upsilon_j\Upsilon_j^*\right] = \text{Re}[\Upsilon_j](\text{Re}[\Upsilon_j])' + \text{Im}[\Upsilon_j](\text{Im}[\Upsilon_j])',
\end{aligned}$$

which is positive semidefinite. Hence, we get a sum of m positive semidefinite matrices for $\hat{D}(d,\theta)$. Lets further define for a fixed $\kappa \in (0,1)$

$$\begin{aligned}
\hat{D}_\kappa(d,\theta) &= \frac{1}{m}\sum_{j=[\kappa m]}^{m}\text{Re}\left[P_j(d-d^0,\theta-\theta^0)\Lambda_j(d^0,\theta^0)^{-1}I_{zj}\Lambda_j^*(d^0,\theta^0)^{-1}P_j^*(d-d^0,\theta-\theta^0)\right] \\
K_\kappa(d,\theta) &= \frac{1}{m}\sum_{j=[\kappa m]}^{m}\text{Re}\left[P_j(d-d^0,\theta-\theta^0)G^0P_j^*(d-d^0,\theta-\theta^0)\right].
\end{aligned}$$

Following Lütkepohl (1996) we know that

$$\det \hat{D}(d,\theta) \ge \det \hat{D}_\kappa(d,\theta).$$

We can use $K_\kappa(d,\theta)$ in order to uniformly approximate $\hat{D}_\kappa(d,\theta)$. Let the (a,b)-th element of the difference between $\hat{D}_\kappa(d,\theta)$ and $K_\kappa(d,\theta)$ be defined as

$$\begin{aligned}
&\frac{1}{m}\sum_{j=[\kappa m]}^{m} \text{Re}\left[e^{i(\lambda_j-\pi)((d_a-d_a^0)-(d_b-d_b^0))/2}\left(\frac{j}{p}\right)^{(d_a-d_a^0)+(d_b-d_b^0)} \alpha_j(d_a,\theta_a)^{-1/2}\alpha_j(d_b,\theta_b)^{-1/2}\right.\\
&\qquad\left.\times\left(\frac{w_{aj}w_{bj}^*}{\Lambda_{ja}(d^0,\theta^0)\Lambda_{jb}^*(d^0,\theta^0)}-G_{ab}^0\right)\right]\\
&=\left(\frac{m}{p}\right)^{(d_a-d_a^0)+(d_b-d_b^0)} \text{Re}\left[\frac{1}{m}\sum_{j=[\kappa m]}^{m} e^{i(\lambda_j-\pi)((d_a-d_a^0)-(d_b-d_b^0))/2}\left(\frac{j}{m}\right)^{(d_a-d_a^0)+(d_b-d_b^0)}\right.\\
&\qquad\left.\times\alpha_j(d_a,\theta_a)^{-1/2}\alpha_j(d_b,\theta_b)^{-1/2}\left(\frac{w_{aj}w_{bj}^*}{\Lambda_{ja}(d^0,\theta^0)\Lambda_{jb}^*(d^0,\theta^0)}-G_{ab}^0\right)\right].
\end{aligned}$$

This difference is $o_p(1)$ uniformly over $(d-d^0,\theta-\theta^0)\in D_1\times\Theta$. This result is derived in a similar way as Eq. (17) from summation by parts, Lemma 2, Lemma 5.4. of Shimotsu and Phillips (2005), and the upper bound of $\alpha_j(d,\theta)$. With this result, we have for any $\kappa\in(0,1)$

$$\sup_{D_1\times\Theta}\left|\det\,\hat{D}_\kappa(d,\theta)-\det\,K_\kappa(d,\theta)\right| = o_p(1) \text{ as } n\longrightarrow\infty.$$

Now we only need to derive a lower bound for $K_\kappa(d,\theta)$ for $\hat{d}\in D_1$ in order to complete the proof. Therefore, we rewrite $K_\kappa(d,\theta)$ as

$$K_\kappa(d,\theta)=M_m^\kappa(d-d^0,\theta-\theta^0)\odot G^0,$$

with $M_m^\kappa(d-d^0,\theta-\theta^0)$ being a positive semidefinite matrix given by

$$\begin{aligned}
M_m^\kappa &= \frac{1}{m}\sum_{j=[\kappa m]}^{m}\text{Re}\left[Z_jZ_j^*\right],\\
Z_j &= \left(e^{i(\lambda_j-\pi)(d_1-d_1^0)/2}\left(\frac{j}{p}\right)^{(d_1-d_1^0)}\alpha_j(d_1,\theta_1)^{-1/2},...,e^{i(\lambda_j-\pi)(d_q-d_q^0)/2}\left(\frac{j}{p}\right)^{(d_q-d_q^0)}\alpha_j(d_q,\theta_q)^{-1/2}\right).
\end{aligned}$$

By Oppenheim's inequality, Lemma 5.5 of Shimotsu and Phillips (2005), and Lemma 2 of Shimotsu (2006), there exists an $\epsilon\in(0,0.1)$ and $\bar{\kappa}\in(0,1/4)$ such that we get, for a sufficiently large m and all $\kappa\in(0,\bar{\kappa})$,

$$\begin{aligned}
\inf_{(d,\theta)\in D_1\times\Theta}\det\,K_\kappa(d,\theta) &\geq \det\,G^0\inf_{d\in D_1}\prod_{a=1}^{q}\frac{1}{m}\sum_{j=[\kappa m]}^{m}\left(\frac{j}{p}\right)^{2(d_a-d_a^0)}\\
&\geq \det\,G^0(1+2\epsilon)(1-\kappa^{2\varepsilon})^{q-1}+o(1).
\end{aligned}$$

If we choose κ small enough then we get $(1+2\epsilon)(1-\kappa^{2\varepsilon})^{q-1}\geq 1+\epsilon$. Hence, we get

$$\inf_{(d,\theta)\in D_1\times\Theta}\det\,\hat{D}_\kappa(d,\theta)=\inf_{(d,\theta)\in D_1\times\Theta}\det\,K_\kappa(d,\theta)+o_p(1)\geq\det\,G^0(1+\epsilon)+o_p(1).$$

We already now that det $\hat{D}(d^0,\theta^0) = \det \hat{G}(d^0,\theta^0) \xrightarrow{p} \det G^0$ as $n \to \infty$. This gives us the desired result of

$$P(\inf_{(d,\theta)\in D_1\times\Theta} \det \hat{D}_\kappa(d,\theta) - \det \hat{D}(d^0,\theta^0) \leq 0) \longrightarrow 0 \ \text{ as } n \longrightarrow \infty$$

and Eq. (18) follows which completes the proof.
We strengthen this result by showing that $d_a - d_a^0 = o_p((\log\, n)^{-6})$ for all $a = 1, ..., q$. We achieve a faster convergence rate compared to Shimotsu (2007) caused by the more appropriate approximation of the spectral density near the zero frequency. Therefore, we can use the proof of Hurvich et al. (2005) with slight adaptions of the logarithmic rate to $(\log\, n)^{-6}$ and use the maximum value of the long memory parameter in order to adapt for the multivariate version.

Proof of Theorem 2

This proof mainly uses the approach of Andrews and Sun (2004), and extends it to the multivariate case, combining it with the method of Shimotsu (2007). We start with the score and the Hessian of the objective function. In addition, we define i_a to be a $q \times q$ matrix where the off-diagonal elements are zero and the a-th diagonal element is equal to one, $\Lambda_j(d,\theta)^{-1} = \Lambda_j^{-1} = \text{diag}(\lambda_j^{d_a} e^{i(\lambda_j-\pi)d_a/2}(1+\theta_a\lambda_j^{2d_a})^{-1/2})$, and $I_z(\lambda_j) = I_{zj}$. Let the score $m\nabla R(d,\theta) = S(d,\theta)$, where ∇ is the gradient of the respective function, given by

$$\begin{aligned} S(d_a,\theta_a) = &-\frac{2}{m}\sum_{j=1}^m X_{ja} + \text{tr}\left[\hat{G}(d_a,\theta_a)^{-1}\left(\frac{1}{m}\sum_{j=1}^m X_{ja}\text{Re}\left[(\Lambda_j^0)^{-1}(i_a I_{zj} + I_{zj} i_a)(\Lambda_j^{0*})^{-1}\right] X_{ja}\right)\right] \\ &+ \text{tr}\left[\hat{G}(d_a,\theta_a)^{-1}\left(\frac{1}{m}\sum_{j=1}^m L_j\text{Im}\left[(\Lambda_j^0)^{-1}(-i_a I_{zj} + I_{zj} i_a)(\Lambda_j^{0*})^{-1}\right]\right)\right] \\ = &-\frac{2}{m}\sum_{j=1}^m X_{ja} + \text{tr}\left[\hat{G}(d_a,\theta_a)^{-1}Q_{1a}\right] + \text{tr}\left[\hat{G}(d_a,\theta_a)^{-1}Q_{2a}\right], \end{aligned} \tag{19}$$

with $Q_{1a} = \frac{1}{m}\sum_{j=1}^m X_{ja}\text{Re}\left[(\Lambda_j^0)^{-1}(i_a I_{zj} + I_{zj} i_a)(\Lambda_j^{0*})^{-1}\right] X_{ja}$, $Q_{2a} = \frac{1}{m}\sum_{j=1}^m L_j\text{Im}\left[(\Lambda_j^0)^{-1}(-i_a I_{zj} + I_{zj} i_a)\right.$ $\left.(\Lambda_j^{0*})^{-1}\right]$, $L_j = ((\lambda_j - \pi)/2, 0)'$, and $X_{ja} = (\log \lambda_j/(1+\theta_a\lambda_j^{2d_a})^{1/2}, -\lambda_j^{d_a}/(2(1+\theta_a\lambda_j^{2d_a})^{1/2})'$.
The Hessian $m\nabla^2 R(d,\theta) = H(d,\theta)$ is given by the following terms

$$\begin{aligned} H(d,\theta) &= H_1(d,\theta) + H_2(d,\theta), \qquad (20)\\ H_1(d,\theta) &= \text{tr}\left[-\hat{G}^{-1}(d,\theta)\nabla\hat{G}(d_a,\theta_a)\hat{G}^{-1}(d,\theta)\nabla\hat{G}(d_b,\theta_b) + \hat{G}^{-1}(d,\theta)\nabla^2\hat{G}(d_{a,b},\theta_{a,b})\right], \\ H_2(d,\theta) &= -\frac{2}{m}\sum_{j=1}^m \nabla X_j. \end{aligned}$$

The derivatives of $\hat{G}(d,\theta)$ are given by

$$\nabla\hat{G}(d_a,\theta_a) = \frac{1}{m}\sum_{j=1}^{m}\text{Re}\left[\begin{pmatrix}\frac{\log\,\lambda_j}{1+\theta_a\lambda_j^{2d_a}}+\frac{\lambda_j-\pi}{2}i\\ -\frac{\lambda_j^{2d_a}}{2(1+\theta_a\lambda_j^{2d_a})}\end{pmatrix} i_a\Lambda_j(d,\theta)^{-1}I_{zj}\Lambda_j^*(d,\theta)^{-1}\right] + \frac{1}{m}\sum_{j=1}^{m}\text{Re}\left[\begin{pmatrix}\frac{\log\,\lambda_j}{1+\theta_a\lambda_j^{2d_a}}+\frac{\lambda_j-\pi}{2}i\\ -\frac{\lambda_j^{2d_a}}{2(1+\theta_a\lambda_j^{2d_a})}\end{pmatrix} \Lambda_j(d,\theta)^{-1}I_{zj}\Lambda_j^*(d,\theta)^{-1}i_a\right]$$

and

$$\begin{aligned}\nabla^2\hat{G}(d_{a,b},\theta_{a,b}) =& \frac{1}{m}\sum_{j=1}^{m}\text{Re}\left[\begin{pmatrix}\frac{\log\,\lambda_j}{\sqrt{1+\theta_a\lambda_j^{2d_a}}}+\frac{\lambda_j-\pi}{2}i\\ -\frac{\lambda_j^{d_a}}{2(\sqrt{1+\theta_a\lambda_j^{2d_a}})}\end{pmatrix}\begin{pmatrix}\frac{\log\,\lambda_j}{\sqrt{1+\theta_b\lambda_j^{2d_b}}}+\frac{\lambda_j-\pi}{2}i\\ -\frac{\lambda_j^{d_b}}{2(\sqrt{1+\theta_b\lambda_j^{2d_b}})}\end{pmatrix} i_a i_b\Lambda_j(d,\theta)^{-1}I_{zj}\Lambda_j^*(d,\theta)^{-1}\right]\\
&+\frac{1}{m}\sum_{j=1}^{m}\text{Re}\left[\begin{pmatrix}\frac{\log\,\lambda_j}{\sqrt{1+\theta_a\lambda_j^{2d_a}}}+\frac{\lambda_j-\pi}{2}i\\ -\frac{\lambda_j^{d_a}}{2(\sqrt{1+\theta_a\lambda_j^{2d_a}})}\end{pmatrix} i_a\Lambda_j(d,\theta)^{-1}I_{zj}\Lambda_j^*(d,\theta)^{-1}i_b\begin{pmatrix}\frac{\log\,\lambda_j}{\sqrt{1+\theta_b\lambda_j^{2d_b}}}+\frac{\lambda_j-\pi}{2}i\\ -\frac{\lambda_j^{d_b}}{2(\sqrt{1+\theta_b\lambda_j^{2d_b}})}\end{pmatrix}\right]\\
&+\frac{1}{m}\sum_{j=1}^{m}\text{Re}\left[\begin{pmatrix}\frac{\log\,\lambda_j}{\sqrt{1+\theta_b\lambda_j^{2d_b}}}+\frac{\lambda_j-\pi}{2}i\\ -\frac{\lambda_j^{d_b}}{2(\sqrt{1+\theta_b\lambda_j^{2d_b}})}\end{pmatrix} i_b\Lambda_j(d,\theta)^{-1}I_{zj}\Lambda_j^*(d,\theta)^{-1}i_a\begin{pmatrix}\frac{\log\,\lambda_j}{\sqrt{1+\theta_a\lambda_j^{2d_a}}}+\frac{\lambda_j-\pi}{2}i\\ -\frac{\lambda_j^{d_a}}{2(\sqrt{1+\theta_a\lambda_j^{2d_a}})}\end{pmatrix}\right]\\
&+\frac{1}{m}\sum_{j=1}^{m}\text{Re}\left[\begin{pmatrix}\frac{\log\,\lambda_j}{\sqrt{1+\theta_a\lambda_j^{2d_a}}}+\frac{\lambda_j-\pi}{2}i\\ -\frac{\lambda_j^{d_a}}{2(\sqrt{1+\theta_a\lambda_j^{2d_a}})}\end{pmatrix}\begin{pmatrix}\frac{\log\,\lambda_j}{\sqrt{1+\theta_b\lambda_j^{2d_b}}}+\frac{\lambda_j-\pi}{2}i\\ -\frac{\lambda_j^{d_b}}{2(\sqrt{1+\theta_b\lambda_j^{2d_b}})}\end{pmatrix}\Lambda_j(d,\theta)^{-1}I_{zj}\Lambda_j^*(d,\theta)^{-1}i_a i_b\right].\end{aligned}$$

Further, we have

$$J_{na} = \sum_{j=1}^{m}\left(X_{ja}-\frac{1}{m}\sum_{k=1}^{m}X_{ka}\right)\left(X_{ja}-\frac{1}{m}\sum_{k=1}^{m}X_{ka}\right)'.$$

Also we need to define

$$\boldsymbol{J}_a = \begin{pmatrix}1 & -\frac{2d_a^0}{2(1+2d_a^0)^2}\\ -\frac{2d_a^0}{2(1+2d_a^0)^2} & \frac{2(d_a^0)^2}{2(1+2d_a^0)^2(1+4d_a^0)}\end{pmatrix}$$

Next we establish a multivariate version of the lemma used by Frederiksen et al. (2012) in order to be able to show joint asymptotic normality of $\hat{d}$ and $\hat{\theta}$. Therefore, we introduce the notation $D(\tau)=\{d\in D:\log^6(n)||d-d^0||<\tau\}$ for $\tau>0$. The proof is given in the next Section.

Lemma 1. *Given the assumptions of Theorem 2 together with $B_n = diag(m^{1/2}, m^{1/2}\lambda_m^{2d_a})$ we have, as $n\longrightarrow\infty$*

(a) $B_n^{-1}J_nB_n^{-1}\longrightarrow \boldsymbol{J}$,

(b) $\sup_{d\in D(\tau),\theta\in\Theta}||B_n^{-1}H_1(d,\theta)B_n^{-1}||\longrightarrow\boldsymbol{\Omega}$ *and* $\sup_{d\in D(\tau),\theta\in\Theta}||B_n^{-1}H_2(d,\theta)B_n^{-1}||=o_p(1)$*, for all se-*

quences of constants τ for which $n > 1$ and $\tau = o(1)$

(c) $B_n^{-1} S(d^0, \theta^0) B_n^{-1} \xrightarrow{d} N(0, \mathbf{\Omega})$.

By the same argument as in Frederiksen et al. (2012) we know that the MLWN estimator exists because the MLWN likelihood is a continuous function on a compact set. We know by Lemma 1 and by Lemma 1 of Andrews and Sun (2004) that there is a unique solution to the first order conditions with probability tending to one which is satisfied by the results of Theorem 2. Additionally, we only need to prove positive definiteness of our Hessian to prove convexity of our log likelihood function. For the first part of the Hessian, $H_1(d, \theta)$, we have the positive definiteness directly given by the assumption that our matrix $G(d, \theta)$ is positive definite. Since $H_1(d, \theta)$ contains the Hadamard product of two positive definite matrices positive definiteness for the first part of the Hessian follows immediately.
For the second part of the Hessian, $H_2(d, \theta)$, we know by Lemma 1 that $||B_n^{-1} H_2(d, \theta) B_n^{-1}|| = o_p(1)$ uniformly over the set $(d, \theta) \in D(\tau) \times \Theta$. Using this together with the result of Theorem 1 that $\hat{d} \in D(\tau)$ with probability tending to one shows that our Hessian is positive definite with probability tending to one. This concludes the proof.
It should be noted that we omit part (c) and (d) of Lemma 1 as given in Frederiksen et al. (2012) as they do not depend on the univariate structure of the model.

Proof of Lemma 1

Proof of (a)

The convergence result directly follows from Lemma 2 in Andrews and Guggenberger (2003) by approximating sums by integrals and holds in our multivariate case as well. We only need to replace the single d values by the vector $d = (d_1, ..., d_q)'$.

Proof of (b)

Let's define for $k_1 = 0, 1, 2$ and $k_2 = 0, d_1, ..., d_q$ the following expressions

$$\hat{G}_{k_1,k_2}(d, \theta) = \frac{1}{m} \sum_{j=1}^{m} (\log \lambda_j)^{k_1} (j/m)^{2k_2} \text{Re}\left[\Lambda_j(d, \theta)^{-1} I_{zj} \Lambda_j^*(d, \theta)^{-1}\right],$$

$$\bar{G}_{k_1}(d, \theta) = \frac{1}{m} \sum_{j=1}^{m} (\log \lambda_j)^{k_1} \text{Im}\left[\Lambda_j(d, \theta)^{-1} I_{zj} \Lambda_j^*(d, \theta)^{-1}\right],$$

$$J_{k_1,k_2}(d, \theta) = \frac{1}{m} \sum_{j=1}^{m} (\log \lambda_j)^{k_1} (j/m)^{2k_2} \Lambda_j(d - d^0, \theta - \theta^0)^{-1} G^0 \Lambda_j^*(d - d^0, \theta - \theta^0)^{-1}.$$

We start with the proof of the first part of the Hessian $H_1(d, \theta)$, respectively. The proof is similar to the proof of Shimotsu (2007). Given the expression stated above we can express the elements in $B_n^{-1}H_1(d,\theta)B_n^{-1}$ as

$$\nabla\hat{G}(d,\theta) = i_a\hat{G}_{1,d_a}(d,\theta) + \hat{G}_{1,d_a}(d,\theta)i_a + (\pi/2,0)'\, i_a\bar{G}_0(d,\theta) - (\pi/2,0)'\,\bar{G}_0(d,\theta)i_a + o_p((\log\,(m))^{-1})$$
$$\begin{aligned}\nabla^2\hat{G}(d,\theta) &= i_a i_b\hat{G}_{2,d_a+d_b}(d,\theta) + i_a\hat{G}_{2,d_a+d_b}(d,\theta)i_b + i_b\hat{G}_{2,d_a+d_b}(d,\theta)i_a + \hat{G}_{2,d_a+d_b}(d,\theta)i_a i_b \\ &\quad + \left((\pi/2)^2\,,0\right)'\left[-i_a i_b\hat{G}_{0,0}(d,\theta) + i_a\hat{G}_{0,0}(d,\theta)i_b + i_b\hat{G}_{0,0}(d,\theta)i_a - \hat{G}_{0,0}(d,\theta)i_a i_b\right] \\ &\quad + (\pi,0)'\, i_a i_b\bar{G}_1(d,\theta) - (\pi,0)'\,\bar{G}_1(d,\theta)i_a i_b + o_p(1).\end{aligned}$$

We need to show that

$$\sup_{d\in D(\tau),\theta\in\Theta}\left\|\frac{1}{m}\sum_{j=1}^{m}(\log\,\lambda_j)^{k_1}\left(\frac{j}{m}\right)^{2k_2}\Lambda_j(d,\theta)^{-1}I_{zj}\Lambda_j^*(d,\theta)^{-1} - J_{k_1,k_2}(d,\theta)\right\| = o_p((\log\,m)^{-2}), \tag{21}$$

$$\sup_{d\in D(\tau),\theta\in\Theta}\left\|J_{k_1,k_2}(d,\theta) - G^0\frac{1}{m}\sum_{j=1}^{m}(\log\,\lambda_j)^{k_1}\left(\frac{j}{m}\right)^{2k_2}\right\| = o((\log\,m)^{-2}) \tag{22}$$

in order to get the results

$$\hat{G}_{k_1,k_2}(d,\theta) = G^0\frac{1}{m}\sum_{j=1}^{m}(\log\,\lambda_j)^{k_1}\left(\frac{j}{m}\right)^{2k_2} + o_p((\log\,m)^{-2}), \qquad \bar{G}_{k_1}(d,\theta) = o_p((\log\,m)^{-2})$$

uniformly on $(d,\theta)\in D(\tau)\times\Theta$. Given these results, we can rewrite the expressions of our Hessian as (apart from smaller order terms)

$$\begin{aligned}\hat{G}^{-1}(d,\theta) &= (G^0)^{-1}, \\ \nabla\hat{G}(d,\theta) &= \frac{1}{m}\sum_{j=1}^{m}\widetilde{X}_{ja}G^0_{1a}, \\ \nabla^2\hat{G}(d,\theta) &= \frac{1}{m}\sum_{j=1}^{m}\widetilde{X}_{ja}\widetilde{X}'_{jb}G^0_{2ab} + \left((\pi/2)^2,0\right)' G^0_{3ab},\end{aligned}$$

where $\widetilde{X}_{ja} = (\log\,\lambda_j/(1+\theta_a\lambda_j^{2d_a}), -(j/m)^{2d_a}/2(1+\theta_a\lambda_j^{2d_a}))'$, $G^0_{1a} = i_aG^0 + G^0 i_a$, $G^0_{2ab} = i_a i_b G^0 + i_a G^0 i_b + i_b G^0 i_a + G^0 i_a i_b$, and $G^0_{3ab} = -i_a i_b G^0 + i_a G^0 i_b + i_b G^0 i_a - G^0 i_a i_b$. This gives us

$$\begin{aligned}\hat{G}^{-1}(d,\theta)\nabla\hat{G}(d_a,\theta_a)\hat{G}^{-1}(d,\theta)\nabla\hat{G}(d_b,\theta_b) &= (G^0)^{-1}\left[\frac{1}{m}\sum_{j=1}^{m}\widetilde{X}_{ja}G^0_{1a}\right](G^0)^{-1}\left[\frac{1}{m}\sum_{j=1}^{m}\widetilde{X}_{jb}G^0_{1b}\right] \\ &= \left[\frac{1}{m}\sum_{j=1}^{m}\widetilde{X}_{ja}\right]\left[\frac{1}{m}\sum_{j=1}^{m}\widetilde{X}_{jb}\right]\left[(G^0)^{-1}G^0_{1a}(G^0)^{-1}G^0_{1b}\right]\end{aligned}$$

and

$$\hat{G}^{-1}(d,\theta)\nabla^2\hat{G}(d,\theta) = (G^0)^{-1}\left[\frac{1}{m}\sum_{j=1}^{m}\widetilde{X}_{ja}\widetilde{X}'_{jb}G^0_{2ab} + \left((\pi/2)^2,0\right)' G^0_{3ab}\right]$$
$$= \frac{1}{m}\sum_{j=1}^{m}\widetilde{X}_{ja}\widetilde{X}'_{jb}(G^0)^{-1}G^0_{2ab} + \left((\pi/2)^2,0\right)'(G^0)^{-1}G^0_{3ab}.$$

Caused by the fact that tr $\left[(G^0)^{-1}G^0_{1a}(G^0)^{-1}G^0_{1b}\right]$ = tr $\left[(G^0)^{-1}G^0_{2ab}\right]$ we receive for the first part of the normalised Hessian

$$B_n^{-1}H_1(d,\theta)B_n^{-1} = \frac{1}{m}\sum_{j=1}^{m}\left(\widetilde{X}_{ja} - \frac{1}{m}\sum_{j=1}^{m}\widetilde{X}_{ja}\right)\left(\widetilde{X}_{jb} - \frac{1}{m}\sum_{j=1}^{m}\widetilde{X}_{jb}\right)' \text{tr}\left[(G^0)^{-1}G^0_{2ab}\right]$$
$$+ \text{tr}\left[\left((\pi/2)^2\right)'(G^0)^{-1}G^0_{3ab}\right].$$

Combining this with the result given in Lemma 1 (a) proves the first part of Lemma 1 (b). Now, the only thing we need to show are the arguments given in Eq. (21) and Eq. (22). Writing the left hand side of Eq. (21) as

$$\frac{1}{m}\sum_{j=1}^{m} g_j(d-d^0,\theta-\theta^0)\left[e^{i(\lambda_j-\pi)(d^0_a-d^0_b)/2}\lambda_j^{d^0_a+d^0_b}h_{ab,j}(d^0,\theta^0)w_{aj}w^*_{bj} - G^0_{ab}\right].$$

We define $g_j(d-d^0,\theta-\theta^0) = (\log\,\lambda_j)^{k_1}\left(\frac{j}{m}\right)^{2k_2} e^{i(\lambda_j\pi)((d_a-d^0_a)-(d_b-d^0_b))/2}\lambda_j^{(d_a-d^0_a)+(d_b-d^0_b)}h_{ab,j}(d-d^0,\theta-\theta^0)$ where $h_{ab,j}(d^0,\theta^0) = 1/\sqrt{1+\theta^0_a\lambda_j^{-2d^0_a}}\sqrt{1+\theta^0_b\lambda^{-2d^0_b}}$. Applying summation by parts yields

$$\frac{1}{m}\sum_{k=1}^{m-1}\left[g_k(d-d^0,\theta-\theta^0) - g_{k+1}(d-d^0,\theta-\theta^0)\right]\sum_{j=1}^{k}\left[e^{i(\lambda_j-\pi)(d^0_a-d^0_b)/2}\lambda_j^{d^0_a+d^0_b}h_{ab,j}(d^0,\theta^0)w_{aj}w^*_{bj} - G^0_{ab}\right]$$
$$+ \frac{g_m(d-d^0,\theta-\theta^0)}{m}\sum_{j=1}^{m}\left[e^{i(\lambda_j-\pi)(d^0_a-d^0_b)/2}\lambda_j^{d^0_a+d^0_b}h_{ab,j}(d^0,\theta^0)w_{aj}w^*_{bj} - G^0_{ab}\right].$$

Given that $(\log\,\lambda_k)^{k_1} = O_p((\log\,m)^{k_1})$, $(\log\,\lambda_k)^{k_1} - (\log\,\lambda_{k+1})^{k_1} = O(k^{-1})$, $h_{ab,k}(d-d^0,\theta-\theta^0) = O(1)$, $h_{ab,k}(d-d^0,\theta-\theta^0) - h_{ab,k+1}(d-d^0,\theta-\theta^0) = O(k^{-1})$, $\lambda_k^{(d_a-d^0_a)+(d_b-d^0_b)} = O(1)$, $\lambda_k^{(d_a-d^0_a)+(d_b-d^0_b)} - \lambda_{k+1}^{(d_a-d^0_a)+(d_b-d^0_b)} = O(k^{-1})$, $(k/m)^{2k_2} = O(1)$, $(k/m)^{2k_2} - ((k+1)/m)^{2k_2} = O(k^{-1})$, $e^{i(\lambda_k-\pi)((d_a-d^0_a)-(d_b-d^0_b))/2} = O(1)$, and $e^{i(\lambda_k-\pi)((d_a-d^0_a)-(d_b-d^0_b))/2} - e^{i(\lambda_{k+1}-\pi)((d_a-d^0_a)-(d_b-d^0_b))/2} = O(k^{-1})$ we get for $g_k(d-d^0,\theta-\theta^0) - g_{k+1}(d-d^0,\theta-\theta^0) = O((\log\,m)^{k_1}k^{-1})$ and $g_m(d-d^0,\theta-\theta^0) = O((\log\,m)^{k_1})$. Together with Lemma 3 we get for Eq. (21)

$$O_p\left(\left((\log\,m)^{k_1}\frac{1}{m}\sum_{k=1}^{m}(k^{\beta}n^{-\beta} + k^{-1/2}\log\,k + k^{d^0_a+d^0_b}n^{-(d^0_a+d^0_b)})\right)\right)$$
$$= O_p((\log\,m)^{k_1}m^{\beta}n^{-\beta} + (\log\,m)^{k_1}m^{-1/2}\log\,m + (\log\,m)^{k_1}m^{d^0_a+d^0_b}n^{-(d^0_a+d^0_b)}) = o_p((\log\,m)^{-2}).$$

Next we need to prove the result for Eq. (22). We have the (a,b)-th element of the argument given in Eq. (22) by

$$\frac{1}{m}\sum_{j=1}^{m}(\log\,\lambda_j)^{k_1}\left(\frac{j}{m}\right)^{2k_2}\left[e^{i(\lambda_j-\pi)((d_a-d_a^0)-(d_b-d_b^0))/2}\lambda_j^{(d_a-d_a^0)+(d_b-d_b^0)}h_{ab,j}(d-d^0,\theta-\theta^0)-1\right]G_{ab}^0,$$

where we receive an upper bound for

$$\begin{aligned}
&e^{i(\lambda_j-\pi)((d_a-d_a^0)-(d_b-d_b^0))/2}\lambda_j^{(d_a-d_a^0)+(d_b-d_b^0)}h_{ab,j}(d-d^0,\theta-\theta^0)-1\\
&\le e^{i(\lambda_j-\pi)((d_a-d_a^0)-(d_b-d_b^0))/2}\lambda_j^{(d_a-d_a^0)+(d_b-d_b^0)}-1\\
&=(e^{i(\lambda_j-\pi)((d_a-d_a^0)-(d_b-d_b^0))/2}-1)\lambda_j^{(d_a-d_a^0)+(d_b-d_b^0)}+(\lambda_j^{(d_a-d_a^0)+(d_b-d_b^0)}-1)\\
&\le C(|d_a-d_a^0|+|d_b-d_b^0|)+C(|d_a-d_a^0|+|d_b-d_b^0|)\log\,m\\
&=O((\log\,m)^{-5}).
\end{aligned}$$

The first inequality uses the fact that $h_{ab,j}(d,\theta)<\infty$ because Θ is compact. The second inequality is caused by $|\lambda_j^{(d_a-d_a^0)+(d_b-d_b^0)}-1|/|(d_a-d_a^0)+(d_b-d_b^0)|\le|\log\,\lambda_j|m^{|d_a-d_a^0|+|d_b-d_b^0|}\le|\log\,j|m^{1/\log\,m}\le C\,\log\,j$ for some constant $C<\infty$.
Therefore, we have

$$\frac{1}{m}\sum_{j=1}^{m}(\log\,\lambda_j)^{k_1}\left(\frac{j}{m}\right)^{2k_2}\left[e^{i(\lambda_j-\pi)((d_a-d_a^0)-(d_b-d_b^0))/2}\lambda_j^{(d_a-d_a^0)+(d_b-d_b^0)}-1\right]G_{ab}^0=o_p((\log\,m)^{-2})$$

which completes the first part of the proof.

Turning now to the second part $B_n^{-1}H_2(d,\theta)B_n^{-1}=o_p(1)$. Lets define $q_j(d_a^0,\theta_a^0)$ as in Frederiksen et al. (2012) then we can express the a-th element in $B_n^{-1}H_2(d,\theta)B_n^{-1}$ as

$$\frac{1}{m}\sum_{j=1}^{m}\frac{(j/m)^{2d_a}}{2\sqrt{1+\theta_a\lambda_j^{2d_a}}}q_j(d_a,\theta_a),$$

where $q_j(d_a,\theta_a)=O((\log\,n)^2)$ and $q_j(d,\theta)-q_{j-1}(d,\theta)=O(j^{-1}(\log\,n))$. In order to prove the second part, we need to show

$$\sup_{d\in D(\tau),\theta\in\Theta}\left\|\frac{1}{m}\sum_{j=1}^{m}(j/m)^{d_a+d_b}h_{ab,j}(d,\theta)q_{ab,j}(d,\theta)-(j/m)^{d_a^0+d_b^0}h_{ab,j}(d^0,\theta^0)q_{ab,j}(d^0,\theta^0)\right\|=o_p(1). \tag{23}$$

Rewriting yields

$$\sup_{d\in D(\tau),\theta\in\Theta}\left\|\frac{1}{m}\sum_{j=1}^{m}\left(\left(\frac{j}{m}\right)^{d_a^0+d_b^0}-\left(\frac{j}{m}\right)^{d_a+d_b}\right)h_{ab,j}(d-d^0,\theta-\theta^0)q_{ab,j}(d-d^0,\theta-\theta^0)\right\|,$$

where we have that $h_{ab,j}(d-d^0,\theta-\theta^0)=O((j/n)^{d_a+d_b})$ and $q_j(d_a,\theta_a)=O((\log\,\mathrm{n})^2)$. We know from the mean value theorem that

$$\sup_{d\in D(\tau),j=1,\ldots,m}\left|\left(\frac{j}{m}\right)^{d_a^0+d_b^0}-\left(\frac{j}{m}\right)^{d_a+d_b}\right| = O(\sup_{d\in D(\tau)}((d_a^0+d_b^0)-(d_a+d_b))\log m)$$
$$= O((\log n)^{-6}\tau\log m)$$

with $\tau = o(1)$. All in all we get for Eq. (23)

$$O_p\left((\log n)^{-6}\tau\log m\left(\frac{m}{n}\right)^{d_a+d_b}(\log n)^2\right) = O_p\left((\log n)^{-4}\log m\left(\frac{m}{n}\right)^{d_a+d_b}\right) = o_p(1)$$

which concludes the proof.

Proof of (c)

In order to prove part (c) of Lemma 1, we proceed with the score in the same way as Shimotsu (2007) and write

$$\frac{1}{\sqrt{q}}\sum_{a=1}^{q}\left\{-\frac{2}{m}\sum_{j=1}^{m}X_{ja}+\text{tr}\left[\hat{G}(d_a,\theta_a)^{-1}Q_{1a}\right]\right\}+\text{tr}\left[\hat{G}(d_a,\theta_a)^{-1}Q_{2a}\right] = R_1+R_2,$$

so that we need to find an approximation for R_1 and R_2. The approximation for R_2 is a straightforward adaption of the proof used by Shimotsu (2007) and is therefore omitted. Focusing now on the remainder term R_1. Define $\zeta_{ja} = (\widetilde{X}_{ja} - \frac{1}{m}\sum_{k=1}^{m}\widetilde{X}_{ka})$, ignoring smaller order terms, we get

$$-\frac{2}{\sqrt{m}}\sum_{j=1}^{m}X_{ja}^0+\text{tr}\left[\hat{G}(d^0,\theta^0)^{-1}Q_{1a}\right]$$
$$= \text{tr}\left[\hat{G}(d^0,\theta^0)^{-1}\left(Q_{1a}-\frac{2}{m}\sum_{j=1}^{m}X_{ja}\hat{G}(d^0,\theta^0)i_aX_{ja}\right)\right]$$
$$= \text{tr}\left[\hat{G}(d^0,\theta^0)^{-1}\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\zeta_{ja}^0\text{Re}\left[(\Lambda_j^0)^{-1}I_{zj}(\Lambda_j^{0*})^{-1})\right]i_a\right]\zeta_{ja}^0$$
$$= g_a\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\zeta_{ja}^0\left\{\text{Re}\left[(\Lambda_j^0)^{-1}I_{zj}(\Lambda_j^{0*})^{-1}\right]\right\}_a\zeta_{ja}^0,$$

with g_a representing the a-th row of $(G^0)^{-1}$ and $\{\cdot\}_a$ gives the a-th column of a matrix. It follows from summation by parts and Lemma 2

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\zeta_{ja}^0(\Lambda_j^0)^{-1}I_{zj}(\Lambda_j^{0*})^{-1}\zeta_{ja}^0$$
$$= \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\zeta_{ja}^0\left[(\Lambda_j^0)^{-1}A(\lambda_j)I_{\epsilon j}A^*(\lambda_j)(\Lambda_j^{0*})^{-1}-G^0\right]$$

Which leads to

$$R_1 = \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \zeta^0_{ja} \left(g_a \left\{ \text{Re} \left[(\Lambda^0_j)^{-1} I_{zj} (\Lambda^{0*}_j)^{-1} \right]_a \right\} - 1 \right) \zeta^0_{ja}. \tag{24}$$

We can further rewrite the right hand side of Eq. (24) as $T_{1a} + T_{2a} + T_{3a} + T_{4a}$ defining the single elements in the summation as

$$\begin{aligned}
T_{1a} &= \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \zeta^0_{ja} \Big(g_a \left\{ \text{Re}\left[(\Lambda^0_j)^{-1} I_{zj} (\Lambda^{0*}_j)^{-1} \right] - \text{Re}\left[(\Lambda^0_j)^{-1} A(\lambda_j) I_{\epsilon j} A(\lambda_j) (\Lambda^{0*}_j)^{-1} \right] \right\}_a \\
&\quad - E\Big(g_a \left\{ \text{Re}\left[(\Lambda^0_j)^{-1} I_{zj} (\Lambda^{0*}_j)^{-1} \right] - \text{Re}\left[(\Lambda^0_j)^{-1} A(\lambda_j) I_{\epsilon j} A(\lambda_j) (\Lambda^{0*}_j)^{-1} \right] \right\}_a \Big) \Big) \zeta^0_{ja}, \\
T_{2a} &= \zeta^0_{ja} \left(E\left[w_{aj} w^*_{aj} / f_{z,aa}(\lambda_j) \right] - 1 \right) \left(g_a \left\{ \text{Re}\left[(\Lambda^0_j)^{-1} f_z(\lambda_j) (\Lambda^{0*}_j)^{-1} \right] \right\}_a \right) \zeta^0_{ja}, \\
T_{3a} &= \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \zeta^0_{ja} \left(g_a \left\{ \text{Re}\left[(\Lambda^0_j)^{-1} A(\lambda_j) I_{\epsilon j} A^*(\lambda_j) (\Lambda^{0*}_j)^{-1} \right] \right\}_a - 1 \right) \zeta^0_{ja}, \\
T_{4a} &= \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \zeta^0_{ja} \left(g_a \left\{ \text{Re}\left[(\Lambda^0_j)^{-1} f_z(\lambda_j) (\Lambda^{0*}_j)^{-1}) \right] \right\}_a - 1 \right) \zeta^0_{ja}.
\end{aligned}$$

Now, we need to show that $T_{3a} \xrightarrow{d} N(0, \mathbf{\Omega})$, while the remaining terms are $o_p(1)$. The proof for T_{3a} follows directly from Shimotsu (2007) by replacing ν_j with ζ_j and using the result in part (a) of Lemma 1.

In order to prove the result for T_{1a} we use summation by parts:

$$\begin{aligned}
T_{1a} &= \frac{2}{\sqrt{m}} \sum_{k=1}^{m-1} \left(\widetilde{X}^0_{k,a} - \widetilde{X}^0_{k+1,a} \right) \sum_{j=1}^{k} \Big(g_a \left\{ \text{Re}\left[(\Lambda^0_j)^{-1} I_{zj} (\Lambda^{0*}_j)^{-1} \right] - \text{Re}\left[(\Lambda^0_j)^{-1} A(\lambda_j) I_{\epsilon j} A(\lambda_j) (\Lambda^{0*}_j)^{-1} \right] \right\}_a \\
&\quad - E\Big(g_a \left\{ \text{Re}\left[(\Lambda^0_j)^{-1} I_{zj} (\Lambda^{0*}_j)^{-1} \right] - \text{Re}\left[(\Lambda^0_j)^{-1} A(\lambda_j) I_{\epsilon j} A(\lambda_j) (\Lambda^{0*}_j)^{-1} \right] \right\}_a \Big) \Big) \sum_{k=1}^{m-1} \left(\widetilde{X}^0_{k,a} - \widetilde{X}^0_{k+1,a} \right) \\
&\quad + \zeta^0_{ma} \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \Big(g_a \left\{ \text{Re}\left[(\Lambda^0_j)^{-1} I_{zj} (\Lambda^{0*}_j)^{-1} \right] - \text{Re}\left[(\Lambda^0_j)^{-1} A(\lambda_j) I_{\epsilon j} A(\lambda_j) (\Lambda^{0*}_j)^{-1} \right] \right\}_a \\
&\quad - E\Big(g_a \left\{ \text{Re}\left[(\Lambda^0_j)^{-1} I_{zj} (\Lambda^{0*}_j)^{-1} \right] - \text{Re}\left[(\Lambda^0_j)^{-1} A(\lambda_j) I_{\epsilon j} A(\lambda_j) (\Lambda^{0*}_j)^{-1} \right] \right\}_a \Big) \Big) \zeta^0_{ma} \\
&= \frac{2}{\sqrt{m}} \sum_{k=1}^{m} O(k^{-1}) + O_p(k^{1/2} \log k + k^{\beta+1/2} n^{-\beta} + k^{1/2+2d^0_a} m^{-2d^0_a}) \\
&\quad + O(1) 2 m^{1/2} O_p(m^{1/2} \log m + m^{\beta+1/2} n^{-\beta} + m^{1/2+2d^0_a} n^{-2d^0_a}) \\
&= O_p((\log m)^2 + (m/n)^{\min\{\beta, 2d^0_a\}}).
\end{aligned}$$

The proof uses Lemma 2 and that $\widetilde{X}^0_{k,a} - \widetilde{X}^0_{k+1,a} = O(k^{-1})$ uniformly over $k = 1, ..., m$ and $\widetilde{X}^0_{m,a} - \frac{1}{m} \sum_{k=1}^{m} \widetilde{X}^0_{k,a} = O(1)$, which is a result by approximating sums by integrals, for the second equation. Caused by the fact that d^0_a belongs to the interior of the parameter space, it follows that $T_{1a} = o_p(1)$.

Next we show the result for T_{2a}. Therefore, we use Theorem 2 of Robinson (1995b) yielding $EI_{yj} = f_{yj}\{1 + O(j^{-1} \log(j+1)\}$ uniformly over $j = 1, ..., m$ and the result of Frederiksen et al. (2012) that $EI_{zj} = f_{zj}\{1 + O(j^{-1} \log(j+1)\}$ and further that $g_a \left\{ \text{Re}\left[\Lambda_j(d^0, \theta^0)^{-1} f_z(\lambda_j) \Lambda^*_j(d^0, \theta^0)^{-1} \right] \right\}_a -$

$1 = O((j/n)^{\beta} + (j/n)^{2d_a^0})$ so that T_{2a} can be bounded by the same argument as in Frederiksen et al. (2012) by $T_{2a} = O((\log m)^3 m^{-1/2})$ which is $o(1)$ because $d_a^0 < \Delta_2 < 1/2$.
The proof for T_{4a} uses the accuracy of the approximation of $f_z(\lambda_j)$ by the periodogram and summation by parts and yields the result

$$
\begin{aligned}
T_{4a} &= \frac{2}{\sqrt{m}} \sum_{k=1}^{m-1} \left(\widetilde{X}_{k,a}^0 - \widetilde{X}_{k+1,a}^0 \right) \sum_{j=1}^{k} \left(g_a \left\{ \mathrm{Re}\left[(\Lambda_j^0)^{-1} f_{zj} (\Lambda_j^{0*})^{-1} \right] \right\}_a - 1 \right) \sum_{k=1}^{m-1} \left(\widetilde{X}_{k,a}^0 - \widetilde{X}_{k+1,a}^0 \right) \\
&+ \zeta_{ma}^0 \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \left(g_a \left\{ \mathrm{Re}\left[(\Lambda_j^0)^{-1} f_{zj} (\Lambda_j^{0*})^{-1} \right] \right\}_a - 1 \right) \zeta_{ma}^0 \\
&= \frac{2}{\sqrt{m}} \sum_{k=1}^{m-1} O(k^{-1}) \sum_{j=1}^{m} O((j/n)^{\beta} + (j/n)^{2d_a^0}) + O(1) \frac{2}{\sqrt{m}} \sum_{j=1}^{m} O((j/n)^{\beta} + (j/n)^{2d_a^0}) \\
&= O((m/n)^{\beta} m^{1/2} + (m/n)^{2d_a^0} m^{1/2})
\end{aligned}
$$

which is $o_p(1)$ by Assumption 6.

Proof of Theorem 3

Based on the results given above we can contemplate a Taylor expansion

$$
\frac{1}{\sqrt{q}} S(\hat{d}, \hat{\theta})^r = \frac{1}{\sqrt{q}} S(d^0, \theta^0)^r + \frac{1}{\sqrt{q}} H(\bar{d}, \bar{\theta})^r ((\hat{d}, \hat{\theta}) - (d^0, \theta^0))
$$

with $||(\bar{d}, \bar{\theta}) - (d^0, \theta^0)|| \leq ||(\hat{d}, \hat{\theta}) - (d^0, \theta^0)||$ and $S(d, \theta)^r$ and $H(d, \theta)^r$ being defined as in Eq. (19) and Eq. (20), but with the summation being executed until $\lfloor mr \rfloor$ instead of m to prove Theorem 3.
We can rewrite the first part of the right hand side of the equation as

$$
\frac{1}{\sqrt{q}} \sum_{a=1}^{q} S(d^0, \theta^0)^r = \frac{1}{\sqrt{q}} R_1^r + o_p(1) + \mathrm{tr}\left[\frac{\hat{G}(d_a^0, \theta_a^0)^{-1}}{\sqrt{m}} \sum_{j=1}^{\lfloor mr \rfloor} L_j \mathrm{Im}\left[(\Lambda_j^0)^{-1} (-i_a I_{zj} + I_{zj} i_a) (\Lambda_j^{0*})^{-1} \right] \right] \tag{25}
$$

$$
\begin{aligned}
&= \frac{2}{\sqrt{m}} \sum_{a=1}^{q} \frac{(g_a)^{-1}}{\sqrt{q}} \sum_{j=1}^{\lfloor mr \rfloor} \zeta_{ja}^0 \left[g_a' g_a \left\{ \mathrm{Re}\left[\left(\Lambda_j^0\right)^{-1} I_{zj} \left(\Lambda_j^{0*}\right)^{-1} \right] \right\}_a \zeta_{ja}^0 - 1 \right] \\
&- \frac{2(g_a)^{-1}}{m^{3/2}} \left(\sum_{j=1}^{\lfloor mr \rfloor} \zeta_{ja}^0 \right) \sum_{j=1}^{m} \left[g_a' g_a \left[\mathrm{Re}\left\{ \left(\Lambda_j^0\right)^{-1} A(\lambda_j) I_{\epsilon j} A^*(\lambda_j) \left(\Lambda_j^{0*}\right)^{-1} \right] \right\}_a - 1 \right] \\
&\times \left(\sum_{j=1}^{\lfloor mr \rfloor} \zeta_{ja}^0 \right) + o_p(1) + \mathrm{tr}\left[\frac{\hat{G}(d_a^0, \theta_a^0)^{-1}}{\sqrt{m}} \sum_{j=1}^{\lfloor mr \rfloor} L_j \mathrm{Im}\left[\left(\Lambda_j^0\right)^{-1} (-i_a I_{zj} + I_{zj} i_a) \left(\Lambda_j^{0*}\right)^{-1} \right] \right].
\end{aligned}
$$

We can represent the first part in the equation as before as $T_{1a}^r + T_{2a}^r + T_{3a}^r + T_{4a}^r$. In the same way it can be shown that T_{1a}^r, T_{2a}^r and T_{4a}^r converge to 0. So we can focus on the remaining term T_{3a}^r, which can be rewritten together with the imaginary unit in the same way as in Shimotsu (2007) and Sibbertsen et al. (2018) by $\sum_{t=1}^{n} z_t + o_p(1)$ with $z_1 = 0$ and $z_t = \epsilon_t' \sum_{s=1}^{n-1} \left[\Phi_{t-s} + \tilde{\Phi}_{t-s} \right] \epsilon_s$.
Let $\Phi_s = \frac{1}{\pi\sqrt{mn}} \sum_{j=1}^{m} \zeta_j \left[\mathrm{Re}\left[\psi_j + \psi_j' \right] \cos(s\lambda_j) \right] \zeta_j$, $\tilde{\Phi}_s = \frac{\pi}{2} \frac{1}{\pi\sqrt{mn}} \sum_{j=1}^{m} \mathrm{Re}\left[\psi_j - \psi_j' \right] \sin(s\lambda_j)$, where

$\psi_j = \sum_{a=1}^{q} \left[A^*(\lambda_j) \left(\Lambda_j^{0*}\right)^{-1} \right]_a g_a \left(\Lambda_j^0\right)^{-1} A(\lambda_j)$, $A(\lambda) = \sum_{j=0}^{\infty} A_j \epsilon_{t-j}$ and A_j is given in Assumption 3. The asymptotic normality of z_t follows directly from Theorem 2 of Robinson (1995a) by replacing ν_j by ζ_j. We can handle the covariance of z_t over $0 \leq r_1 \leq r_2 \leq 1$ in the same way as in Sibbertsen et al. (2018) where we need to replace $\sum_{j=1}^{\lfloor mr \rfloor} \nu_j$ by $\sum_{j=1}^{\lfloor mr \rfloor} \zeta_j$ again. Therefore, we have by applying Lemma 3 for $\frac{1}{m} \sum_{j=1}^{\lfloor mr_1 \rfloor} \zeta_j \zeta_j' \to \int_0^{r_1} \mathbf{\Psi}^2(r) ds$ which yields

$$\operatorname{Cov}\left(\sum_{t=1}^{n} z_{t,r_1}, \sum_{t=1}^{n} z_{t,r_2} \right) \to \int_0^{r_1} \mathbf{\Psi}^2(r) ds,$$

with $\mathbf{\Psi}(r)$ being defined as in Theorem 3.

Turning now to the second term of the second equality in Eq. (25). Again, we can use the same arguments as in Sibbertsen et al. (2018) and Qu (2011) but adapting for ζ_j and the new asymptotic behaviour of the spectral density near the origin. Together with the results given in Lemma 2 we achieve

$$\frac{(g_a)^{-1}}{\sqrt{q}} \sum_{j=1}^{m} \left[g_a' g_a \left\{ \left(\Lambda_j^0\right)^{-1} I_{zj} \left(\Lambda_j^*\right)^{-1} \right\}_a - 1 \right] \Rightarrow B(1),$$

with $B(s)$ being a standard Brownian motion. For ζ_j we have by the use of Lemma 3 the convergence against $\int_0^{r_1} \mathbf{\Psi}(r_1) ds$.

Now we have to handle the second part of the Taylor expansion. We need to focus on the Hessian where the summation is again executed until $\lfloor mr \rfloor$ rather than m. Based on the foregoing proofs, we can write it as

$$\begin{aligned} H^r(d,\theta) &= H_1^r(d,\theta) + H_2^r(d,\theta) \\ &= \operatorname{tr}\left[\sum_{a=1}^{q} \frac{1}{m} \sum_{j=1}^{\lfloor mr \rfloor} \zeta_{ja} \zeta_{jb}' \hat{G}(d,\theta)^{-1} \hat{G}_{2ab}(d,\theta) + \left(\left(\frac{\pi}{2}\right)^2, 0 \right)' \hat{G}(d,\theta)^{-1} \hat{G}_{3ab}(d,\theta) \right] \\ &\quad + \left(\frac{-2}{m} \sum_{j=1}^{\lfloor mr \rfloor} \nabla X_j \right). \end{aligned}$$

The second part of the Hessian $H_2^r(d,\theta)$ converges to zero. The proof can be done in the same way as before in Lemma 1. The remaining part of the Hessian is behaving asymptotically as

$$\begin{aligned} H_1^r(d,\theta) &= \operatorname{tr}\left[\sum_{a=1}^{q} \frac{1}{m} \sum_{j=1}^{\lfloor mr \rfloor} \zeta_{ja} \zeta_{jb}' \hat{G}(d,\theta)^{-1} \hat{G}_{2ab}(d,\theta) + \left(\left(\frac{\pi}{2}\right)^2, 0 \right)' \hat{G}(d,\theta)^{-1} \hat{G}_{3ab}(d,\theta) \right] \\ &\Rightarrow \operatorname{tr}\left[\int_0^r \mathbf{\Psi}^{2,0}(r) ds \left(G^0\right)^{-1} G_{2ab}^0 + \left(\frac{\pi^2}{4}, 0 \right)' \left(G^0\right)^{-1} G_{3ab}^0 \right]. \end{aligned}$$

We know by Lemma 2 of Sibbertsen et al. (2018) that the single elements of G^0 given in the Hessian collapse in such a way that we are left with the asymptotic convergence result of the

Hessian by

$$H^r(d,\theta) = H_1^r(d,\theta) \Rightarrow \int_0^r \mathbf{\Psi}^{2,0}(r)ds = F(r).$$

The last step is proving tightness. As in Qu (2011) and Sibbertsen et al. (2018) we use Theorem 13.5 of Billingsley (1999). We need to show that for every m and $r_1 \leq r \leq r_2$

$$\mathrm{E}\left(\left|\sum_{t=1}^n z_{t,r} - \sum_{t=1}^n z_{t,r_1}\right|^2 \left|\sum_{t=1}^n z_{t,r_2} - \sum_{t=1}^n z_{t,r}\right|^2\right) \leq K(\phi_m(r_2) - \phi_m(r_1))^2$$

where $\phi_m(\cdot)$ is a finite, non-decreasing function over $[0,1]$ that fulfils $\lim\limits_{\delta\to 0} \limsup\limits_{m\to\infty} |\phi_m(s+\delta) - \phi(s)| \to 0$ uniformly over the set $s \in [0,1]$ and K is some constant. For the sake of conciseness we write $z_t(s,r) = z_{t,r} - z_{t,s}$, $c_t(r,s) = c_{t,r} - c_{t,s}$ and $c_t = \mathrm{tr}\left[\Phi_t + \tilde{\Phi}_t\right]$. We can use Lemma B.8. from Qu (2011) in order to prove that $K\left(\sum_{t=1}^n \sum_{s=1}^{t-1} c_{t-s}(r_1,r)^2\right)\left(\sum_{t=1}^n \sum_{h=1}^{t-1} c_{t-h}(r,r_2)^2\right)$ is an upper bound for $\mathrm{E}\left(|\sum_{t=1}^n z_{t,r} - \sum_{t=1}^n z_{t,r_1}|^2 |\sum_{t=1}^n z_{t,r_2} - \sum_{t=1}^n z_{t,r}|^2\right)$ with K being some positive constant.
Next, we have

$$\sum_{t=1}^n \sum_{s=1}^{t-1} c_{t-s}(r_1,r)^2 \leq \frac{1}{nm} \sum_{j=\lfloor mr_1\rfloor+1}^{\lfloor mr\rfloor} \sum_{k\neq j}^{\lfloor mr\rfloor} (\zeta_j\zeta_j' + \zeta_k\zeta_k') + \frac{1}{m} \sum_{j=\lfloor mr_1\rfloor+1}^{\lfloor mr\rfloor} \zeta_j\zeta_j' \leq \frac{3}{m} \sum_{j=\lfloor mr_1\rfloor+1}^{\lfloor mr\rfloor} \zeta_j\zeta_j'.$$

The same holds true for $\sum_{t=1}^n \sum_{s=1}^{t-1} c_{t-h}(r,r_2)^2 \leq \frac{3}{m}\sum_{j=\lfloor mr\rfloor+1}^{\lfloor mr_2\rfloor} \zeta_j\zeta_j'$. If $\phi_m(s) = \frac{1}{m}\sum_{j=1}^{\lfloor ms\rfloor} \zeta_j\zeta_j'$ then we have for

$$\lim_{\delta\to 0} \limsup_{m\to\infty} |\phi_m(s+\delta) - \phi(s)| = \lim_{\delta\to 0} \int_s^{s+\delta} \mathbf{\Psi}^2(r)dx \to 0.$$

This completes the proof.

Proof of Theorem 4

In order to prove the consistency of our testing procedure we orientate on the proof of Sibbertsen et al. (2018). Therefore, to analyse the test statistic we split into two parts accordingly to

$$\begin{aligned}
MLWNS = \frac{1}{2}\sup_{r\in[\epsilon,1]} \Bigg|\Bigg| &\frac{2}{\sqrt{m}} \sum_{a=1}^q \frac{1}{\sqrt{q}} \sum_{j=1}^{\lfloor mr\rfloor} \zeta_{ja}\left(g_a\left\{\mathrm{Re}\left[\left(\Lambda_j(\hat{d},\hat{\theta})\right)^{-1} I_{zj}\left(\Lambda_j^*(\hat{d},\hat{\theta})\right)^{-1}\right]_a\right\} - 1\right)\zeta_{ja} \\
&+ \frac{1}{\sqrt{m}} \sum_{a=1}^q \frac{g_a}{\sqrt{q}} \sum_{j=1}^{\lfloor mr\rfloor} L_j \mathrm{Im}\left[\left(\Lambda_j(\hat{d},\hat{\theta})\right)^{-1} I_{zj}\left(\Lambda_j^*(\hat{d},\hat{\theta})\right)^{-1}\right]_a \Bigg|\Bigg| \\
= \sup_{r\in[\epsilon,1]} &||R + L|| . \qquad (26)
\end{aligned}$$

Let us focus on the first term R in Eq. (26). Note that ζ_{ja} is a monotonically increasing function in j with $\zeta_{1a} < 0$ and $\zeta_{ma} > 0$ and further define $j^* = \min\{j : \zeta_{ja} \geq 0\}$. This proof uses the fact that the divergence of the quantity in the MLWNS statistic for at least one r given in the a-th

time series implies convergence of the supremum over all r. This implies that it is sufficient to focus only on the special case were $r = 1$.
From Lemma 3 it follows that $j^* = Km$, where K is some constant. To ease the notation we will write $\left(\Lambda_j(\hat{d}, \hat{\theta})\right)^{-1} = (\Lambda_j)^{-1}$. We proceed with

$$R^I = \left\|\frac{1}{\sqrt{q}} \sum_{a=1}^{q} \left(\left(\sum_{j=1}^{m} \zeta_{ja}\zeta_{jb}' \right) \left(\sum_{j=1}^{m} \zeta_{ja}\zeta_{jb}' \right) \right)^{-1/2} \sum_{j=j^*}^{m} \zeta_{ja} \left(g_a \left\{ \mathrm{Re} \left[(\Lambda_j)^{-1} I_{zj} \left(\Lambda_j^*\right)^{-1} \right] \right\}_a - 1 \right) \zeta_{ja} \right\| \tag{27}$$

and

$$R^{II} = \left\|\frac{1}{\sqrt{q}} \sum_{a=1}^{q} \left(\left(\sum_{j=1}^{m} \zeta_{ja}\zeta_{jb}' \right) \left(\sum_{j=1}^{m} \zeta_{ja}\zeta_{jb}' \right) \right)^{-1/2} \sum_{j=1}^{j^*-1} \zeta_{ja} \left(g_a \left\{ \mathrm{Re} \left[(\Lambda_j)^{-1} I_{zj} \left(\Lambda_j^*\right)^{-1} \right] \right\}_a - 1 \right) \zeta_{ja} \right\|.$$

When applying the reverse triangle inequality together with the results given in Sibbertsen et al. (2018), we know that it is sufficient to show that $R^I \xrightarrow{p} \infty$ if $n \to \infty$ caused by R being bounded from above by $R^I/2$. Hence, we rewrite Eq. (27) as

$$\begin{aligned} R^I = & \left\|\frac{1}{\sqrt{q}} \sum_{a=1}^{q} \left(\left(\sum_{j=1}^{m} \zeta_{ja}\zeta_{jb}' \right) \left(\sum_{j=1}^{m} \zeta_{ja}\zeta_{jb}' \right) \right)^{-1/2} \sum_{j=j^*}^{m} \zeta_{ja} \left(g_a \left\{ \mathrm{Re} \left[(\Lambda_j)^{-1} A(\lambda_j) I_{\epsilon j} A^*(\lambda_j) \right.\right.\right.\right. \\ & \left.\left.\left.\left. \times \left(\Lambda_j^*\right)^{-1} \right] \right\}_a - 1 \right) \zeta_{ja} \right\|. \end{aligned}$$

Applying the reverse triangle inequality yields

$$\begin{aligned} R^I \geq & \frac{1}{\sqrt{q}} \sum_{a=1}^{q} \left(\left(\sum_{j=1}^{m} \zeta_{ja}\zeta_{jb}' \right) \left(\sum_{j=1}^{m} \zeta_{ja}\zeta_{jb}' \right) \right)^{-1/2} \sum_{j=j^*}^{m} \zeta_{ja} \\ & - \frac{1}{\sqrt{q}} \sum_{a=1}^{q} \sum_{j=j^*}^{m} \zeta_{ja} \left(g_a \left\{ \mathrm{Re} \left[(\Lambda_j)^{-1} A(\lambda_j) I_{\epsilon j} A^*(\lambda_j) \left(\Lambda_j^*\right)^{-1} \right] \right\}_a \right) \zeta_{ja}. \end{aligned}$$

We know from Lemma 3 that the first term has for each component the form

$$\left(\left(\sum_{j=1}^{m} \zeta_{ja}\zeta_{jb}' \right) \left(\sum_{j=1}^{m} \zeta_{ja}\zeta_{jb}' \right) \right)^{-1/2} \sum_{j=j^*}^{m} \zeta_{ja} = m^{-1/2} \int_K^1 \Psi(s) ds + o(m^{1/2})$$

which is strictly positive and of order $m^{1/2}$.
Focusing now on the second term, using that $m/n^{1/2} \to \infty$ it needs to hold for $j^*/n^{1/2} \to \infty$. As a result, $I(\lambda_j) = O_p(1)$ for every $j^* \leq j \leq m$. In addition, we have

$$\begin{aligned} \zeta_{ja} \left(g_a \left\{ \mathrm{Re} \left[(\Lambda_j)^{-1} A(\lambda_j) I_{\epsilon j} A^*(\lambda_j) \left(\Lambda_j^*\right)^{-1} \right] \right\}_a \right) \zeta_{ja} &= O(\log(m)) O_p \left(\lambda_j^{\min\left(\beta, \hat{d}_a - d_a^0 + \min_b(\hat{d}_b - d_b^0)\right)} \right) \\ &= O_p(\log(m)) \qquad \forall a = 1, ..., q \text{ and } \forall b = 1, ..., q. \end{aligned}$$

This holds by applying Lemma 2 for $j^* \leq j \leq m$ and in addition it is caused by the positive definiteness of $\hat{G}(d,\theta)$, $P(\hat{d}_a - d_a^0 \geq 0) \to 1\ \forall a$ and $\lambda_j = o(1)$. As a consequence, the second term is asymptotically dominated by the first one. All in all we have $R^I \xrightarrow{p} \infty$ if $n \to \infty$.
Turning now to the second Term L in Eq. (26). It can be treated in the exact same way as R before. We write

$$L^I = \left\| \frac{1}{\sqrt{q}} \sum_{a=1}^{q} g_a \left(\left(\sum_{j=1}^{m} \zeta_{ja} \zeta_{jb}' \right) \left(\sum_{j=1}^{m} \zeta_{ja} \zeta_{jb}' \right) \right)^{-1/2} \sum_{j=j^*}^{m} L_j \mathrm{Im} \left[(\Lambda_j)^{-1} I_{zj} \left(\Lambda_j^*\right)^{-1} \right]_a \right\| \tag{28}$$

and

$$L^{II} = \left\| \frac{1}{\sqrt{q}} \sum_{a=1}^{q} g_a \left(\left(\sum_{j=1}^{m} \zeta_{ja} \zeta_{jb}' \right) \left(\sum_{j=1}^{m} \zeta_{ja} \zeta_{jb}' \right) \right)^{-1/2} \sum_{j=1}^{j^*-1} L_j \mathrm{Im} \left[(\Lambda_j)^{-1} I_{zj} \left(\Lambda_j^*\right)^{-1} \right]_a \right\|.$$

Again, we apply the reverse triangle inequality and get as a result that the second term in Eq. (26) is bounded from above by Eq. (28). Hence, we can proceed in the same way as before, meaning that the same arguments as before can be applied to Eq. (27). We see that the term L is of lower order than $m^{1/2}$. This implies that L is asymptotically dominated by R in the limit. This proves the theorem.

Technical Lemmas

Lemma 2. *Let $h_{ab,j}(d^0,\theta^0) = 1/(1+\theta_a^0 \lambda_j^{-2d_a^0})^{1/2}(1+\theta_b^0 \lambda^{-2d_b^0})^{1/2}$ and that $A_a(\lambda_j)$ defines the a-th row of $A(\lambda_j) = \sum_{k=0}^{\infty} A_k e^{ik\lambda_j}$ and $A_b^*(\lambda_j)$ is the b-th column of the considered matrix. Given that Assumptions 1-6 hold then as $n \to \infty$, for $1 \leq s < r \leq m$,*

$$\max_{a,b} \sum_{j=s}^{r} e^{i(\lambda_j - \pi)(d_a^0 - d_b^0)/2} \lambda_j^{d_a^0 + d_b^0} h_{ab,j}(d^0,\theta^0)(w_{aj} w_{bj}^* - A_a(\lambda_j) I_{\epsilon j} A_b^*(\lambda_j))$$
$$= O_p(r^{1/3}(\log r)^{2/3} + \log r + r^{1/2} n^{-1/4} + r^{d_a^0 + d_b^0} n^{-1/2(d_a^0 + d_b^0)} \log r + r^{1/2(d_a^0 + d_b^0)} n^{-1/2(d_a^0 + d_b^0)}),$$

$$\max_{a,b} \sum_{j=s}^{r} (e^{i(\lambda_j - \pi)(d_a^0 - d_b^0)/2} \lambda_j^{d_a^0 + d_b^0} h_{ab,j}(d^0,\theta^0) w_{aj} w_{bj}^* - G_{ab}^0) = O_p(r^{\beta+1} n^{-\beta} + r^{1/2} \log r + r^{1 + d_a^0 + d_b^0} n^{-(d_a^0 + d_b^0)}),$$

and

$$\max_{a,b} \sum_{j=s}^{r} e^{i(\lambda_j - \pi)(d_a^0 - d_b^0)/2} \lambda_j^{d_a^0 + d_b^0} h_{ab,j}(d^0,\theta^0)(w_{aj} w_{bj}^* - A_a(\lambda_j) I_{\epsilon j} A_b^*(\lambda_j))$$
$$- E(e^{i(\lambda_j - \pi)(d_a^0 - d_b^0)/2} \lambda_j^{d_a^0 + d_b^0} h_{ab,j}(d^0,\theta^0)(w_{aj} w_{bj}^* - A_a(\lambda_j) I_{\epsilon j} A_b^*(\lambda_j)))$$
$$= O_p(r^{1/3}(\log r)^{2/3} + \log r + r^{1/2} n^{-1/4}),$$

$$\max_{a,b} \sum_{j=s}^{r} (e^{i(\lambda_j - \pi)(d_a^0 - d_b^0)/2} \lambda_j^{d_a^0 + d_b^0} h_{ab,j}(d^0,\theta^0) w_{aj} w_{bj}^* - G_{ab}^0) - E(e^{i(\lambda_j - \pi)(d_a^0 - d_b^0)/2} \lambda_j^{d_a^0 + d_b^0} h_{ab,j}(d^0,\theta^0) w_{aj} w_{bj}^* - G_{ab}^0)$$
$$= O_p(r^{\beta+1/2} n^{-\beta} + r^{1/2} \log r + r^{1/2 + d_a^0 + d_b^0} n^{-(d_a^0 + d_b^0)}).$$

Proof of Lemma 2

In order to prove the results in Lemma 2 we start by decomposing the single entries inside the summation similarly as in Shimotsu (2007)

$$
\begin{aligned}
H_{1j} &= e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0}h_{ab,j}(d_0,\theta_0)[w_{aj}w_{bj}^* - A_a(\lambda_j)I_{\epsilon j}A_b^*(\lambda_j)],\\
H_{2j} &= e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0}h_{ab,j}(d_0,\theta_0)[A_a(\lambda_j)I_{\epsilon j}A_b^*(\lambda_j) - f_{ab}(\lambda_j)],\\
H_{3j} &= e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0}h_{ab,j}(d_0,\theta_0)f_{ab}(\lambda_j) - G_{ab}^0.
\end{aligned}
$$

Focusing on the first two results stated in the lemma. The smoothness conditions given in the assumptions give us directly the result $\max_{a,b}|\sum_{j=s}^r H_{3j}|= O(r^{\beta+1}n^{-\beta} + r^{1+d_a^0+d_b^0}n^{-(d_a^0+d_b^0)})$. Caused by the independence of $I_y(\lambda_j)$ and $I_w(\lambda_j)$ we can further decompose the terms into the signal and the noise part, and treat them separately. The results for the signal processes are directly applicable from Shimotsu (2007) so that $\max_{a,b}|\sum_{j=s}^r H_{1j,y}|= O_p(r^{1/3}(\log\, r)^{2/3} + \log\, r + r^{1/2}n^{-1/4})$ and $\max_{a,b}|\sum_{j=s}^r H_{2j,y}|= O_p(r^{1/2}\log\, r)$. Turning now to the contribution of the noise part, we know from Lemma 3 of Frederiksen et al. (2012) that the approximation of the noise part fulfils the bound $O((r/n)^{d_a^0+d_b^0})$ and that caused by the independence assumption of the signal and the noise term we further get the additional bound $O_p((r/n)^{d_a^0+d_b^0}(\log\, r + r^{1+\min\{\beta,1/2(d_a^0+d_b^0)\}}n^{-\min\{\beta,1/2(d_a^0,d_b^0)\}})$ giving the contribution to the last missing part in the first argument of the lemma. This completes the proof.

Lemma 3. *$1/m\sum_{j=1}^{[mr]}\zeta_j = \mathbf{\Psi}(x)$ and $1/m\sum_{j=1}^{[mr]}\zeta_j^2 = \mathbf{\Psi}^2(x)$, where $\mathbf{\Psi}(x) = (\Psi_1(x),\Psi_2(x))'$ with the single entries being defined as $\Psi_{1,a}(x) = \int_0^r \log\, x/(1+\theta_a x^{2d_a})dx - r\int_0^1 \log\, x/(1+\theta_a x^{2d_a})dx + O(\frac{1}{m^{1-\epsilon}})$ and $\Psi_{2,a}(x) = \int_0^r x^{2d_a}/(1+\theta_a x^{2d_a})dx - r\int_0^1 x^{2d_a}/(1+\theta_a x^{2d_a})dx + O(\frac{1}{m^{1-\varepsilon}})$ for $a = 1,\ldots,q$, uniformly in $r\in[0,1]$ with ε being some arbitrary small positive number.*

Proof of Lemma 3

We can prove Lemma 3 in the same way as in Qu (2011) by using the Euler–Maclaurin formula, which is given by

$$
\sum_{j=1}^{k} g(j) = \int_1^k g(x)dx + \frac{g(1)+g(k)}{2} + \frac{1}{12}(g'(k)-g'(1)) + R
$$

while R is bounded by $|R|\leq \frac{2}{(2\pi)^2}\int_1^k|g''(k)|dx$. We start the proof with the first expression given in ζ. Let $k=[mr]$ and apply the Euler–Maclaurin formula to the individual elements with $g_a(x) = \frac{\log(\frac{x}{m})}{1+\theta_a(\frac{x}{m})^{2d_a}}$, then we get for the a-th element of the first expression in ζ

$$
\begin{aligned}
\frac{1}{m}\sum_{j=1}^{[mr]}\frac{\log(\frac{x}{m})}{1+\theta_a(\frac{x}{m})^{2d_a}} &= \int_{1/m}^{r}\frac{\log(x)}{1+\theta_a x^{2d_a}}dx + \int_r^{[mr]/m}\frac{\log(x)}{1+\theta_a x^{2d_a}}dx + \frac{\frac{\log(r)-\log(m)}{1+\theta_a(\frac{m}{r})^{2d_a}} - \frac{\log(m)}{1+\theta_a(\frac{1}{m})^{2d_a}}}{2m}\\
&\quad + \frac{1}{12m}\left(\frac{1}{1+\theta_a(\frac{1}{m})^{2d_a}} - \frac{2d_a\theta_a(\frac{1}{m})^{2d_a}\log(\frac{1}{m})}{(1+\theta_a(\frac{1}{m})^{2d_a})^2}\right.\\
&\quad \left. - \left(\frac{1}{[mr](1+\theta_a r^{2d_a})} - \frac{2d_a\theta_a r^{2d_a-1}\log(r)}{m(1+\theta_a r^{2d_a})^2}\right)\right) + R,
\end{aligned}
$$

with $|R|\leq \frac{2}{(2\pi)^2}\int_1^{[mr]} \frac{\theta_a^2(\frac{x}{m})^{4d_a}(2(2d_a+1)d_a\log(\frac{x}{m})-4d_a-1)-2\theta_a(\frac{x}{m})^{2d_a}((2d_a-1)d_a\log(\frac{x}{m})+2d_a+1)-1}{x^2(\theta_a(\frac{x}{m})^{2d_a}+1)^3}dx = O(1)$. We have that $\frac{1}{m}\sum_{j=1}^{[mr]}\frac{\log(\frac{x}{m})}{1+\theta_a(\frac{x}{m})^{2d_a}} = \int_{1/m}^{r}\frac{\log(x)}{1+\theta_a x^{2d_a}}dx + O(\frac{\log(m)}{m}) = \int_0^r \frac{\log(x)}{1+\theta_a x^{2d_a}}dx + O(\frac{1}{m^{1-\varepsilon}})$. Thus, we get for the a-th entry

$$\frac{1}{m}\sum_{j=1}^{[rm]}\left(X_{1,ja}-\frac{1}{m}\sum_{j=1}^{m}X_{1,ja}\right) = \int_0^r \frac{\log(x)}{1+\theta_a x^{2d_a}}dx - r\int_0^1 \frac{\log(x)}{1+\theta_a x^{2d_a}}dx.$$

The other arguments can be shown in the same way and are therefore omitted.

References

Andrews, Donald WK and Patrik Guggenberger (2003). "A Bias–Reduced Log–Periodogram Regression Estimator for the Long–Memory Parameter". *Econometrica* 71(2), pp. 675–712.

Andrews, Donald WK and Yixiao Sun (2004). "Adaptive Local Polynomial Whittle Estimation of Long-Range Dependence". *Econometrica* 72(2), pp. 569–614.

Arteche, Josu (2004). "Gaussian Semiparametric Estimation in Long Memory in Stochastic Volatility and Signal Plus Noise Models". *Journal of Econometrics* 119(1), pp. 131–154.

Arteche, Josu (2006). "Semiparametric Estimation in Perturbed Long Memory Series". *Computational Statistics & Data Analysis* 51(4), pp. 2118–2141.

Billingsley, Patrick (1999). "Convergence of Probability Measures". *INC, New York* 2(2.4).

Breidt, F. Jay, Nuno Crato, and Pedro De Lima (1998). "The Detection and Estimation of Long Memory in Stochastic Volatility". *Journal of Econometrics* 83(1-2), pp. 325–348.

Davidson, James and Dooruj Rambaccussing (2015). "A Test of the Long Memory Hypothesis Based on Self-Similarity". *Journal of Time Series Econometrics* 7(2), pp. 115–141.

Deo, Rohit S and Clifford M Hurvich (2001). "On the log Periodogram Regression Estimator of the Memory Parameter in Long Memory Stochastic Volatility Models". *Econometric Theory* 17(4), pp. 686–710.

Dolado, Juan J., Jesus Gonzalo, and Laura Mayoral (2005). "What is What?: A Simple Time-Domain Test of Long-Memory vs. Structural Breaks". *Unpublished Manuscript, Department of Economics, Universidad Carlos III de Madrid.*

Frederiksen, Per, Frank S Nielsen, and Morten Ø Nielsen (2012). "Local Polynomial Whittle Estimation of Perturbed Fractional Processes". *Journal of Econometrics* 167(2), pp. 426–447.

Geweke, John and Susan Porter-Hudak (1983). "The Estimation and Application of Long Memory Time Series Models". *Journal of Time Series Analysis* 4(4), pp. 221–238.

Granger, Clive WJ and Zhuanxin Ding (1996). "Varieties of Long Memory Models". *Journal of Econometrics* 73(1), pp. 61–77.

Granger, Clive WJ and Namwon Hyung (2004). "Occasional Structural Breaks and Long Memory with an Application to the S&P 500 Absolute Stock Returns". *Journal of Empirical Finance* 11(3), pp. 399–421.

Haldrup, Niels and Robinson Kruse (2014). *Discriminating Between Fractional Integration and Spurious Long Memory.* Tech. rep. Department of Economics and Business Economics, Aarhus University.

Hurvich, Clifford M., Eric Moulines, and Philippe Soulier (2005). "Estimating Long Memory in Volatility". *Econometrica* 73(4), pp. 1283–1328.

Hurvich, Clifford M and Bonnie K Ray (2003). "The local Whittle Estimator of Long-Memory Stochastic Volatility". *Journal of Financial Econometrics* 1(3), pp. 445–470.

Leschinski, Christian, Michelle Voges, and Philipp Sibbertsen (2021). "A Comparison of Semiparametric Tests for Fractional Cointegration". *Statistical Papers* 62(4), pp. 1997–2030.

Lu, Yang K and Pierre Perron (2010). "Modeling and Forecasting Stock Return Volatility Using a Random Level Shift Model". *Journal of Empirical Finance* 17(1), pp. 138–156.

Lütkepohl, Helmut (1996). "Handbook of Matrices". *Wiley, New York*.

McCloskey, Adam and Pierre Perron (2013). "Memory Parameter Estimation in the Presence of Level Shifts and Deterministic Trends". *Econometric Theory* 29(6), pp. 1196–1237.

Nielsen, Frank S (2011). "Local Whittle Estimation of Multi-Variate Fractionally Integrated Processes". *Journal of Time Series Analysis* 32(3), pp. 317–335.

Nielsen, Morten Ø and Katsumi Shimotsu (2007). "Determining the Cointegrating Rank in Nonstationary Fractional Systems by the Exact Local Whittle Approach". *Journal of Econometrics* 141(2), pp. 574–596.

Ohanissian, Arek, Jeffrey R Russell, and Ruey S Tsay (2008). "True or Spurious Long Memory? A new Test". *Journal of Business & Economic Statistics* 26(2), pp. 161–175.

Perron, Pierre and Zhongjun Qu (2010). "Long-Memory and Level Shifts in the Volatility of Stock Market Return Indices". *Journal of Business & Economic Statistics* 28(2), pp. 275–290.

Qu, Zhongjun (2011). "A Test Against Spurious Long Memory". *Journal of Business & Economic Statistics* 29(3), pp. 423–438.

Robinson, Peter M (1995a). "Gaussian Semiparametric Estimation of Long Range Dependence". *The Annals of Statistics* 23(5), pp. 1630–1661.

Robinson, Peter M (1995b). "Log-Periodogram Regression of Time Series with Long Range Dependence". *The Annals of Statistics*, pp. 1048–1072.

Robinson, Peter M (2008). "Diagnostic Testing for Cointegration". *Journal of Econometrics* 143(1), pp. 206–225.

Shimotsu, Katsumi (2006). *Simple (but Effective) Tests of Long Memory Versus Structural Breaks*. Tech. rep. Queen's Economics Department Working Paper.

Shimotsu, Katsumi (2007). "Gaussian Semiparametric Estimation of Multivariate Fractionally Integrated Processes". *Journal of Econometrics* 137(2), pp. 277–310.

Shimotsu, Katsumi and Peter CB Phillips (2005). "Exact local Whittle Estimation of Fractional Integration". *The Annals of Statistics* 33(4), pp. 1890–1933.

Sibbertsen, Philipp, Christian Leschinski, and Marie Busch (2018). "A Multivariate Test Against Spurious Long Memory". *Journal of Econometrics* 203(1), pp. 33–49.

So, Mike KP and Susanna WY Kwok (2006). "A Multivariate Long Memory Stochastic Volatility Model". *Physica A: Statistical Mechanics and its Applications* 362(2), pp. 450–464.

Sun, Yixiao and Peter CB Phillips (2003). "Nonlinear Log-Periodogram Regression for Perturbed Fractional Processes". *Journal of Econometrics* 115(2), pp. 355–389.

Varneskov, Rasmus T and Pierre Perron (2018). "Combining Long Memory and Level Shifts in Modelling and Forecasting the Volatility of Asset Returns". *Quantitative Finance* 18(3), pp. 371–393.

Xu, Jiawen and Pierre Perron (2014). "Forecasting Return Volatility: Level Shifts with Varying Jump Probability and Mean Reversion". *International Journal of Forecasting* 30(3), pp. 449–463.