A Generalization of a Gaussian Semiparametric Estimator on Multivariate Long-Range Dependent Processes

Guilherme Pumi [a,*] **and Sílvia R.C. Lopes** [a]

Mathematics Institute
Federal University of Rio Grande do Sul

Abstract

In this paper we propose and study a general class of Gaussian Semiparametric Estimators (GSE) of the fractional differencing parameter in the context of long-range dependent multivariate time series. We establish large sample properties of the estimator without assuming Gaussianity. The class of models considered here satisfies simple conditions on the spectral density function, restricted to a small neighborhood of the zero frequency and includes important class of VARFIMA processes. We also present a simulation study to assess the finite sample properties of the proposed estimator based on a smoothed version of the GSE which supports its competitiveness.

Keywords: Fractional integration; Long-range dependence; Semiparametric estimation; Smoothed periodogram; Tapered periodogram; VARFIMA processes.

Mathematical Subject Classification (2010). Primary 62H12, 62F12, 62M10, 60G10, 62M15;

1 Introduction

Let $\boldsymbol{d} = (d_1, \cdots, d_q)' \in (-1/2, 1/2)^q$ and let $\mathcal{B}$ be the shift operator. Consider the q-dimensional weakly stationary process $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ obtained as a stationary solution of the difference equations

$$\underset{k\in\{1,\cdots,q\}}{\operatorname{diag}}\big\{(1-\mathcal{B})^{d_k}\big\}\big(\boldsymbol{X}_t - \mathbb{E}(\boldsymbol{X}_t)\big) = \boldsymbol{Y}_t, \tag{1.1}$$

where $\{\boldsymbol{Y}_t\}_{t=0}^{\infty}$ is a q-dimensional weakly stationary process whose spectral density function $f_{\boldsymbol{Y}}$ is bounded and bounded away from zero. Each coordinate process in (1.1) exhibits long-range dependence whenever the respective parameter $d_i > 0$, in the sense that the spectral density function satisfies $f(\lambda) \sim K\lambda^{-2d_i}$, as $\lambda \to 0^+$, for some constant $K > 0$ and $i \in \{1, \cdots, q\}$.

Processes of the form (1.1) constitute the so-called fractionally integrated processes. As a particular case, consider the situation where the i-th coordinate process $\{Y_t^{(i)}\}_{t=0}^{\infty}$ follows an ARMA model. In this case, the associated coordinate process $\{X_t^{(i)}\}_{t=0}^{\infty}$ will be a classical ARFIMA process with the same AR and MA orders and differencing parameter d_i. If the process $\{\boldsymbol{Y}_t\}_{t=0}^{\infty}$ is a vectorial ARMA process, then the resulting multivariate process will be the so-called VARFIMA process with differencing parameter $\boldsymbol{d} = (d_1, \cdots, d_q)'$. VARFIMA and, more generally, fractionally integrated processes, are widely used to model multivariate processes with long-range dependence. See, for instance, the recent work of Chiriac and Voev (2011) on modeling and forecasting high frequency data by using VARFIMA and fractionally integrated processes.

[*]Corresponding author.

[a]Mathematics Institute - Federal University of Rio Grande do Sul - 9500, Bento Gonçalves Avenue - 91509-900, Porto Alegre - RS - Brazil.

E-mail addresses: guipumi@gmail.com (G. Pumi), silvia.lopes@ufrgs.br (S.R.C. Lopes).

This version: September 22, 2013.

The parameter $\boldsymbol{d}$ in (1.1) determines the spectral density function behavior at the zero frequency as well as the long run autocovariance/autocorrelation structure. Hence, estimation becomes an important matter whenever the long run structure of the process is of interest.

Estimation of the parameter $\boldsymbol{d}$ in the multivariate case has seen a growing interest in the last years. A maximum likelihood approach was first considered in Sowell (1989), but the computational cost of the author's method at the time was very high A few years later, Luceño (1996) presented a computationally cheaper alternative for the maximum likelihood approach based on rewriting and approximating the quadratic form of the Gaussian likelihood function. In a recent work, Tsay (2010) proposed an even faster approach to calculate the exact conditional likelihood based on the multivariate Durbin-Levinson algorithm. Although the maximum likelihood approach usually provides good results, it is still a computationally expensive method.

The works of Fox and Taqqu (1986), Giraitis and Surgailis (1990), among others, provided a rigorous asymptotic theory for (univariate) Gaussian parametric estimates which includes, for instance, $n^{1/2}$-consistency and asymptotic normality. One drawback is the crucial role played by the Gaussianity assumption in the theory, which also requires strong distributional and regularity conditions and is non-robust with respect to the parametric specification of the model, leading to inconsistent estimates under misspecification.

In the univariate case, Gaussian Semiparametric Estimation (GSE) was first introduced in Künsch (1987) and later rigorously developed by Robinson (1995b). It provides a more robust alternative compared to the parametric one, requiring less distributional assumptions and being more efficient. In the multivariate case, Robinson (1995a) was the first to study and develop a rigorous treatment of a semiparametric estimator. A two-step multivariate GSE has been studied in the work of Lobato (1999), which showed its asymptotic normality under mild conditions, but without relying on Gaussianity. A few years later, Shimotsu (2007) introduced a refinement of Lobato's two-step GSE, which is consistent and asymptotically normal under very mild conditions (Gaussianity is, again, nowhere assumed), but with smaller asymptotic variance than Lobato's estimator. The technique applied in Shimotsu (2007) was a multivariate extension of that in Robinson (1995b), powerful enough to show not only the consistency of the proposed estimator, but also the consistency of Lobato's two-step GSE. Recently, Nielsen (2011) extended the work of Shimotsu (2007) to include the non-stationary case by using the so-called extended periodogram while Pumi and Lopes (2013) extends the work of Lobato (1999) by considering general estimators of the spectral density function in Lobato's objective function.

The estimator introduced in Shimotsu (2007) is based on the specification of the spectral density function in a neighborhood of the zero frequency. Estimation of the differencing parameter $\boldsymbol{d}$ is obtained through minimization of an objective function, which is derived from the expression of the Gaussian log-likelihood function near the zero frequency. To obtain the objective function, the spectral density is estimated by the periodogram of the process. Although asymptotically unbiased, it is well known that the periodogram is not a consistent estimator of the spectral density, presents wild fluctuations near the zero frequency and, understood as a sequence of random variables, it does not converge to a random variable at all (cf. Grenander, 1951). Some authors actually consider the periodogram "*an extremely poor (if not useless) estimate of the spectral density function*" (Priestley, 1981, page 420). The estimators introduced in Lobato (1999) and Shimotsu (2007) are known to present a good finite sample performance, but given the wild behavior of the periodogram, a natural question is: can we do better with a better behaved spectral density estimator? In this work, our primary goal is to provide an answer to this question.

Our contribution to the theory of GSE is two-folded. First, being consistency a highly desirable property of an estimator, we study the consequences of substituting the periodogram in Shimotsu (2007)'s objective function by an arbitrary consistent estimator of the spectral density function. We prove the consistency of the proposed estimator under the same assumptions as in

Shimotsu (2007) and no assumption on the spectral density estimator other than consistency. Second, considering Shimotsu (2007)'s objective function with the periodogram substituted by an arbitrary spectral density estimator, we derive necessary conditions under which GSE is consistent and satisfy a multivariate CLT. Gaussianity is nowhere assumed. In order to assess the finite sample properties of the estimators studied here and its competitiveness, we present a simulation study based on simulated VARFIMA process. We apply the so-called smoothed periodogram and the tapered periodogram as estimators of the spectral density function.

The paper is organized as follows. In the next section, we present some preliminaries concepts and results necessary for this work and introduce a general class of estimators based on appropriate modifications of the Shimotsu's objective function. Section 3 is devoted to derive the consistency of the proposed estimator while in Section 4 we derive conditions for the proposed estimator to satisfy a multivariate CLT. In Section 5 we present some Monte Carlo simulation results to assess the finite sample performance of the proposed estimator. Conclusions and final remarks are reserved to Section 6. The usually long proofs of our results are postponed to the Appendix A.

2 Preliminaries

Let $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ be a q-dimensional process specified by (1.1) and assume that the spectral density matrix of $\boldsymbol{Y}_t$ satisfies $f_{\boldsymbol{Y}} \sim G$ for a real, symmetric, finite and positive definite matrix G. Let f be the spectral density matrix function of $\boldsymbol{X}_t$, so that

$$\mathbb{E}\big[\big(\boldsymbol{X}_t - \mathbb{E}(\boldsymbol{X}_t)\big)\big(\boldsymbol{X}_{t+h} - \mathbb{E}(\boldsymbol{X}_t)\big)'\big] = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}h\lambda} f(\lambda) d\lambda,$$

for $h \in \mathbb{N}^* := \mathbb{N}\backslash\{0\}$. Following the reasoning in Shimotsu (2007), the spectral density matrix of $\boldsymbol{X}_t$ at the Fourier frequencies $\lambda_j = 2\pi j/n$, with $j = 1, \cdots, m$ and $m = o(n)$, can be written as

$$f(\lambda_j) \sim \Lambda_j(\boldsymbol{d}) G \overline{\Lambda_j(\boldsymbol{d})}', \quad \text{for} \quad \Lambda_j(\boldsymbol{d}) = \underset{k\in\{1,\cdots,q\}}{\operatorname{diag}} \{\Lambda_j^{(k)}(\boldsymbol{d})\} \quad \text{and} \quad \Lambda_j^{(k)}(\boldsymbol{d}) = \lambda_j^{-d_k} \mathrm{e}^{\mathrm{i}(\pi-\lambda_j)d_k/2}, \tag{2.1}$$

where, for a complex matrix A, $\overline{A}'$ denotes the conjugate transpose of A. Let

$$I_n(\lambda) := w_n(\lambda)\overline{w_n(\lambda)}', \quad \text{where} \quad w_n(\lambda) := \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} \boldsymbol{X}_t \mathrm{e}^{\mathrm{i}t\lambda},$$

be the periodogram and the discrete Fourier transform of $\boldsymbol{X}_t$ at λ, respectively. From the local form of the spectral density at zero frequency, given in (2.1), replaced in the frequency domain Gaussian log-likelihood localized at the origin, Shimotsu (2007) proposed a semiparametric estimator for the fractional differencing parameter $\boldsymbol{d}$ based on the objective function

$$R(\boldsymbol{d}) := \log\big(\det\{\tilde{G}(\boldsymbol{d})\}\big) - 2\sum_{k=1}^{q} d_k \frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j), \tag{2.2}$$

where

$$\tilde{G}(\boldsymbol{d}) := \frac{1}{m}\sum_{j=1}^{m} \mathrm{Re}\big[\Lambda_j(\boldsymbol{d})^{-1} I_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}'\big]. \tag{2.3}$$

with $\Lambda_j(\boldsymbol{d})$ defined in (2.1). The estimator of $\boldsymbol{d}$ is then given by

$$\widetilde{\boldsymbol{d}} = \underset{\boldsymbol{d}\in\Theta}{\arg\min}\{R(\boldsymbol{d})\}, \tag{2.4}$$

where the space of admissible estimates is of the form $\Theta = [-1/2 + \epsilon_1, 1/2 - \epsilon_2]$, for arbitrarily small $\epsilon_i > 0$, $i = 1, 2$, henceforth fixed except stated otherwise. Shimotsu (2007) shows that the estimator based on the objective function (2.2) and (2.3) is consistent under mild conditions.

Given the wild behavior of the periodogram as an estimator of the spectral density function, specially near the origin, in this work we consider substituting the periodogram by some other spectral density estimator, say f_n. Our interest lies on estimators based on objective functions of the form

$$S(\boldsymbol{d}) := \log\big(\det\{\widehat{G}(\boldsymbol{d})\}\big) - 2\sum_{k=1}^{q} d_k \frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j), \tag{2.5}$$

with

$$\widehat{G}(\boldsymbol{d}) := \frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\big[\Lambda_j(\boldsymbol{d})^{-1} f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}\,'\big]. \tag{2.6}$$

Notice that (2.5) is just (2.2) with the periodogram I_n in (2.3) replaced by f_n. The estimator of $\boldsymbol{d}$ is then defined analogously as

$$\widehat{\boldsymbol{d}} = \underset{\boldsymbol{d}\in\Theta}{\arg\min}\{S(\boldsymbol{d})\}. \tag{2.7}$$

In the sections to come, we shall study the asymptotic behavior of estimator (2.7). The study is focused on two different classes of spectral density estimator. First we shall consider the class of consistent estimators of the spectral density function and we shall show that, under no further hypothesis on the f_n, estimator (2.7) is consistent. Second, we consider a class of spectral density functions satisfying a moment condition and we shall derive conditions for the consistency and asymptotic normality of the estimator (2.7).

Before proceeding with the results, we shall establish some notation. Let $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ be a q-dimensional process specified by (1.1) and let $\{\boldsymbol{\varepsilon}_t\}_{t\in\mathbb{Z}}$ such that $\boldsymbol{X}_t - \mathbb{E}(\boldsymbol{X}_t) = \sum_{k=0}^{\infty} A_k\boldsymbol{\varepsilon}_{t-k}$. We define a function A by setting

$$A(\lambda) := \sum_{k=0}^{\infty} A_k \mathrm{e}^{\mathrm{i}k\lambda}. \tag{2.8}$$

The periodogram function associated to $\{\boldsymbol{\varepsilon}_t\}_{t\in\mathbb{Z}}$ is denoted by $I_{\boldsymbol{\varepsilon}}$, that is,

$$I_{\boldsymbol{\varepsilon}}(\lambda) := w_{\boldsymbol{\varepsilon}}(\lambda)\overline{w_{\boldsymbol{\varepsilon}}(\lambda)}\,', \qquad \text{where} \qquad w_{\boldsymbol{\varepsilon}}(\lambda) := \frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n}\boldsymbol{\varepsilon}_t \mathrm{e}^{\mathrm{i}t\lambda}. \tag{2.9}$$

For a matrix M, we shall denote the r-th row and the s-th column of M by $(M)_{r\cdot}$ and $(M)_{\cdot s}$, respectively.

3 Consistency of the estimator

Let $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ be a q-dimensional process specified by (1.1) and f be its spectral density matrix. Suppose that the spectral density matrix of the weakly stationary process $\{\boldsymbol{Y}_t\}_{t=0}^{\infty}$ in (1.1) satisfies $f_{\boldsymbol{Y}}(\lambda) \sim G_0$ for a real, symmetric and positive definite matrix $G_0 = (G_0^{rs})_{r,s=1}^{q}$. Let $\boldsymbol{d}_0 = (d_1^0, \cdots, d_q^0)'$ be the true fractional differencing vector parameter and assume that the following assumptions are satisfied:

A1. As $\lambda \to 0^+$,

$$f_{rs}(\lambda) = \mathrm{e}^{\mathrm{i}\pi(d_r^0 - d_s^0)/2} G_0^{rs}\lambda^{-d_r^0 - d_s^0} + o(\lambda^{-d_r^0 - d_s^0}), \qquad \text{for all } r,s = 1,\cdots,q.$$

A2. Denoting the sup-norm by $\|\cdot\|_\infty$, assume that

$$\boldsymbol{X}_t - \mathbb{E}(\boldsymbol{X}_t) = \sum_{k=0}^{\infty} A_k\boldsymbol{\varepsilon}_{t-k}, \quad \sum_{k=0}^{\infty}\big\|A_k\big\|_\infty^2 < \infty, \tag{3.1}$$

where $\{\boldsymbol{\varepsilon}_t\}_{t\in\mathbb{Z}}$ is a process such that

$$\mathbb{E}(\boldsymbol{\varepsilon}_t|\mathscr{F}_{t-1}) = 0 \quad \text{and} \quad \mathbb{E}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'|\mathscr{F}_{t-1}) = \mathrm{I}_q, \quad \text{a.s.}$$

for all $t \in \mathbb{Z}$, where I_q is the $q \times q$ identity matrix and $\mathscr{F}_t$ denotes the σ-field generated by $\{\boldsymbol{\varepsilon}_s, s \leq t\}$. Also assume that there exist a scalar random variable ξ and a constant $K > 0$ such that $\mathbb{E}(\xi^2) < \infty$ and $\mathbb{P}\big(\|\boldsymbol{\varepsilon}_t\|_\infty^2 > \eta\big) \leq K\mathbb{P}(\xi^2 > \eta)$, for all $\eta > 0$.

A3. In a neighborhood $(0, \delta)$ of the origin, A given by (2.8) is differentiable and, as $\lambda \to 0^+$,

$$\frac{\partial}{\partial \lambda}\big(\overline{A(\lambda)}\,'\big)_{r\cdot} = O\big(\lambda^{-1}\|\big(\overline{A(\lambda)}\,'\big)_{r\cdot}\|_\infty\big).$$

A4. As $n \to \infty$,

$$\frac{1}{m} + \frac{m}{n} \longrightarrow 0.$$

Remark 3.1. Assumptions **A**1-**A**4 are the same as in Shimotsu (2007) and are multivariate extensions of the assumptions made in Robinson (1995b) and analogous to the ones used in Robinson (1995a) and Lobato (1999). Assumption **A**1 describes the true spectral density matrix behavior at the origin. Notice that, since $\lim_{\lambda\to 0^+} \mathrm{e}^{\mathrm{i}\lambda} - 1 = 0$, replacing $\mathrm{e}^{\mathrm{i}\pi(d_r^0 - d_s^0)/2}$ by $\mathrm{e}^{\mathrm{i}(\pi-\lambda)(d_r^0 - d_s^0)/2}$ makes no difference. Assumption **A**2 regards the causal representation of $\boldsymbol{X}_t$, and more specifically, the behavior of the innovation process which is assumed to be a not necessarily uncorrelated square integrable martingale difference uniformly dominated (in probability) by a scalar random variable with finite second moment. Assumption **A**3 is a regularity condition (also imposed in Fox and Taqqu, 1986 and Giraitis and Surgailis, 1990, among others, in the parametric case) and will be useful in proving Lemmas 3.1 and 3.2 below. Assumption **A**4 is minimal but necessary since m must go to infinity for consistency, but slower than n in view of Assumption **A**1.

Observe that assumptions **A**1-**A**4 are only concerned to the behavior of the spectral density matrix on a neighborhood of the origin and, apart from integrability (implied by the process's weakly stationarity property), no assumption whatsoever is made on the spectral density matrix behavior outside this neighborhood. For $\beta \in (0,1)$, let f_n be an n^β-consistent (that is, $n^\beta(f_n - f) \overset{\mathbb{P}}{\longrightarrow} 0$, as n goes to infinity) estimator of the spectral density for all $\boldsymbol{d}_0 \in B$, where $B \subset \mathbb{R}^q$ is a closed set. If $d_k^0 \in (0, 0.5)$, the respective component of the spectral density matrix of $\{\boldsymbol{X}_t\}_{t=0}^\infty$ is unbounded at the origin. Hence, there is no hope in obtaining a consistent estimator of the spectral density function when $\boldsymbol{d}_0 \in (0, 0.5)^q$. For $q \in \mathbb{N}^*$, let

$$\Omega_\beta := \Big[-\frac{\beta}{2}, 0\Big)^q \bigcap \Big(-\frac{1}{2}, 0\Big)^q \bigcap B \subseteq \Big(-\frac{1}{2}, 0\Big)^q. \tag{3.2}$$

Lemma 3.1 establishes the consistency of $\widehat{G}(\boldsymbol{d}_0)$ given in (2.6) under the assumption of n^β-consistency of f_n in B. Due to their lengths, the proofs of all results in the paper are postponed to the Appendix A.

Lemma 3.1. *Let $\{\boldsymbol{X}_t\}_{t=0}^\infty$ be a q-dimensional process specified by* (1.1) *and f be its spectral density matrix. Let f_n be a n^β-consistent estimator (under the sup-norm) for f, for all $\boldsymbol{d}_0 \in B$. If $\boldsymbol{d}_0 \in \Omega_\beta$, then*

$$\widehat{G}(\boldsymbol{d}_0) = G_0 + o_{\mathbb{P}}(1).$$

Theorem 3.1 establishes the consistency of $\widehat{\boldsymbol{d}}$, given in (2.7) with Θ substituted by Ω_β, under assumptions **A**1-**A**4 and assuming n^β-consistency of the spectral density function estimator.

Theorem 3.1. *Let $\{\boldsymbol{X}_t\}_{t=0}^\infty$ be a q-dimensional process specified by* (1.1) *and f be its spectral density matrix. Let f_n be a n^β-consistent estimator (under the sup-norm) of f, for all $\boldsymbol{d}_0 \in B$ and $\beta \in (0,1)$, and let $\widehat{\boldsymbol{d}}$ be as in* (2.7) *with Θ substituted by Ω_β. Assume that assumptions **A**1-**A**4 hold and let $\boldsymbol{d}_0 \in \Omega_\beta$. Then, $\widehat{\boldsymbol{d}} \overset{\mathbb{P}}{\longrightarrow} \boldsymbol{d}_0$, as $n \to \infty$.*

Assuming the consistency of the spectral density estimator f_n in Theorem 3.1 excluded the case $\boldsymbol{d}_0 \in (0, 0.5)^q$, so that, under this assumption, the process $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ can have no long-range dependent component. To overcome this limitation, we now consider the class $\mathscr{D}$ of estimators $f_n := (f_n^{rs})_{r,s=1}^{q}$ satisfying, for all $r, s \in \{1, \cdots, q\}$,

$$\mathbb{E}\Big(\lambda_j^{d_r^0+d_s^0}\Big|f_n^{rs}(\lambda_j) - \big(A(\lambda_j)\big)_{r\cdot}I_{\boldsymbol{\varepsilon}}(\lambda_j)\big(\overline{A(\lambda_j)}\,'\big)_{\cdot s}\Big|\Big) = o(1), \quad \text{as } n \to \infty, \tag{3.3}$$

where A and $I_{\boldsymbol{\varepsilon}}$ are given by (2.8) and (2.9), respectively, and $\boldsymbol{d}_0 \in \Theta \subset [-0.5, 0.5]^q$. Compare condition (3.3) with expressions on the proof of Lemma 1 in Shimotsu (2007), with the ones in lemma 1 in Lobato (1999) and also in Robinson (1995a). Condition (3.3) is satisfied by the ordinary periodogram and the tapered periodogram and, thus, $\mathscr{D}$ is non-empty. The next lemma will be useful in proving Theorem 3.2.

Lemma 3.2. *Let $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ be a q-dimensional process specified by* (1.1) *and f be its spectral density matrix. Let $f_n \in \mathscr{D}$ and assume that assumptions **A**1-**A**4 hold. Then, for $1 \leq u < v \leq m$,*

$$\max_{r,s\in\{1,\cdots,q\}}\left\{\sum_{j=u}^{v} \mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(d_r^0-d_s^0)/2}\lambda_j^{d_r^0+d_s^0} f_n^{rs}(\lambda_j) - G_0^{rs}\right\} = \mathscr{A}_{uv} + \mathscr{B}_{uv},$$

where $\mathscr{A}_{uv}$ and $\mathscr{B}_{uv}$ satisfy

$$\mathbb{E}\big(|\mathscr{A}_{uv}|\big) = o(v-u+1) \qquad \textit{and} \qquad \max_{1\leq u<v\leq m}\big\{|v^{-1}\mathscr{B}_{uv}|\big\} = o_{\mathbb{P}}(1).$$

In Theorem 3.2 we derive a necessary condition for the consistency of $\widehat{\boldsymbol{d}}$ given in (2.7), when the consistency condition on f_n is relaxed and we assume $f_n \in \mathscr{D}$ instead.

Theorem 3.2. *Let $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ be a q-dimensional process specified by* (1.1) *and f be its spectral density matrix. Let $f_n \in \mathscr{D}$ be an estimator of f, and consider the estimator $\widehat{\boldsymbol{d}}$, based on f_n, given in* (2.7)*. Assume that assumptions **A**1-**A**4 hold. Then, $\widehat{\boldsymbol{d}} \xrightarrow{\mathbb{P}} \boldsymbol{d}_0$, as $n \to \infty$.*

4 Asymptotic normality of the estimator

In this section we present a sufficient condition for the asymptotic normality of the GSE given by (2.7), under similar assumptions as in Shimotsu (2007), with f_n an estimator of the spectral density function satisfying a single regularity condition. The asymptotic distribution of the estimator (2.7) will be the same as (2.4), established by Shimotsu (2007).

Again, let $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ be a q-dimensional process specified by (1.1) and f be its spectral density matrix. Suppose that the spectral density matrix of the weakly stationary process $\{\boldsymbol{Y}_t\}_{t=0}^{\infty}$ in (1.1) satisfies $f_{\boldsymbol{Y}}(\lambda) \sim G_0$ for a real, symmetric and positive definite matrix $G_0 = (G_0^{rs})_{r,s=1}^{q}$. Let $\boldsymbol{d}_0 = (d_1^0, \cdots, d_q^0)'$ be the true fractional differencing vector parameter. Assume that the following assumptions are satisfied

B1. For $\alpha \in (0, 2\,]$ and $r, s \in \{1, \cdots, q\}$,

$$f_{rs}(\lambda) = \mathrm{e}^{\mathrm{i}(\pi-\lambda)(d_r^0-d_s^0)/2}\lambda^{-d_r^0-d_s^0}G_0^{rs} + O(\lambda^{-d_r^0-d_s^0+\alpha}), \quad \text{as } \lambda \to 0^+.$$

B2. Assumption **A**2 holds and the process $\{\boldsymbol{\varepsilon}_t\}_{t\in\mathbb{Z}}$ has finite fourth moment.

B3. Assumption **A**3 holds.

B4. For any $\delta > 0$,

$$\frac{1}{m} + \frac{m^{1+2\alpha}\log(m)^2}{n^{2\alpha}} + \frac{\log(n)}{m^{\delta}} \longrightarrow 0, \quad \text{as } n \to \infty.$$

B5. There exists a finite real matrix M such that

$$\Lambda_j(\boldsymbol{d}_0)^{-1}A(\lambda_j) = M + o(1), \quad \text{as } \lambda_j \to 0.$$

Remark 4.1. Assumption **B**1 is a smoothness condition often imposed in spectral analysis. Compared to assumption 1 in Robinson (1995a), assumption **B**1 is slightly more restrictive. It is satisfied by certain VARFIMA processes. Assumption **B**2 imposes that the process $\{\boldsymbol{X}_t\}_{t\in\mathbb{N}^*}$ is linear with finite fourth moment. This restriction in the innovation process is necessary since at a certain point of the proof of the asymptotic normality, we shall need a CLT result regarding a certain martingale difference derived from a quadratic form involving $\{\boldsymbol{\varepsilon}_t\}_{t\in\mathbb{Z}}$, which must have finite variance. Assumption **B**4 is the same as assumption 4′ in Shimotsu (2007) and is slightly stronger than the ones imposed in Robinson (1995b) and Lobato (1999) (see Shimotsu, 2007 p.283 for a discussion). It implies that $(m/n)^b = o\big(m^{-\frac{b}{2\alpha}}\log(m)^{-\frac{b}{\alpha}}\big)$, for $b \neq 0$. Assumption **B**5 is the same as assumption 5' in Shimotsu (2007) and is a mild regularity condition in the degree of approximation of $A(\lambda_j)$ by $\Lambda_j(\boldsymbol{d}_0)$. It is satisfied by general VARFIMA processes.

The next lemma will be useful in proving Theorem 4.1. The proofs of the results in this section are presented in Appendix A.

Lemma 4.1. *Let $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ be a q-dimensional process specified by* (1.1) *and f be its spectral density matrix. Let f_n be an estimator of f, and consider the estimator $\widehat{\boldsymbol{d}}$, based on f_n, given in* (2.7)*. Assume that assumptions **B1**-**B**5 hold and that f_n satisfies*

$$\max_{1\leq v\leq m}\bigg\{\sum_{j=1}^{v}\big[f_n^{rs}(\lambda_j) - \big(A(\lambda_j)\big)_{r\cdot}I_{\boldsymbol{\varepsilon}}(\lambda_j)\big(\overline{A(\lambda_j)}\,'\big)_{\cdot s}\big]\bigg\} = o_{\mathbb{P}}\bigg(\frac{m}{n^{1+|d_r^0+d_s^0|}}\bigg), \tag{4.1}$$

for all $r, s \in \{1,\cdots,q\}$ and $\boldsymbol{d}_0 \in \Theta$, where A and $I_{\boldsymbol{\varepsilon}}$ are defined in (2.8) *and* (2.9)*, respectively. Then,*

(a) *uniformly in $1 \leq v \leq m$,*

$$\max_{r,s\in\{1,\cdots,q\}}\bigg\{\sum_{j=1}^{v}\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(d_r^0-d_s^0)/2}\lambda_j^{d_r^0+d_s^0}\Big[f_n^{rs}(\lambda_j) - \big(A(\lambda_j)\big)_{r\cdot}I_{\boldsymbol{\varepsilon}}(\lambda_j)\big(\overline{A(\lambda_j)}\,'\big)_{\cdot s}\Big]\bigg\} = o_{\mathbb{P}}\bigg(\frac{m^{1/2}}{\log(m)}\bigg); \tag{4.2}$$

(b) *uniformly in $1 \leq v \leq m$,*

$$\max_{r,s\in\{1,\cdots,q\}}\bigg\{\sum_{j=1}^{v}\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(d_r^0-d_s^0)/2}\lambda_j^{d_r^0+d_s^0}f_n^{rs}(\lambda_j) - G_0^{rs}\bigg\} = O_{\mathbb{P}}\bigg(\frac{m^{\alpha+1}}{n^{\alpha}} + m^{1/2}\log(m)\bigg). \tag{4.3}$$

The next theorem presents a necessary condition for the asymptotic normality of the GSE given in (2.7). We notice that the variance-covariance matrix of the limiting distribution is the same as the estimator in (2.4), as derived in Shimotsu (2007).

Theorem 4.1. *Let $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ be a q-dimensional process specified by* (1.1) *and f be its spectral density matrix. Let f_n be an estimator of f, and consider the estimator $\widehat{\boldsymbol{d}}$, based on f_n, given in* (2.7)*. Assume that assumptions **B1**-**B**5 hold. Suppose that f_n satisfies* (4.1)*, for all $r, s \in \{1,\cdots,q\}$ and $\boldsymbol{d}_0 \in \Theta$. Then*

$$m^{1/2}(\widehat{\boldsymbol{d}} - \boldsymbol{d}_0) \xrightarrow{d} N(\boldsymbol{0}, \Sigma^{-1}),$$

as n tends to infinity, where

$$\Sigma := 2\bigg[G_0 \odot G_0^{-1} + \mathrm{I}_q + \frac{\pi^2}{4}\big(G_0 \odot G_0^{-1} - \mathrm{I}_q\big)\bigg],$$

with I_q the $q\times q$ identity matrix and $\odot$ denotes the Hadamard product.

Observe that condition (4.1) is a refinement of (3.3) and implies the former.

5 Monte Carlo Simulation Study

In this section we perform a Monte Carlo simulation study to assess the finite sample performance of the estimator proposed in (2.7). We apply, as spectral density estimators, the so-called smoothed and tapered periodogram and compare them to the estimator (2.4). We start recalling some facts about the smoothed and tapered periodogram.

5.1 The Smoothed Periodogram

Let $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ be a q-dimensional process specified by (1.1). Under some mild conditions, a class of consistent estimators of the spectral density of $\boldsymbol{X}_t$ is the so-called class of smoothed periodogram. For an array of functions $W_n(k) := \big(W_n^{ij}(k)\big)_{i,j=1}^{q}$ (called weights) and $\{\ell(k)\}_{k=0}^{\infty}$ an increasing sequence of positive integers, the smoothed periodogram of $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ at the Fourier frequency λ_j is defined as

$$\hat{f}_n(\lambda_j) := \sum_{|k|\leq \ell(n)} W_n(k) \odot w_n(\lambda_{j+k})\overline{w_n(\lambda_{j+k})}\,', \tag{5.1}$$

where $\odot$ denotes the Hadamard product.

The smoothed periodogram (5.1) is a multivariate extension of the univariate smoothed periodogram. Notice that the use of the Hadamard product in (5.1) allows the use of different weight functions for different components of the spectral density matrix. This flexibility accommodates the necessity, often observed in practice, of modeling different characteristics of the spectral density matrix components (including the cross spectrum ones) with different weight functions. Types and properties of the different weight functions are subject of most textbooks in spectral analysis and will not be discussed here. See, for instance, Priestley (1981) and references therein.

In the presence of long-range dependence, the spectral density function has a pole at the zero frequency, so that some authors restrict the summation on (5.1) to $k \neq -j$. In practice, however, since the sample paths of $\boldsymbol{X}_t$ are finite with probability one, there is no problem in applying (5.1) as defined.

Assume that the weight functions $\big(W_n^{ij}(k)\big)_{i,j=1}^{q}$ and the sequence $\{\ell(k)\}_{k=0}^{\infty}$ satisfy the following conditions:

C1. $1/\ell(n) + \ell(n)/n \longrightarrow 0$, as n tends to infinity;

C2. $W_n^{ij}(k) = W_n^{ij}(-k)$ and $W_n^{ij}(k) \geq 0$, for all k;

C3. $\sum_{|k|\leq \ell(n)} W_n^{ij}(k) = 1$;

C4. $\sum_{|k|\leq \ell(n)} W_n^{ij}(k)^2 \longrightarrow 0$, as n tends to infinity.

It can be shown that under assumptions **C**1 - **C**4 and if $\boldsymbol{d} \in (-0.5, 0)^q$, then the smoothed periodogram is a $n^{1/2}$-consistent estimator of the spectral density. Theorem 3.1 thus applies and we conclude that the estimator (2.7) based on the smoothed periodogram is consistent for all $\boldsymbol{d}_0 \in \Omega_{1/2} = (-1/4, 0)^q$. At this moment, we have not been able to show the conditions of Theorems 3.2 and 4.1 hold for the smoothed periodogram, but we have empirical evidence that the estimator is indeed consistent and asymptotically normally distributed for $\boldsymbol{d} \in (-0.5, 0.5)^q$. See Section 5.3.

5.2 The Tapered Periodogram

In case of long-range dependent components in the process $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$, the ordinary periodogram is not only non-consistent, but it is also asymptotically biased (cf. Hurvich and Beltrão, 1993).

A simple way to reduce this asymptotic bias is by tapering the data prior calculating the periodogram of the series. Let $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ be a q-dimensional process specified by (1.1). Let $\{h_i\}_{i=1}^{q}$ be a collection of real functions defined on $[0,1]$. Consider the function $L_n : \mathbb{R} \to \mathbb{R}^q$ given by $L_n(\lambda) = \big(L_n^1(\lambda), \cdots, L_n^q(\lambda)\big)$ where $L_n^i(\lambda) := h_i\big(\lambda/n\big)$ and let

$$S_n(\lambda) := \left(\frac{L_n^i(\lambda)}{\sqrt{\sum_{t=1}^{n} L_n^i(t)^2}}\right)_{i=1}^{q}.$$

The tapered periodogram $I_T(\lambda; n)$ of the time series $\{\boldsymbol{X}_t\}_{t=1}^{n}$ is defined by setting

$$I_T(\lambda;n) := w_T(\lambda;n)\overline{w_T(\lambda;n)}\,', \quad \text{where} \quad w_T(\lambda;n) := \frac{1}{\sqrt{2\pi}}\sum_{t=1}^{n} S_n(t) \odot \boldsymbol{X}_t \mathrm{e}^{-it\lambda}. \tag{5.2}$$

We shall assume the following:

- Assumption **D**. The tapering functions h_i are of bounded variation and $H_i := \int_0^1 h_i^2(x)dx > 0$, for all $i \in \{1, \cdots, q\}$.

The tapered periodogram is not a consistent estimator of the spectral density function, since the reduction on the bias induces, in this case, an augmentation of the variance. Just as the ordinary periodogram, the increase in the variance can be dealt by smoothing the tapered periodogram in order to obtain a consistent estimator of the spectral density function in the case $\boldsymbol{d} \in (-0.5, 0)$ (see, for instance, the recent work of Fryzlewicz et al., 2008). More details can be found in Priestley (1981), Dahlhaus (1983), Hurvich and Beltrão (1993), Fryzlewicz et al. (2008) and references therein.

Under Assumption **D**, $\sum_{t=1}^{n} L_n^i(t)^2 \sim nH_i$ (cf. Fryzlewicz et al., 2008) so that $I_T(\lambda;n) = O\big(I_n(\lambda)\big)$. This allows to show that the estimator (2.7) based on the tapered periodogram is also consistent and asymptotically normally distributed. These are the contents of the next Corollaries.

Corollary 5.1. *Let $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ be a weakly stationary q-dimensional process specified by* (1.1) *and with spectral density function f satisfying Assumptions **A1**-**A4**. Let f_n be the tapered periodogram defined in* (5.2) *satisfying Assumption **D**. For $\boldsymbol{d}_0 \in \Theta$, consider the estimator $\widehat{\boldsymbol{d}}$ based on f_n, as given in* (2.7)*. Then, $\widehat{\boldsymbol{d}} \xrightarrow{\mathbb{P}} \boldsymbol{d}_0$, as n tends to infinity.*

Corollary 5.2. *Let $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ be a weakly stationary q-dimensional process specified by* (1.1) *and with spectral density function f satisfying Assumptions **B1**-**B5**, with **B4** holding for $\alpha = 1$. Let f_n be the tapered periodogram given in* (5.2) *satisfying Assumption **D**. For $\boldsymbol{d}_0 \in \Theta$, consider the estimator $\widehat{\boldsymbol{d}}$ based on f_n, as given in* (2.7)*. Then, for $\boldsymbol{d}_0 \in \Theta$,*

$$m^{1/2}(\widehat{\boldsymbol{d}} - \boldsymbol{d}_0) \xrightarrow{d} N(\boldsymbol{0}, \Omega),$$

as n tends to infinity, with Ω as given in Theorem 4.1.

5.3 Simulation Results

In this section we present a Monte Carlo simulation study to assess the finite sample performance of the estimator (2.7). Recall that a q-dimensional stationary process $\{\boldsymbol{X}_t\}_{t\in\mathbb{Z}}$ is called a VARFIMA$(p, \boldsymbol{d}, q)$ if it is a stationary solution of the difference equations

$$\boldsymbol{\Phi}(\mathcal{B})\,\mathrm{diag}\big\{(1-\mathcal{B})^{\boldsymbol{d}}\big\}\big(\boldsymbol{X}_t - \mathbb{E}(\boldsymbol{X}_t)\big) = \boldsymbol{\Theta}(\mathcal{B})\boldsymbol{\varepsilon}_t,$$

where $\mathcal{B}$ is the backward shift operator, $\{\boldsymbol{\varepsilon}_t\}_{t\in\mathbb{Z}}$ is a q-dimensional stationary process (the innovation process), $\boldsymbol{\Phi}(\mathcal{B})$ and $\boldsymbol{\Theta}(\mathcal{B})$ are $q \times q$ matrices in $\mathcal{B}$, given by the equations

$$\boldsymbol{\Phi}(\mathcal{B}) = \sum_{\ell=0}^{p} \boldsymbol{\phi}_\ell \mathcal{B}^\ell \quad \text{and} \quad \boldsymbol{\Theta}(\mathcal{B}) = \sum_{r=0}^{q} \boldsymbol{\theta}_r \mathcal{B}^r,$$

assumed to have no common roots, where $\boldsymbol{\phi}_1, \cdots, \boldsymbol{\phi}_p, \boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_q$ are real $q \times q$ matrices and $\boldsymbol{\phi}_0 = \boldsymbol{\theta}_0 = \mathrm{I}_q$.

All Monte Carlo simulations are based on time series of fixed sample size $n = 1,000$ obtained from bidimensional Gaussian VARFIMA$(0, \boldsymbol{d}, 0)$ processes for several different parameters $\boldsymbol{d}$ and correlation $\rho \in \{0, 0.3, 0.6, 0.8\}$. We perform 1,000 replications of each experiment. To generate the time series, we apply the traditional method of truncating the multidimensional infinite moving average representation of the process. The truncation point is fixed in 50,000 for all cases. For comparison purposes, we calculate the estimator (2.4) (denoted by Sh) and the estimator (2.7) with the smoothed periodogram with and without the restriction $k \neq -j$ (denoted by SSh and SSh*, respectively) and with the tapered periodogram (denoted by TSh).

For the smoothed periodogram, we apply the same weights for all spectral density components, given by the so-called Bartlett's window, that is,

$$W_n^{ij}(k) := \frac{\sin^2(\ell(n)k/2)}{n\ell(n)\sin^2(k/2)}, \quad \text{for all } i, j = 1, 2.$$

For the tapered periodogram, we apply the cosine-bell tapering function, namely,

$$h_i(u) = \begin{cases} \frac{1}{2}\big[1 - \cos(2\pi u)\big], & \text{if } 0 \leq u \leq 1/2, \\ h_i(1-u), & \text{if } 1/2 < u \leq 1, \end{cases} \quad \text{for all } i = 1, 2.$$

The cosine-bell taper is widely used in applications as, for instance, in Hurvich and Ray (1995), Velasco (1999) and Olbermann et al. (2006).

The truncation point of the smoothed periodogram function is of the form $\ell(n, \beta) := \lfloor n^{\beta} \rfloor$, for $\beta \in \{0.7, 0.9\}$, while the truncation point of the estimator (2.7) is of the form $m := m(n, \alpha) = \lfloor n^{\alpha} \rfloor$, for $\alpha \in \{0.65, 0.85\}$ for all estimators. The routines are implemented in FORTRAN 95 language optimized by using OpenMP directives for parallel computing. All simulations were performed by using the computational resources from the (Brazilian) National Center of Super Computing (CESUP-UFRGS).

Tables 5.1 and 5.2 report the simulation results. Presented are the estimated values (mean), their standard deviations (st.d.) and the mean square error of the estimates (mse). Overall all estimators present a good performance with small mse and standard deviation. The bias is generally small, except when the correlation in the noise is very high ($\rho = 0.8$) and the respective component of $\boldsymbol{d}$ is small (specially 0.1), in which case the bias is high. The SSh and SSh* estimators generally perform better than Sh and TSh in terms of both, mse and bias. The same can be said about the standard deviations of the estimators. As the correlation in the innovation process increases, the estimated values degrade in some degree according to the magnitude of the respective parameter component.

The best performance in terms of bias is obtained for $\alpha = 0.85$ for most cases (83 out of 128 cases). The value $\alpha = 0.85$ also gives uniformly smaller mean square errors for all estimators. For the SSh estimator, there is an overall equilibrium over the values of β presenting the smallest bias and usually the combination $\alpha = 0.85$ and $\beta = 0.9$ gives the best results in terms of mean square error. For the SSh*, $\alpha = 0.85$ gives the best results in most cases (21 out of 32) while for β there is an equilibrium between the values. The Sh and the TSh estimators present similar behavior and in most cases they agree in the value of α which yields smallest bias. There is a small advantage for $\alpha = 0.85$ in terms of bias in both cases (19 and 18 out of 32 for the Sh and TSh, respectively). In terms of mean square error and standard deviation, $\alpha = 0.85$ generally presents best results and overall the Sh has advantage over the TSh estimator. The variance of the estimators generally responds strongly to changes in α than in β in the opposite direction, that is, the higher the α, the smaller the variance.

Table 5.1: Simulation results of the estimator (2.7) based on the smoothed periodogram (SSh and SSh*), the ordinary periodogram (Sh) and the tapered periodogram (TSh) in VARFIMA$(0, \boldsymbol{d}, 0)$ processes. Presented are the estimated values (mean), its standard deviation (st.d) and the mean square error of the estimates (mse).

ρ	Method	β	$\hat{d}_i$	$\boldsymbol{d} = (0.1, 0.4)$						$\boldsymbol{d} = (0.2, 0.3)$					
				$\alpha = 0.65$			$\alpha = 0.85$			$\alpha = 0.65$			$\alpha = 0.85$		
				mean	st.d.	mse	mean	st.d.	mse	mean	st.d.	mse	mean	st.d.	mse
0	SSh	0.7	$\hat{d}_1$	0.1043	0.0542	0.0029	0.0957	0.0268	0.0007	0.2089	0.0566	0.0033	0.1945	0.0277	0.0008
			$\hat{d}_2$	0.4453	0.0787	0.0082	0.4136	0.0401	0.0018	0.3160	0.0640	0.0043	0.2998	0.0320	0.0010
		0.9	$\hat{d}_1$	0.1062	0.0558	0.0032	0.0956	0.0267	0.0007	0.2091	0.0561	0.0032	0.1924	0.0268	0.0008
			$\hat{d}_2$	0.4216	0.0622	0.0043	0.3948	0.0306	0.0010	0.3074	0.0577	0.0034	0.2918	0.0288	0.0009
	SSh*	0.7	$\hat{d}_1$	0.0940	0.0547	0.0030	0.0920	0.0270	0.0008	0.1929	0.0564	0.0032	0.1885	0.0275	0.0009
			$\hat{d}_2$	0.3823	0.0616	0.0041	0.3853	0.0308	0.0012	0.2783	0.0590	0.0040	0.2849	0.0298	0.0011
		0.9	$\hat{d}_1$	0.0995	0.0571	0.0033	0.0934	0.0270	0.0008	0.1998	0.0574	0.0033	0.1892	0.0271	0.0009
			$\hat{d}_2$	0.3882	0.0605	0.0038	0.3813	0.0295	0.0012	0.2866	0.0595	0.0037	0.2843	0.0292	0.0011
	Sh	-	$\hat{d}_1$	0.1069	0.0577	0.0034	0.0958	0.0270	0.0007	0.2076	0.0576	0.0034	0.1916	0.0271	0.0008
			$\hat{d}_2$	0.3940	0.0603	0.0037	0.3820	0.0292	0.0012	0.2933	0.0598	0.0036	0.2859	0.0291	0.0010
	TSh	-	$\hat{d}_1$	0.1090	0.0777	0.0061	0.0962	0.0376	0.0014	0.2087	0.0773	0.0060	0.1917	0.0376	0.0015
			$\hat{d}_2$	0.4051	0.0767	0.0059	0.3863	0.0384	0.0017	0.2987	0.0764	0.0058	0.2880	0.0385	0.0016
0.3	SSh	0.7	$\hat{d}_1$	0.1272	0.0526	0.0035	0.1114	0.0255	0.0008	0.2175	0.0517	0.0030	0.1998	0.0257	0.0007
			$\hat{d}_2$	0.4288	0.0764	0.0067	0.4054	0.0394	0.0016	0.3088	0.0587	0.0035	0.2957	0.0302	0.0009
		0.9	$\hat{d}_1$	0.1281	0.0537	0.0037	0.1102	0.0253	0.0007	0.2167	0.0513	0.0029	0.1972	0.0250	0.0006
			$\hat{d}_2$	0.4057	0.0600	0.0036	0.3864	0.0296	0.0011	0.3009	0.0530	0.0028	0.2879	0.0270	0.0009
	SSh*	0.7	$\hat{d}_1$	0.1141	0.0523	0.0029	0.1064	0.0255	0.0007	0.1991	0.0513	0.0026	0.1931	0.0255	0.0007
			$\hat{d}_2$	0.3660	0.0598	0.0047	0.3769	0.0298	0.0014	0.2727	0.0544	0.0037	0.2811	0.0279	0.0011
		0.9	$\hat{d}_1$	0.1196	0.0544	0.0033	0.1073	0.0255	0.0007	0.2061	0.0524	0.0028	0.1937	0.0252	0.0007
			$\hat{d}_2$	0.3719	0.0585	0.0042	0.3727	0.0282	0.0015	0.2808	0.0550	0.0034	0.2805	0.0273	0.0011
	Sh	-	$\hat{d}_1$	0.1263	0.0548	0.0037	0.1094	0.0255	0.0007	0.2135	0.0526	0.0029	0.1959	0.0252	0.0007
			$\hat{d}_2$	0.3781	0.0583	0.0039	0.3735	0.0279	0.0015	0.2878	0.0552	0.0032	0.2822	0.0272	0.0011
	TSh	-	$\hat{d}_1$	0.1301	0.0747	0.0065	0.1110	0.0361	0.0014	0.2159	0.0712	0.0053	0.1967	0.0353	0.0013
			$\hat{d}_2$	0.3897	0.0735	0.0055	0.3773	0.0366	0.0019	0.2926	0.0704	0.0050	0.2838	0.0359	0.0016
0.6	SSh	0.7	$\hat{d}_1$	0.1850	0.0488	0.0096	0.1562	0.0241	0.0037	0.2353	0.0450	0.0033	0.2126	0.0226	0.0007
			$\hat{d}_2$	0.3987	0.0709	0.0050	0.3915	0.0383	0.0015	0.2968	0.0513	0.0026	0.2876	0.0269	0.0009
		0.9	$\hat{d}_1$	0.1817	0.0482	0.0090	0.1512	0.0231	0.0032	0.2323	0.0440	0.0030	0.2089	0.0219	0.0006
			$\hat{d}_2$	0.3780	0.0557	0.0036	0.3724	0.0279	0.0015	0.2900	0.0464	0.0023	0.2801	0.0238	0.0010
	SSh*	0.7	$\hat{d}_1$	0.1609	0.0474	0.0060	0.1463	0.0232	0.0027	0.2118	0.0444	0.0021	0.2042	0.0223	0.0005
			$\hat{d}_2$	0.3391	0.0556	0.0068	0.3629	0.0282	0.0022	0.2633	0.0478	0.0036	0.2736	0.0247	0.0013
		0.9	$\hat{d}_1$	0.1666	0.0487	0.0068	0.1460	0.0230	0.0026	0.2190	0.0453	0.0024	0.2045	0.0221	0.0005
			$\hat{d}_2$	0.3448	0.0539	0.0059	0.3584	0.0260	0.0024	0.2712	0.0481	0.0031	0.2730	0.0240	0.0013
	Sh	-	$\hat{d}_1$	0.1727	0.0490	0.0077	0.1475	0.0229	0.0028	0.2261	0.0454	0.0027	0.2065	0.0220	0.0005
			$\hat{d}_2$	0.3518	0.0537	0.0052	0.3593	0.0256	0.0023	0.2786	0.0483	0.0028	0.2748	0.0238	0.0012
	TSh	-	$\hat{d}_1$	0.1802	0.0662	0.0108	0.1510	0.0327	0.0037	0.2300	0.0604	0.0045	0.2081	0.0309	0.0010
			$\hat{d}_2$	0.3638	0.0677	0.0059	0.3628	0.0339	0.0025	0.2831	0.0611	0.0040	0.2760	0.0316	0.0016
0.8	SSh	0.7	$\hat{d}_1$	0.2373	0.0522	0.0216	0.2099	0.0274	0.0128	0.2491	0.0445	0.0044	0.2252	0.0220	0.0011
			$\hat{d}_2$	0.3944	0.0697	0.0049	0.3982	0.0388	0.0015	0.2945	0.0498	0.0025	0.2858	0.0258	0.0009
		0.9	$\hat{d}_1$	0.2295	0.0494	0.0192	0.1998	0.0238	0.0105	0.2443	0.0424	0.0038	0.2202	0.0207	0.0008
			$\hat{d}_2$	0.3746	0.0561	0.0038	0.3776	0.0280	0.0013	0.2875	0.0449	0.0022	0.2780	0.0225	0.0010
	SSh*	0.7	$\hat{d}_1$	0.2016	0.0489	0.0127	0.1937	0.0237	0.0093	0.2217	0.0433	0.0023	0.2150	0.0212	0.0007
			$\hat{d}_2$	0.3364	0.0562	0.0072	0.3693	0.0283	0.0017	0.2614	0.0464	0.0036	0.2718	0.0233	0.0013
		0.9	$\hat{d}_1$	0.2075	0.0495	0.0140	0.1915	0.0230	0.0089	0.2289	0.0439	0.0028	0.2149	0.0209	0.0007
			$\hat{d}_2$	0.3417	0.0541	0.0063	0.3634	0.0256	0.0020	0.2689	0.0464	0.0031	0.2709	0.0224	0.0013
	Sh	-	$\hat{d}_1$	0.2137	0.0498	0.0154	0.1925	0.0229	0.0091	0.2358	0.0440	0.0032	0.2167	0.0208	0.0007
			$\hat{d}_2$	0.3489	0.0540	0.0055	0.3642	0.0252	0.0019	0.2764	0.0465	0.0027	0.2728	0.0223	0.0012
	TSh	-	$\hat{d}_1$	0.2248	0.0659	0.0199	0.1974	0.0322	0.0105	0.2409	0.0572	0.0049	0.2187	0.0289	0.0012
			$\hat{d}_2$	0.3614	0.0681	0.0061	0.3681	0.0335	0.0021	0.2814	0.0587	0.0038	0.2741	0.0297	0.0016

Table 5.2: Simulation results of the estimator (2.7) based on the smoothed periodogram (SSh and SSh*), the ordinary periodogram (Sh) and the tapered periodogram (TSh) in VARFIMA$(0, \boldsymbol{d}, 0)$ processes. Presented are the estimated values (mean), its standard deviation (st.d) and the mean square error of the estimates (mse).

ρ	Method	β	$\hat{d}_i$	$\boldsymbol{d} = (0.1, 0.3)$						$\boldsymbol{d} = (0.3, 0.4)$					
				$\alpha = 0.65$			$\alpha = 0.85$			$\alpha = 0.65$			$\alpha = 0.85$		
				mean	st.d.	mse	mean	st.d.	mse	mean	st.d.	mse	mean	st.d.	mse
0	SSh	0.7	$\hat{d}_1$	0.1038	0.0541	0.0029	0.0956	0.0268	0.0007	0.3237	0.0610	0.0043	0.2987	0.0299	0.0009
			$\hat{d}_2$	0.3156	0.0641	0.0044	0.2998	0.0321	0.0010	0.4454	0.0782	0.0082	0.4136	0.0401	0.0018
		0.9	$\hat{d}_1$	0.1057	0.0558	0.0031	0.0955	0.0267	0.0007	0.3173	0.0570	0.0035	0.2916	0.0273	0.0008
			$\hat{d}_2$	0.3068	0.0577	0.0034	0.2918	0.0288	0.0009	0.4225	0.0620	0.0044	0.3948	0.0306	0.0010
	SSh*	0.7	$\hat{d}_1$	0.0935	0.0546	0.0030	0.0919	0.0270	0.0008	0.2975	0.0586	0.0034	0.2880	0.0284	0.0010
			$\hat{d}_2$	0.2777	0.0591	0.0040	0.2848	0.0298	0.0011	0.3831	0.0613	0.0040	0.3854	0.0308	0.0012
		0.9	$\hat{d}_1$	0.0990	0.0570	0.0032	0.0933	0.0270	0.0008	0.3032	0.0581	0.0034	0.2863	0.0274	0.0009
			$\hat{d}_2$	0.2859	0.0596	0.0037	0.2843	0.0292	0.0011	0.3893	0.0603	0.0037	0.3814	0.0294	0.0012
	Sh	-	$\hat{d}_1$	0.1063	0.0576	0.0034	0.0957	0.0270	0.0007	0.3103	0.0579	0.0035	0.2881	0.0272	0.0009
			$\hat{d}_2$	0.2926	0.0599	0.0036	0.2859	0.0291	0.0010	0.3951	0.0601	0.0036	0.3821	0.0292	0.0012
	TSh	-	$\hat{d}_1$	0.1079	0.0775	0.0061	0.0960	0.0376	0.0014	0.3124	0.0770	0.0061	0.2884	0.0375	0.0015
			$\hat{d}_2$	0.2976	0.0765	0.0059	0.2879	0.0385	0.0016	0.4071	0.0765	0.0059	0.3865	0.0384	0.0017
0.3	SSh	0.7	$\hat{d}_1$	0.1204	0.0506	0.0030	0.1061	0.0251	0.0007	0.3333	0.0557	0.0042	0.3043	0.0277	0.0008
			$\hat{d}_2$	0.3022	0.0603	0.0036	0.2927	0.0307	0.0010	0.4380	0.0719	0.0066	0.4096	0.0381	0.0015
		0.9	$\hat{d}_1$	0.1213	0.0520	0.0032	0.1054	0.0250	0.0007	0.3253	0.0521	0.0033	0.2965	0.0254	0.0007
			$\hat{d}_2$	0.2941	0.0542	0.0030	0.2848	0.0273	0.0010	0.4161	0.0569	0.0035	0.3910	0.0288	0.0009
	SSh*	0.7	$\hat{d}_1$	0.1075	0.0508	0.0026	0.1016	0.0252	0.0006	0.3036	0.0531	0.0028	0.2923	0.0263	0.0008
			$\hat{d}_2$	0.2654	0.0556	0.0043	0.2779	0.0282	0.0013	0.3778	0.0565	0.0037	0.3818	0.0290	0.0012
		0.9	$\hat{d}_1$	0.1132	0.0530	0.0030	0.1028	0.0252	0.0006	0.3093	0.0529	0.0029	0.2906	0.0255	0.0007
			$\hat{d}_2$	0.2735	0.0561	0.0038	0.2773	0.0276	0.0013	0.3838	0.0556	0.0034	0.3778	0.0276	0.0013
	Sh	-	$\hat{d}_1$	0.1201	0.0534	0.0033	0.1049	0.0252	0.0007	0.3161	0.0528	0.0030	0.2922	0.0253	0.0007
			$\hat{d}_2$	0.2806	0.0564	0.0036	0.2790	0.0274	0.0012	0.3899	0.0555	0.0032	0.3786	0.0273	0.0012
	TSh	-	$\hat{d}_1$	0.1228	0.0726	0.0058	0.1061	0.0356	0.0013	0.3200	0.0711	0.0054	0.2934	0.0352	0.0013
			$\hat{d}_2$	0.2854	0.0717	0.0053	0.2805	0.0362	0.0017	0.4011	0.0704	0.0050	0.3824	0.0359	0.0016
0.6	SSh	0.7	$\hat{d}_1$	0.1571	0.0451	0.0053	0.1334	0.0226	0.0016	0.3543	0.0493	0.0054	0.3186	0.0247	0.0010
			$\hat{d}_2$	0.2806	0.0542	0.0033	0.2803	0.0282	0.0012	0.4249	0.0627	0.0046	0.4015	0.0347	0.0012
		0.9	$\hat{d}_1$	0.1553	0.0452	0.0051	0.1310	0.0221	0.0014	0.3425	0.0449	0.0038	0.3086	0.0223	0.0006
			$\hat{d}_2$	0.2736	0.0486	0.0031	0.2725	0.0246	0.0014	0.4049	0.0500	0.0025	0.3834	0.0258	0.0009
	SSh*	0.7	$\hat{d}_1$	0.1374	0.0449	0.0034	0.1266	0.0223	0.0012	0.3163	0.0457	0.0024	0.3034	0.0229	0.0005
			$\hat{d}_2$	0.2462	0.0501	0.0054	0.2657	0.0256	0.0018	0.3688	0.0498	0.0035	0.3747	0.0259	0.0013
		0.9	$\hat{d}_1$	0.1435	0.0464	0.0040	0.1273	0.0223	0.0012	0.3220	0.0456	0.0026	0.3013	0.0222	0.0005
			$\hat{d}_2$	0.2538	0.0503	0.0047	0.2650	0.0247	0.0018	0.3745	0.0487	0.0030	0.3705	0.0243	0.0015
	Sh	-	$\hat{d}_1$	0.1500	0.0468	0.0047	0.1290	0.0223	0.0013	0.3284	0.0455	0.0029	0.3026	0.0221	0.0005
			$\hat{d}_2$	0.2614	0.0506	0.0040	0.2669	0.0246	0.0017	0.3809	0.0486	0.0027	0.3714	0.0240	0.0014
	TSh	-	$\hat{d}_1$	0.1544	0.0625	0.0069	0.1311	0.0315	0.0020	0.3353	0.0603	0.0049	0.3050	0.0309	0.0010
			$\hat{d}_2$	0.2663	0.0639	0.0052	0.2680	0.0325	0.0021	0.3914	0.0614	0.0038	0.3747	0.0317	0.0016
0.8	*SSh*	0.7	$\hat{d}_1$	0.1882	0.0468	0.0100	0.1641	0.0231	0.0046	0.3710	0.0504	0.0076	0.3332	0.0253	0.0017
			$\hat{d}_2$	0.2787	0.0544	0.0034	0.2822	0.0280	0.0011	0.4228	0.0605	0.0042	0.4003	0.0338	0.0011
		0.9	$\hat{d}_1$	0.1840	0.0452	0.0091	0.1593	0.0217	0.0040	0.3562	0.0440	0.0051	0.3208	0.0215	0.0009
			$\hat{d}_2$	0.2713	0.0486	0.0032	0.2736	0.0239	0.0013	0.4028	0.0485	0.0024	0.3817	0.0247	0.0009
	SSh*	0.7	$\hat{d}_1$	0.1625	0.0455	0.0060	0.1545	0.0220	0.0035	0.3263	0.0447	0.0027	0.3145	0.0218	0.0007
			$\hat{d}_2$	0.2443	0.0501	0.0056	0.2672	0.0249	0.0017	0.3671	0.0485	0.0034	0.3733	0.0248	0.0013
		0.9	$\hat{d}_1$	0.1687	0.0465	0.0069	0.1543	0.0217	0.0034	0.3318	0.0442	0.0030	0.3117	0.0210	0.0006
			$\hat{d}_2$	0.2514	0.0499	0.0048	0.2659	0.0237	0.0017	0.3724	0.0471	0.0030	0.3687	0.0229	0.0015
	Sh	-	$\hat{d}_1$	0.1753	0.0468	0.0079	0.1558	0.0216	0.0036	0.3381	0.0440	0.0034	0.3128	0.0208	0.0006
			$\hat{d}_2$	0.2592	0.0502	0.0042	0.2678	0.0236	0.0016	0.3790	0.0469	0.0026	0.3695	0.0226	0.0014
	TSh	-	$\hat{d}_1$	0.1811	0.0611	0.0103	0.1584	0.0303	0.0043	0.3477	0.0573	0.0055	0.3161	0.0289	0.0011
			$\hat{d}_2$	0.2647	0.0633	0.0053	0.2692	0.0315	0.0019	0.3901	0.0592	0.0036	0.3730	0.0299	0.0016

Figure 5.1 presents the scatter plot, histogram and kernel density estimator of the SSh estimated values for $\boldsymbol{d}_0 = (0.2, 0.3)$ when $\alpha = 0.85$ and $\beta = 0.9$. Figures 5.1(a)–(c) correspond to $\rho = 0$, Figures 5.1(d)–(f) to $\rho = 0.3$, Figures 5.1(g)–(i) to $\rho = 0.6$ and Figures 5.1(j)–(l) to $\rho = 0.8$. At this moment, we were not able to prove the asymptotic normality of the SSh estimator by direct verification of (4.1). However, we conjecture that this is the case and Figure 5.1 supports this opinion.

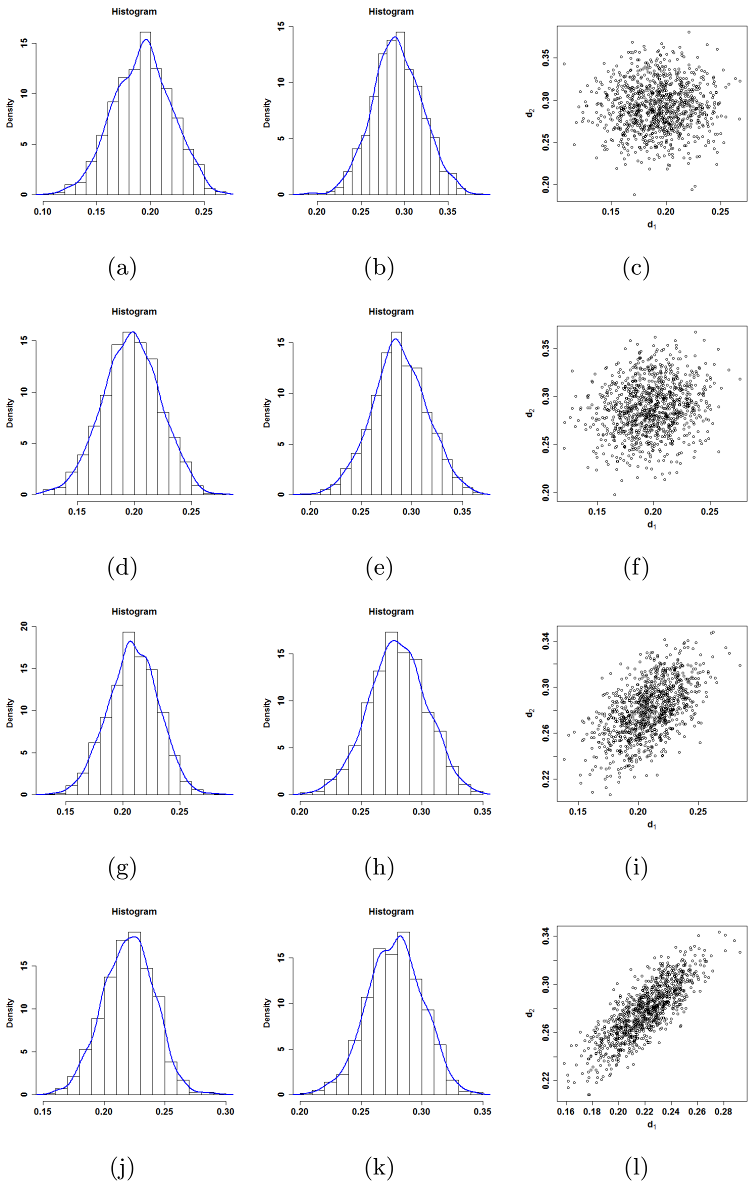

Figure 5.1: Histogram, kernel density and scatter plot of the SSh estimated values of $\boldsymbol{d}_0 = (0.2, 0.3)$ for (a)–(c) $\rho = 0$; (d)–(f) $\rho = 0.3$; (g)–(i) $\rho = 0.6$ and (j)–(l) $\rho = 0.8$.

Finally we observe that, although we have applied "better" spectral density estimators in the simulations, the results obtained from the respective estimators were not as superior as one could expect compared to the ones obtained with the ordinary periodogram. A possible explanation, brought to our attention by a referee, is that the estimators applied here are functional forms of the ordinary periodogram and may inherit too much of its local behavior not allowing the proposed estimators to present substantial improvement.

6 Conclusions

In this work we propose and analyze a class of Gaussian semiparametric estimators of multivariate long-range dependent processes. The work is motivated by the semiparametric methodology presented in Shimotsu (2007). More specifically, we propose a class of estimators based on the method studied in Shimotsu (2007) by substituting the periodogram applied there for an arbitrary spectral density estimator. We analyze two frameworks. First we assume that the spectral density estimator is consistent for the spectral density estimator and we show that the proposed semiparametric estimator is also consistent under mild conditions. Second, we relax the consistency condition and derive necessary conditions for the consistency and asymptotic normality of the proposed estimator. We show that the variance-covariance matrix of the limiting distribution is the same as the one derived in Shimotsu (2007), under the same conditions imposed in the process.

In order to assess the finite sample performance and illustrate the usefulness of the estimator, we perform a Monte Carlo simulation based on VARFIMA$(0, \boldsymbol{d}, 0)$ processes. We applied the smoothed periodogram with the Bartlett's weight function and the tapered periodogram with the cosine-bell taper as the spectral density estimators. For comparison we also compute the estimator proposed in Shimotsu (2007).

The assumptions required in the asymptotic theory are mild ones and are commonly applied in the literature. The semiparametric methodology present several advantages compared to the parametric framework such as weaker distributional assumptions, robustness with respect to misspecification of the short run dynamics of the process and efficiency. The theory includes the fractionally integrated processes as well as the class of VARFIMA processes.

Acknowledgements

G. Pumi was partially supported by CAPES/Fulbright Grant BEX 2910/06-3 and by CNPq-Brazil. S.R.C. Lopes research was partially supported by CNPq-Brazil, by CAPES-Brazil, by Pronex *Probabilidade e Processos Estocásticos* - E-26/170.008/2008 -APQ1 and also by INCT *em Matemática*. The authors are also grateful to the (Brazilian) National Center of Super Computing (CESUP-UFRGS) for the computational resources.

Appendix A: Proofs

In this section we present the proofs of the results in Sections 4 and 5. We establish lemmas and theorems in the same sequence as they appear in the text.

Proof of Lemma 3.1:

By hypothesis, $f_n(\lambda) = f(\lambda) + o_{\mathbb{P}}(n^{-\beta})$ in B. Recalling the definition of Λ_j given in (2.1), we have

$$\widehat{G}(\boldsymbol{d}_0) = \frac{1}{m}\sum_{j=1}^{m} \mathrm{Re}\big[\Lambda_j(\boldsymbol{d}_0)^{-1} f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'\big] = \frac{1}{m}\sum_{j=1}^{m} \mathrm{Re}\big[\Lambda_j(\boldsymbol{d}_0)^{-1}\big(f(\lambda_j) + o_{\mathbb{P}}(n^{-\beta})\big)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'\big]$$

$$= G_0 + \frac{1}{m}\sum_{j=1}^{m} \mathrm{Re}\big[\Lambda_j(\boldsymbol{d}_0)^{-1} o_{\mathbb{P}}(n^{-\beta})\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'\big]. \tag{A.1}$$

The (r, s)-th component of the second part on the RHS of (A.1) is given by

$$\frac{1}{m}\sum_{j=1}^{m} \mathrm{Re}\big[\mathrm{e}^{\mathrm{i}(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0}\big] o_{\mathbb{P}}(n^{-\beta}) = \bigg[\frac{1}{m}\sum_{j=1}^{m} \lambda_j^{d_r^0 + d_s^0}\bigg] o_{\mathbb{P}}(n^{-\beta})$$

$$= \frac{1}{d_r^0 + d_s^0 + 1}\left(\frac{2\pi m}{n}\right)^{d_r^0+d_s^0}\left[\frac{d_r^0 + d_s^0 + 1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{d_r^0+d_s^0}\right]o_{\mathbb{P}}(n^{-\beta})$$

$$= \frac{1}{d_r^0 + d_s^0 + 1}\left(\frac{2\pi m}{n}\right)^{d_r^0+d_s^0}\big[O(m^{\beta-1}) + 1\big]o_{\mathbb{P}}(n^{-\beta}) \;=\; o_{\mathbb{P}}(1),$$

where the penultimate equality follows from lemma 1 in Robinson (1995b), by taking $\gamma = d_r^0 + d_s^0 + 1 > 1 - \beta > 0$, while the last one follows since $\beta \in (0,1)$ and $\boldsymbol{d}_0 \in \Omega_\beta$. Hence, $\widehat{G}(\boldsymbol{d}_0)_{rs} = G_0^{rs} + o_{\mathbb{P}}(1)$, for all $r,s \in \{1,\cdots,q\}$ and the proof is complete. ■

Proof of Theorem 3.1:

Let $\boldsymbol{\theta} = (\theta_1,\cdots,\theta_q)' := \boldsymbol{d} - \boldsymbol{d}_0$ and $L(\boldsymbol{d}) := S(\boldsymbol{d}) - S(\boldsymbol{d}_0)$. Let $0 < \delta < 1/2$ be fixed and let

$$N_\delta := \big\{\boldsymbol{d} : \|\boldsymbol{d} - \boldsymbol{d}_0\|_\infty > \delta\big\}.$$

Let $0 < \epsilon < 1/4$ and define $\Theta_1 := \big\{\boldsymbol{\theta} : \boldsymbol{\theta} \in [-1/2+\epsilon, 1/2]^q\big\}$ and $\Theta_2 = \Omega_\beta \backslash \Theta_1$ (possibly an empty set), where Ω_β is given by (3.2). Following Robinson (1995b) and Shimotsu (2007), we have

$$\mathbb{P}\big(\|\widehat{\boldsymbol{d}} - \boldsymbol{d}_0\|_\infty > \delta\big) \le \mathbb{P}\Big(\inf_{\overline{N_\delta \cap \Omega_\beta}}\big\{L(\boldsymbol{d})\big\} \le 0\Big)$$
$$\le \mathbb{P}\Big(\inf_{\overline{N_\delta \cap \Theta_1}}\big\{L(\boldsymbol{d})\big\} \le 0\Big) + \mathbb{P}\Big(\inf_{\overline{\Theta_2}}\big\{L(\boldsymbol{d})\big\} \le 0\Big) := P_1 + P_2 \qquad \text{(A.2)}$$

where, for a given set $\mathcal{O}$, $\overline{\mathcal{O}}$ denotes the closure of $\mathcal{O}$. We shall first show that $P_1 \to 0$, as n tends to infinity. Rewrite $L(\boldsymbol{d})$ as

$$\begin{aligned}
L(\boldsymbol{d}) &= \log\big(\det\{\widehat{G}(\boldsymbol{d})\}\big) - \log\big(\det\{\widehat{G}(\boldsymbol{d}_0)\}\big) - 2\sum_{k=1}^{q}\theta_k\frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j)\\
&= \log\big(\det\{\widehat{G}(\boldsymbol{d})\}\big) - \log\big(\det\{\widehat{G}(\boldsymbol{d}_0)\}\big) + \log\left(\frac{2\pi m}{n}\right)^{-2\sum_k\theta_k} -\\
&\quad - 2\sum_{k=1}^{q}\theta_k\left(\frac{1}{m}\sum_{j=1}^{m}\log(j) - \log(m)\right) - \sum_{k=1}^{q}\log(2\theta_k+1)\\
&= \log\big(\mathcal{A}(\boldsymbol{d})\big) - \log\big(\mathcal{B}(\boldsymbol{d})\big) - \log\big(\mathcal{A}(\boldsymbol{d}_0)\big) + \log\big(\mathcal{B}(\boldsymbol{d}_0)\big) + \mathcal{R}(\boldsymbol{d})\\
&= Q_1(\boldsymbol{d}) - Q_2(\boldsymbol{d}) + \mathcal{R}(\boldsymbol{d}), \qquad \text{(A.3)}
\end{aligned}$$

where

$$Q_1(\boldsymbol{d}) := \log\big(\mathcal{A}(\boldsymbol{d})\big) - \log\big(\mathcal{B}(\boldsymbol{d})\big), \qquad Q_2(\boldsymbol{d}) := \log\big(\mathcal{A}(\boldsymbol{d}_0)\big) + \log\big(\mathcal{B}(\boldsymbol{d}_0)\big),$$
$$\mathcal{A}(\boldsymbol{d}) := \left(\frac{2\pi m}{n}\right)^{-2\sum_k\theta_k}\det\{\widehat{G}(\boldsymbol{d})\}, \qquad \mathcal{B}(\boldsymbol{d}) := \det\{G_0\}\prod_{k=1}^{q}\frac{1}{2\theta_k+1},$$
$$\text{and} \quad \mathcal{R}(\boldsymbol{d}) := 2\sum_{k=1}^{q}\theta_k\left(\log(m) - \frac{1}{m}\sum_{j=1}^{m}\log(j)\right) - \sum_{k=1}^{q}\log(2\theta_k+1).$$

By lemma 2 in Robinson (1995b), $\log(m) - m^{-1}\sum_{j=1}^{m}\log(j) = 1 + O(m^{-1}\log(m))$, so that

$$\mathcal{R}(\boldsymbol{d}) = \sum_{k=1}^{q}2\theta_k - \log(2\theta_k+1) + O\left(\frac{\log(m)}{m}\right).$$

Since $x - \log(x+1)$ has a unique global minimum in $(-1,\infty)$ at $x = 0$ and $x - \log(x+1) \ge x^2/4$, for $|x| \le 1$, it follows that

$$\inf_{\overline{N_\delta \cap \Theta_1}}\big\{\mathcal{R}(\boldsymbol{d})\big\} \ge \frac{1}{4}\Big(2\max_k\{\theta_k\}\Big)^2 \ge \delta^2 > 0\,.$$

As for $Q_1(\boldsymbol{d})$ and $Q_2(\boldsymbol{d})$ in (A.3), it suffices to show the existence of a function $h(\boldsymbol{d}) > 0$ satisfying

$$\text{(i)}\ \sup_{\Theta_1}\{|\mathcal{A}(\boldsymbol{d}) - h(\boldsymbol{d})|\} = o_{\mathbb{P}}(1); \quad \text{(ii)}\ h(\boldsymbol{d}) \geq \mathcal{B}(\boldsymbol{d}); \quad \text{(iii)}\ h(\boldsymbol{d}_0) = \mathcal{B}(\boldsymbol{d}_0),$$

as n goes to infinity, because (ii) implies $\inf_{\Theta_1}\{h(\boldsymbol{d})\} \geq \inf_{\Theta_1}\{\mathcal{B}(\boldsymbol{d})\} > 0$, so that, uniformly in Θ_1,

$$Q_1(\boldsymbol{d}) \geq \log\big(\mathcal{A}(\boldsymbol{d})\big) - \log\big(h(\boldsymbol{d})\big) = \log\big(h(\boldsymbol{d}) + o_{\mathbb{P}}(1)\big) - \log\big(h(\boldsymbol{d})\big) = o_{\mathbb{P}}(1), \tag{A.4}$$

and (iii) implies $Q_2(\boldsymbol{d}) = \log\big(h(\boldsymbol{d}_0) + o_{\mathbb{P}}(1)\big) - \log\big(h(\boldsymbol{d}_0)\big) = o_{\mathbb{P}}(1)$, from which $P_1 \to 0$ follows. To show (i), recall that

$$\begin{aligned}\Lambda_j(\boldsymbol{d})^{-1} &= \operatorname*{diag}_{k\in\{1,\cdots,q\}}\{\lambda_j^{d_k}\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)d_k/2}\} = \operatorname*{diag}_{k\in\{1,\cdots,q\}}\{\lambda_j^{(d_k-d_k^0)}\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(d_k-d_k^0)/2} \times \lambda_j^{d_k^0}\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)d_k^0/2}\} \\ &= \Lambda_j(\boldsymbol{d}-\boldsymbol{d}_0)^{-1}\Lambda_j(\boldsymbol{d}_0)^{-1} = \Lambda_j(\boldsymbol{\theta})^{-1}\Lambda_j(\boldsymbol{d}_0)^{-1}, \end{aligned} \tag{A.5}$$

so that we can write

$$\begin{aligned}\mathcal{A}(\boldsymbol{d}) &= \left(\frac{2\pi m}{n}\right)^{-2\sum_k\theta_k} \times \det\left\{\frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\big[\Lambda_j(\boldsymbol{\theta})^{-1}\Lambda_j(\boldsymbol{d}_0)^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}'\,\overline{\Lambda_j(\boldsymbol{\theta})^{-1}}'\big]\right\} \\ &= \det\left\{\frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\big[M_j(\boldsymbol{\theta})\Lambda_j(\boldsymbol{d}_0)^{-1}\big(f(\lambda_j)+o_{\mathbb{P}}(n^{-\beta})\big)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}'\,\overline{M_j(\boldsymbol{\theta})}'\big]\right\} \\ &= \det\left\{\frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\big[M_j(\boldsymbol{\theta})G_0\overline{M_j(\boldsymbol{\theta})}'\big] + \frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\big[M_j(\boldsymbol{\theta})\Lambda_j(\boldsymbol{d}_0)^{-1}o_{\mathbb{P}}(n^{-\beta})\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}'\,\overline{M_j(\boldsymbol{\theta})}'\big]\right\}, \end{aligned} \tag{A.6}$$

where $M_j(\boldsymbol{\theta}) := \mathrm{diag}_{k\in\{1,\cdots,q\}}\left\{\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)\theta_k/2}(j/m)^{\theta_k}\right\}$. To determine the function h in (i), we first show that the second term on the RHS of (A.6) is $o_{\mathbb{P}}(1)$. This follows by noticing that its (r,s)-th element is given by

$$\begin{aligned}\frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\left[\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(\hat{d}_r-\hat{d}_s)/2}\left(\frac{j}{m}\right)^{\theta_r+\theta_s}\lambda_j^{d_r^0+d_s^0}\right] &\leq \frac{\mathcal{C}}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{\hat{d}_r+\hat{d}_s}\left(\frac{m}{n}\right)^{d_r^0+d_s^0}o_{\mathbb{P}}(n^{-\beta}) \\ &= \frac{1}{\hat{d}_r+\hat{d}_s+1}\left(\frac{2\pi m}{n}\right)^{\hat{d}_r+\hat{d}_s}\big[O(m^{\beta-1})+1\big]\left(\frac{m}{n}\right)^{d_r^0+d_s^0}o_{\mathbb{P}}(n^{-\beta}) = o_{\mathbb{P}}(1)\end{aligned}$$

for $\mathcal{C}$ a constant, where the penultimate equality follows from lemma 1 in Robinson (1995b) and the last one follows since $\widehat{\boldsymbol{d}} \in \Omega_\beta$. Hence

$$\mathcal{A}(\boldsymbol{d}) = \det\left\{\frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\big[M_j(\boldsymbol{\theta})G_0\overline{M_j(\boldsymbol{\theta})}'\big] + o_{\mathbb{P}}(1)\right\}. \tag{A.7}$$

Upon defining the matrices

$$\mathcal{E}(\boldsymbol{\theta}) := \big(\mathrm{e}^{-\mathrm{i}\pi(\theta_r-\theta_s)/2}\big)_{r,s=1}^{q} \quad \text{and} \quad \mathcal{M}(\boldsymbol{\theta}) := \left(\frac{1}{1+\theta_r+\theta_s}\right)_{r,s=1}^{q}$$

from the proof of theorem 1 in Shimotsu (2007), it follows that the function

$$h(\boldsymbol{d}) := \det\Big\{\mathrm{Re}\big[\mathcal{E}(\boldsymbol{\theta})\big] \odot \mathcal{M}(\boldsymbol{\theta}) \odot G_0\Big\},$$

where $\odot$ denotes the Hadamard product, satisfies the conditions (i), (ii) and (iii) in (A.4) (see the argument following (11) in Shimotsu, 2007, p.292).

Now we move to bound P_2 in (A.2). Expression (A.5) can be used to rewrite $L(\boldsymbol{d})$ as

$$L(\boldsymbol{d}) \;=\; \log\big(\det\{\widehat{G}(\boldsymbol{d})\}\big) - \log\big(\det\{\widehat{G}(\boldsymbol{d}_0)\}\big) - 2\sum_{k=1}^{q}\theta_k\frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j)$$

$$= \log\Big(\det\big\{\widehat{\mathcal{D}}(\boldsymbol{d})\big\}\Big) - \log\Big(\det\big\{\widehat{\mathcal{D}}(\boldsymbol{d}_0)\big\}\Big), \tag{A.8}$$

where

$$\widehat{\mathcal{D}}(\boldsymbol{d}) := \frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\big[\mathcal{P}_j(\boldsymbol{\theta})\Lambda_j(\boldsymbol{d}_0)^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'\,\overline{\mathcal{P}_j(\boldsymbol{\theta})}\,'\big],$$

with

$$\mathcal{P}_j(\boldsymbol{\theta}) := \underset{k\in\{1,\cdots,q\}}{\mathrm{diag}}\left\{\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)\theta_k/2}\left(\frac{j}{p}\right)^{\theta_k}\right\} \quad \text{and} \quad p := \exp\left(\frac{1}{m}\sum_{j=1}^{m}\log(j)\right),$$

and, as m tends to infinity, $p \sim m/\mathrm{e}$. Observe that $\widehat{\mathcal{D}}(\boldsymbol{d})$ is positive semidefinite since each summand of $\widehat{\mathcal{D}}$ is. For $\kappa \in (0,1)$, define

$$\widehat{\mathcal{D}}_\kappa(\boldsymbol{d}) := \frac{1}{m}\sum_{j=[m\kappa]}^{m}\mathrm{Re}\big[\mathcal{P}_j(\boldsymbol{\theta})\Lambda_j(\boldsymbol{d}_0)^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'\,\overline{\mathcal{P}_j(\boldsymbol{\theta})}\,'\big]$$

and

$$\mathcal{Q}_\kappa(\boldsymbol{d}) := \frac{1}{m}\sum_{j=[m\kappa]}^{m}\mathrm{Re}\big[\mathcal{P}_j(\boldsymbol{\theta})G_0\overline{\mathcal{P}_j(\boldsymbol{\theta})}\,'\big],$$

where $[x]$ denotes the integer part of x.

$$\begin{aligned}\widehat{\mathcal{D}}_\kappa(\boldsymbol{d}) &= \frac{1}{m}\sum_{j=[m\kappa]}^{m}\mathrm{Re}\big[\mathcal{P}_j(\boldsymbol{\theta})\Lambda_j(\boldsymbol{d}_0)^{-1}\big(f(\lambda_j)+o_{\mathbb{P}}(n^{-\beta})\big)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'\,\overline{\mathcal{P}_j(\boldsymbol{\theta})}\,'\big]\\ &= \mathcal{Q}_\kappa(\boldsymbol{d}) + \frac{1}{m}\sum_{j=[m\kappa]}^{m}\mathrm{Re}\big[\mathcal{P}_j(\boldsymbol{\theta})\Lambda_j(\boldsymbol{d}_0)^{-1}o_{\mathbb{P}}(n^{-\beta})\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'\,\overline{\mathcal{P}_j(\boldsymbol{\theta})}\,'\big],\end{aligned} \tag{A.9}$$

where the last equality follows from lemma 5.4 in Shimotsu and Phillips (2005). The (r,s)-th element of the third term on the RHS of (A.9) is given by

$$\begin{aligned}\mathrm{Re}\bigg[\frac{1}{m}\sum_{j=[m\kappa]}^{m}\left(\frac{j}{p}\right)^{\theta_r+\theta_s}\left(\frac{2\pi j}{n}\right)^{d_r^0+d_s^0}\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(\hat{d}_r-\hat{d}_s)/2}o_{\mathbb{P}}(n^{-\beta})\bigg] &=\\ = O(1)\left(\frac{m}{p}\right)^{\theta_r+\theta_s}\left(\frac{m}{n}\right)^{d_r^0+d_s^0}o_{\mathbb{P}}(n^{-\beta})\,\frac{1}{m}\sum_{j=[m\kappa]}^{m}\left(\frac{j}{m}\right)^{2(d_r^0+d_s^0)-(\widehat{d}_r+\widehat{d}_s)}&\\ = O(1)o_{\mathbb{P}}(1)\left(\frac{m}{p}\right)^{\theta_r+\theta_s}O(1) = o_{\mathbb{P}}(1),&\end{aligned}$$

uniformly in $\boldsymbol{\theta} \in \Theta_2$, since $\boldsymbol{d}_0 \in \Omega_\beta$, $\beta \in (0,1)$ and Assumption **A**4, where the penultimate equality follows from lemma 5.4 in Shimotsu and Phillips (2005). Hence

$$\sup_{\Theta_2}\big\{\big|\det\{\widehat{\mathcal{D}}(\boldsymbol{d})\} - \det\{\mathcal{Q}_\kappa(\boldsymbol{d})\}\big|\big\} = o_{\mathbb{P}}(1).$$

The proof now follows viz a viz (with the obvious notational identification) from the proof of theorem 1 in Shimotsu (2007), p.294 (see the argument following equation (16)). We thus conclude that $P_2 \to 0$, as n tends to infinity, and the proof is complete. ■

Proof of Lemma 3.2:

For fixed $r,s \in \{1,\cdots,q\}$, let $\mathscr{A}_{uv} := \sum_{j=u}^{v}\mathscr{A}_j$ and $\mathscr{B}_{uv} := \sum_{j=u}^{v}\mathscr{B}_j$, where

$$\mathscr{A}_j := \mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(d_r^0-d_s^0)/2}\lambda_j^{d_r^0+d_s^0}\big[f_n^{rs}(\lambda_j) - \big(A(\lambda_j)\big)_{r\cdot}I_{\boldsymbol{\varepsilon}}(\lambda_j)\big(\overline{A(\lambda_j)}\,'\big)_{\cdot s}\big], \tag{A.10}$$

and

$$\mathscr{B}_j := \mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(d_r^0-d_s^0)/2}\lambda_j^{d_r^0+d_s^0}\big(A(\lambda_j)\big)_{r.}I_{\boldsymbol{\varepsilon}}(\lambda_j)\big(\overline{A(\lambda_j)}\,'\big)_{.s} - G_0^{rs}. \tag{A.11}$$

Hence, for each j, $\mathscr{A}_j + \mathscr{B}_j = \mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(d_r^0-d_s^0)/2}\lambda_j^{d_r^0+d_s^0} f_n^{rs}(\lambda_j) - G_0^{rs}$. For fixed $u \le j \le v$, we have

$$\mathbb{E}\big(|\mathscr{A}_j|\big) = \mathbb{E}\Big(\lambda_j^{d_r^0+d_s^0}\Big|f_n^{rs}(\lambda_j) - \big(A(\lambda_j)\big)_{r.}I_{\boldsymbol{\varepsilon}}(\lambda_j)\big(\overline{A(\lambda_j)}\,'\big)_{.s}\Big|\Big) = o(1),$$

which yields $\max\limits_{r,s}\big\{\mathbb{E}\big(\big|\sum_{j=u}^{v}\mathscr{A}_j\big|\big)\big\} = o(v-u+1)$ and the result on $\mathscr{A}_{uv}$ follows. As for $\mathscr{B}_j$, from the proof of lemma 1(a) in Shimotsu (2007) (notice that $\mathscr{B}_j$ does not depend on f_n) it follows that $\mathscr{B}_{uv} = o_{\mathbb{P}}(v)$ uniformly in u and v and the desired result follows. ■

Proof of Theorem 3.2:

From a careful inspection of the proof of Theorem 3.1, we observe that it suffices to show (with the same notation as in that proof)

$$\frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\Big[M_j(\boldsymbol{\theta})\Lambda_j(\boldsymbol{d}_0)^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'\,\overline{M_j(\boldsymbol{\theta})}\,'\Big] = \frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\big[M_j(\boldsymbol{\theta})G_0\overline{M_j(\boldsymbol{\theta})}\,'\big] + o_{\mathbb{P}}(1), \tag{A.12}$$

uniformly in Θ_1 and that $\widehat{\mathcal{D}}_\kappa(\boldsymbol{d}) - \mathcal{Q}_\kappa(\boldsymbol{d}) = o_{\mathbb{P}}(1)$, uniformly in Θ_2. To show (A.12), notice that the (r,s)-th component of the LHS in (A.12) is given by

$$\frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\bigg[\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(\theta_r-\theta_s)/2}\Big(\frac{j}{m}\Big)^{\theta_r+\theta_s} f_n^{rs}(\lambda_j)\Big(\Lambda_j^{(r)}(\boldsymbol{d}_0)\overline{\Lambda_j^{(s)}(\boldsymbol{d}_0)}\,'\Big)^{-1}\bigg].$$

Summation by parts (see Zygmund, 2002, p.3) yields

$$\begin{aligned}
&\sup_{\Theta_1}\bigg\{\bigg|\frac{1}{m}\sum_{j=1}^{m}\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(\theta_r-\theta_s)/2}\Big(\frac{j}{m}\Big)^{\theta_r+\theta_s}\Big[f_n^{rs}(\lambda_j)\Big(\Lambda_j^{(r)}(\boldsymbol{d}_0)\overline{\Lambda_j^{(s)}(\boldsymbol{d}_0)}\,'\Big)^{-1} - G_0^{rs}\Big]\bigg|\bigg\} \le \\
&\le \frac{1}{m}\sum_{k=1}^{m-1}\sup_{\Theta_1}\bigg\{\bigg|\mathrm{e}^{\mathrm{i}(\lambda_k-\pi)(\theta_r-\theta_s)/2}\Big(\frac{k}{m}\Big)^{\theta_r+\theta_s} - \mathrm{e}^{\mathrm{i}(\lambda_{k+1}-\pi)(\theta_r-\theta_s)/2}\Big(\frac{k+1}{m}\Big)^{\theta_r+\theta_s}\bigg|\bigg\}\times \\
&\quad\times\bigg|\sum_{j=1}^{k}\Big[f_n^{rs}(\lambda_j)\Big(\Lambda_j^{(r)}(\boldsymbol{d}_0)\overline{\Lambda_j^{(s)}(\boldsymbol{d}_0)}\,'\Big)^{-1} - G_0^{rs}\Big]\bigg| + \bigg|\frac{1}{m}\sum_{j=1}^{m}\Big[f_n^{rs}(\lambda_j)\Big(\Lambda_j^{(r)}(\boldsymbol{d}_0)\overline{\Lambda_j^{(s)}(\boldsymbol{d}_0)}\,'\Big)^{-1} - G_0^{rs}\Big]\bigg| \\
&\le \mathcal{C}\sum_{k=1}^{m-1}\Big(\frac{k}{m}\Big)^{2\epsilon}\frac{1}{k^2}\bigg|\sum_{j=1}^{k}\Big[f_n^{rs}(\lambda_j)\Big(\Lambda_j^{(r)}(\boldsymbol{d}_0)\overline{\Lambda_j^{(s)}(\boldsymbol{d}_0)}\,'\Big)^{-1} - G_0^{rs}\Big]\bigg| + \\
&\quad+\bigg|\frac{1}{m}\sum_{j=1}^{m}\Big[f_n^{rs}(\lambda_j)\Big(\Lambda_j^{(r)}(\boldsymbol{d}_0)\overline{\Lambda_j^{(s)}(\boldsymbol{d}_0)}\,'\Big)^{-1} - G_0^{rs}\Big]\bigg|,
\end{aligned} \tag{A.13}$$

where $0 < \mathcal{C} < \infty$ is a constant. Now, from Lemma 3.2,

$$\begin{aligned}
\sum_{k=1}^{m-1}\Big(\frac{k}{m}\Big)^{2\epsilon}\frac{1}{k^2}\bigg|\sum_{j=1}^{k}\Big[f_n^{rs}(\lambda_j)\Big(\Lambda_j^{(r)}(\boldsymbol{d}_0)\overline{\Lambda_j^{(s)}(\boldsymbol{d}_0)}\,'\Big)^{-1} - G_0^{rs}\Big]\bigg| &\le \sum_{k=1}^{m-1}\Big(\frac{k}{m}\Big)^{2\epsilon}\frac{1}{k^2}\Big(|\mathscr{A}_{1k}| + |\mathscr{B}_{1k}|\Big) \\
&= \frac{1}{m^{2\epsilon}}\sum_{k=1}^{m-1}k^{2(\epsilon-1)}|\mathscr{A}_{1k}| + \frac{1}{m^{2\epsilon}}\sum_{k=1}^{m-1}k^{2(\epsilon-1)}o_{\mathbb{P}}(k).
\end{aligned} \tag{A.14}$$

The RHS of (A.14) is $o_{\mathbb{P}}(1)$ uniformly in (r,s) by lemma 1 in Robinson (1995b), Lemma 3.2 and Chebyshev's inequality since $\mathbb{E}\big(m^{-2\epsilon}\sum_{k=1}^{m-1}k^{2(\epsilon-1)}|\mathscr{A}_{1k}|\big) = o(1)$, uniformly in (r,s). The other term in (A.13) is also $o_{\mathbb{P}}(1)$, uniformly in (r,s), by the same argument and, hence, (A.12) follows. On the other hand, the (r,s)-th element of $\widehat{\mathcal{D}}_\kappa(\boldsymbol{d}) - \mathcal{Q}_\kappa(\boldsymbol{d})$ is given by

$$\frac{1}{m}\sum_{j=[m\kappa]}^{m}\mathrm{Re}\bigg[\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(\theta_r-\theta_s)/2}\Big(\frac{j}{p}\Big)^{\theta_r+\theta_s}\Big[f_n^{rs}(\lambda_j)\Big(\Lambda_j^{(r)}(\boldsymbol{d}_0)\overline{\Lambda_j^{(s)}(\boldsymbol{d}_0)}\,'\Big)^{-1} - G_0^{rs}\Big]\bigg] =$$

$$= \left(\frac{m}{p}\right)^{\theta_r+\theta_s}\mathrm{Re}\Bigg[\frac{1}{m}\sum_{j=[m\kappa]}^{m}\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(\theta_r-\theta_s)/2}\left(\frac{j}{m}\right)^{\theta_r+\theta_s}\Big[f_n^{rs}(\lambda_j)\Big(\Lambda_j^{(r)}(\boldsymbol{d}_0)\overline{\Lambda_j^{(s)}(\boldsymbol{d}_0)}\,'\Big)^{-1}-G_0^{rs}\Big]\Bigg] = o_{\mathbb{P}}(1),$$

uniformly in $\boldsymbol{\theta}\in\Theta_2$, where the last equality is derived similarly to (A.13) from summation by parts and lemma 5.4 in Shimotsu and Phillips (2005). This completes the proof. ■

Proof of Lemma 4.1:

(a) For $r,s\in\{1,\cdots,q\}$ fixed, we have

$$\sum_{j=1}^{v}\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(d_r^0-d_s^0)/2}\lambda_j^{d_r^0+d_s^0}\Big[f_n^{rs}(\lambda_j)-\big(A(\lambda_j)\big)_{r\cdot}I_{\boldsymbol{\varepsilon}}(\lambda_j)\big(\overline{A(\lambda_j)}\,'\big)_{\cdot s}\Big]\leq$$

$$\leq\Big[1+O\big((\lambda_j-\pi)(d_r^0-d_s^0)/2\big)\Big]\sum_{j=1}^{v}\lambda_j^{d_r^0+d_s^0}\Big[f_n^{rs}(\lambda_j)-\big(A(\lambda_j)\big)_{r\cdot}I_{\boldsymbol{\varepsilon}}(\lambda_j)\big(\overline{A(\lambda_j)}\,'\big)_{\cdot s}\Big]$$

$$\leq O(1)\max\{1,n^{|d_r^0+d_s^0|}\}\max_{1\leq v\leq m}\left\{\sum_{j=1}^{v}\Big[f_n^{rs}(\lambda_j)-\big(A(\lambda_j)\big)_{r\cdot}I_{\boldsymbol{\varepsilon}}(\lambda_j)\big(\overline{A(\lambda_j)}\,'\big)_{\cdot s}\Big]\right\}$$

$$= o_{\mathbb{P}}\left(\frac{m^2}{n}\right) = o_{\mathbb{P}}\left(\frac{\sqrt{m}}{\log(m)}\right),$$

where the last equality follows from **B**4, and (4.2) follows.

(b) Rewrite the argument of the summation in (4.3) as $\mathscr{A}_j+\mathscr{B}_j+\mathscr{C}_j$, where

$$\mathscr{A}_j := \mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(d_r^0-d_s^0)/2}\lambda_j^{d_r^0+d_s^0}\Big[f_n^{rs}(\lambda_j)-\big(A(\lambda_j)\big)_{r\cdot}I_{\boldsymbol{\varepsilon}}(\lambda_j)\big(\overline{A(\lambda_j)}\,'\big)_{\cdot s}\Big],$$

$$\mathscr{B}_j := \mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(d_r^0-d_s^0)/2}\lambda_j^{d_r^0+d_s^0}\Big[\big(A(\lambda_j)\big)_{r\cdot}I_{\boldsymbol{\varepsilon}}(\lambda_j)\big(\overline{A(\lambda_j)}\,'\big)_{\cdot s}-f_{rs}(\lambda_j)\Big],$$

$$\mathscr{C}_j := \mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(d_r^0-d_s^0)/2}\lambda_j^{d_r^0+d_s^0}f_{rs}(\lambda_j)-G_0^{rs}.$$

Part (a) yields $\max_{r,s}\Big\{\sum_{j=1}^{v}\big|\mathscr{A}_j\big|\Big\} = o_{\mathbb{P}}\Big(m^{1/2}\big(\log(m)\big)^{-1}\Big)$, while, from the proof of lemma 1(b2) in Shimotsu (2007), we obtain $\max_{r,s}\Big\{\sum_{j=1}^{v}\big|\mathscr{B}_j\big|\Big\} = O_{\mathbb{P}}\big(m^{1/2}\log(m)\big)$. Assumption **B**1 implies $\max_{r,s}\Big\{\sum_{j=1}^{v}\big|\mathscr{C}_j\big|\Big\} = O\big(m^{\alpha+1}n^{-\alpha}\big)$. The result now follows by noticing that $m^{1/2}\log(m)^{-1}=O\big(m^{1/2}\log(m)\big)$. ■

Proof of Theorem 4.1:

The idea of the proof is similar to that of Lobato (1999) with similar adaptations as in Shimotsu (2007). By hypothesis,

$$\mathbf{0} = \frac{\partial S(\boldsymbol{d})}{\partial\boldsymbol{d}}\bigg|_{\widehat{\boldsymbol{d}}} = \frac{\partial S(\boldsymbol{d})}{\partial\boldsymbol{d}}\bigg|_{\boldsymbol{d}_0}+\left(\frac{\partial^2 S(\boldsymbol{d})}{\partial\boldsymbol{d}\partial\boldsymbol{d}'}\bigg|_{\overline{\mathbf{d}}}\right)(\widehat{\boldsymbol{d}}-\boldsymbol{d}_0),$$

with probability tending to 1, as n tends to infinity, for some $\overline{\mathbf{d}}$ such that $\|\overline{\mathbf{d}}-\boldsymbol{d}_0\|_\infty\leq\|\widehat{\boldsymbol{d}}-\boldsymbol{d}_0\|_\infty$. We observe that $\widehat{\boldsymbol{d}}$ has the stated limiting distribution if

$$\sqrt{m}\,\frac{\partial S(\boldsymbol{d})}{\partial\boldsymbol{d}}\bigg|_{\boldsymbol{d}_0}\xrightarrow{\ d\ } N(0,\Sigma) \qquad \text{(A.15)}$$

and

$$\frac{\partial^2 S(\boldsymbol{d})}{\partial\boldsymbol{d}\partial\boldsymbol{d}'}\bigg|_{\overline{\mathbf{d}}}\xrightarrow{\ \mathbb{P}\ }\Sigma. \qquad \text{(A.16)}$$

We shall prove (A.15) first. Observe that, for $r \in \{1, \cdots, q\}$,

$$\sqrt{m}\,\frac{\partial S(\boldsymbol{d})}{\partial d_r} = -\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\log(\lambda_j) + \mathrm{tr}\left[\widehat{G}(\boldsymbol{d})^{-1}\sqrt{m}\,\frac{\partial \widehat{G}(\boldsymbol{d})}{\partial d_r}\right].$$

Let $\mathrm{I}_{(r)}$ denote a $q \times q$ matrix whose (r, r)-th element is 1 and all other elements are zero. Define a function $\varphi : (0, \infty) \to \mathbb{C}$ by

$$\varphi(x) := \log(x) + \mathrm{i}\left(\frac{x-\pi}{2}\right). \tag{A.17}$$

Since $\Lambda_j(\boldsymbol{d})^{-1} = \operatorname*{diag}_{k\in\{1,\cdots,q\}}\big\{\lambda_j^{d_k}\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)d_k/2}\big\}$ and $\mathrm{Re}\big[(a+\mathrm{i}b)(c+\mathrm{i}d)\big] = ac - bd$, we can write

$$\begin{aligned}
\sqrt{m}\,\frac{\partial \widehat{G}(\boldsymbol{d})}{\partial d_r}\bigg|_{\boldsymbol{d}_0} &= \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\mathrm{Re}\Big[\varphi(\lambda_j)\Lambda_j(\boldsymbol{d}_0)^{-1}\mathrm{I}_{(r)}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'\Big] + \\
&\quad + \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\mathrm{Re}\Big[\overline{\varphi(\lambda_j)}\Lambda_j(\boldsymbol{d}_0)^{-1}f_n(\lambda_j)\mathrm{I}_{(r)}\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'\Big] \\
&= \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\log(\lambda_j)\mathrm{Re}\Big[\Lambda_j(\boldsymbol{d}_0)^{-1}\big(\mathrm{I}_{(r)}f_n(\lambda_j) + f_n(\lambda_j)\mathrm{I}_{(r)}\big)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'\Big] + \\
&\quad + \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\left[\frac{\lambda_j-\pi}{2}\right]\mathrm{Im}\Big[\Lambda_j(\boldsymbol{d}_0)^{-1}\big(-\mathrm{I}_{(r)}f_n(\lambda_j) + f_n(\lambda_j)\mathrm{I}_{(r)}\big)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'\Big], \\
&:= \mathscr{H}_1(r) + \mathscr{H}_2(r).
\end{aligned} \tag{A.18}$$

Therefore, for η an arbitrary vector in $\mathbb{R}^q$, from (A.18) we obtain

$$\begin{aligned}
\eta'\sqrt{m}\,\frac{\partial S(\boldsymbol{d})}{\partial \boldsymbol{d}}\bigg|_{\boldsymbol{d}_0} &= \sum_{k=1}^{q}\eta_k\sqrt{m}\,\frac{\partial S(\boldsymbol{d})}{\partial d_k}\bigg|_{\boldsymbol{d}_0} = \\
&= \sum_{k=1}^{q}\eta_k\bigg[-\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\log(\lambda_j) + \mathrm{tr}\big[\widehat{G}(\boldsymbol{d}_0)^{-1}\mathscr{H}_1(k)\big]\bigg] + \sum_{k=1}^{q}\eta_k\,\mathrm{tr}\big[\widehat{G}(\boldsymbol{d}_0)^{-1}\mathscr{H}_2(k)\big], \\
&:= \mathscr{R}_1 + \mathscr{R}_2.
\end{aligned}$$

We analyze $\mathscr{R}_1$ first. By letting

$$a_j := \log(\lambda_j) - \frac{1}{m}\sum_{k=1}^{m}\log(\lambda_k) = \log(j) - \frac{1}{m}\sum_{k=1}^{m}\log(k) = O\big(\log(m)\big),$$

we can write

$$\begin{aligned}
-\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\log(\lambda_j) + \mathrm{tr}\big[\widehat{G}(\boldsymbol{d}_0)^{-1}\mathscr{H}_1(k)\big] &= \mathrm{tr}\left[\widehat{G}(\boldsymbol{d}_0)^{-1}\bigg(\mathscr{H}_1(k) - \frac{2}{\sqrt{m}}\sum_{j=1}^{m}\log(\lambda_j)\widehat{G}(\boldsymbol{d}_0)\mathrm{I}_{(k)}\bigg)\right] \\
&= \mathrm{tr}\left[\widehat{G}(\boldsymbol{d}_0)^{-1}\frac{2}{\sqrt{m}}\sum_{j=1}^{m}a_j\mathrm{Re}\Big[\Lambda_j(\boldsymbol{d}_0)^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'\Big]\mathrm{I}_{(k)}\right].
\end{aligned} \tag{A.19}$$

By Lemma 4.1(b), (A.19) can be written as

$$\Big[(G_0^{-1})_{k\cdot} + o_{\mathbb{P}}(1)\Big]\frac{2}{\sqrt{m}}\sum_{j=1}^{m}a_j\bigg(\mathrm{Re}\Big[\Lambda_j(\boldsymbol{d}_0)^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'\Big]\bigg)_{\cdot k}.$$

Now, by Lemma 4.1(a),

$$\bigg(\sum_{j=1}^{m}a_j\Lambda_j(\boldsymbol{d}_0)^{-1}\Big(f_n(\lambda_j) - A(\lambda_j)I_{\boldsymbol{\varepsilon}}(\lambda_j)\overline{A(\lambda_j)}\,'\Big)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'\bigg)_{rs} \leq$$

$$\leq O\big(\log(m)\big)\max_{v=1,\cdots,m}\bigg\{\sum_{j=1}^{v} \mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(d_r^0-d_s^0)/2}\lambda_j^{d_r^0+d_s^0}\Big(f_n^{rs}(\lambda_j) - \big(A(\lambda_j)\big)_{r\cdot}I_{\boldsymbol{\varepsilon}}(\lambda_j)\big(\overline{A(\lambda_j)}\,'\big)_{\cdot s}\Big)\bigg\}$$
$$= O\big(\log(m)\big)o_{\mathbb{P}}\bigg(\frac{\sqrt{m}}{\log(m)}\bigg) \;=\; o_{\mathbb{P}}(\sqrt{m}),$$

uniformly in $r,s\in\{1,\cdots,q\}$. Therefore,

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m} a_j\Lambda_j(\boldsymbol{d}_0)^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,' =$$
$$= \frac{1}{\sqrt{m}}\sum_{j=1}^{m} a_j\Lambda_j(\boldsymbol{d}_0)^{-1}A(\lambda_j)I_{\boldsymbol{\varepsilon}}(\lambda_j)\overline{A(\lambda_j)}\,'\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,' + \frac{1}{\sqrt{m}}o_{\mathbb{P}}(\sqrt{m})$$
$$= \frac{1}{\sqrt{m}}\sum_{j=1}^{m} a_j\Big[\Lambda_j(\boldsymbol{d}_0)^{-1}A(\lambda_j)I_{\boldsymbol{\varepsilon}}(\lambda_j)\overline{A(\lambda_j)}\,'\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,' - G_0\Big] + o_{\mathbb{P}}(1), \tag{A.20}$$

where the last equality follows from $\sum_{j=1}^{m}a_j = 0$. The proof of (A.15) now follows viz a viz from the proof of theorem 2 in Shimotsu (2007) p.296, by noticing that, with the appropriate notational identification, (A.20) is (21) in the aforementioned theorem.

We move to prove (A.16). For fixed $\delta>0$, let $\boldsymbol{\theta} := \boldsymbol{d}-\boldsymbol{d}_0$ and define

$$\mathcal{M} := \Big\{\boldsymbol{d} : \log(n)^4\|\boldsymbol{d}-\boldsymbol{d}_0\|_\infty < \delta\Big\} = \Big\{\boldsymbol{\theta} : \log(n)^4\|\boldsymbol{\theta}\|_\infty < \delta\Big\}.$$

First we show that $\mathbb{P}\big(\overline{\mathbf{d}}\in\mathcal{M}\big)\to 1$, as $n\to\infty$. Assuming the same notation as in the proof of Theorem 3.1, recall the decomposition of $L(\boldsymbol{d}) = S(\boldsymbol{d}) - S(\boldsymbol{d}_0)$ given in expression (A.3). By applying the same argument as in the proof of Theorem 3.1, we first obtain

$$\inf_{\Theta_1\backslash\mathcal{M}}\big\{\mathcal{R}(\boldsymbol{d})\big\} \geq \delta^2\log(n)^8,$$

and upon applying Lemma 4.1, we obtain

$$\sup_{\Theta_1}\Big\{\big|\mathcal{A}(\boldsymbol{d}) - h(\boldsymbol{d})\big|\Big\} = O_{\mathbb{P}}\bigg(\frac{m^\alpha}{n^\alpha} + \frac{\log(m)}{m^{2\epsilon}} + \frac{m}{n}\bigg).$$

It follows, uniformly in Θ_1 (cf. Shimotsu, 2007, p.300), that

$$\log\big(\mathcal{A}(\boldsymbol{d})\big) - \log\big(\mathcal{B}(\boldsymbol{d})\big) \geq \log\big(h(\boldsymbol{d}) + o_{\mathbb{P}}\big(\log(n)^{-8}\big)\big) - \log\big(h(\boldsymbol{d})\big) = o_{\mathbb{P}}\big(\log(n)^{-8}\big)$$
$$\log\big(\mathcal{A}(\boldsymbol{d}_0)\big) - \log\big(\mathcal{B}(\boldsymbol{d}_0)\big) = \log\big(h(\boldsymbol{d}_0) + o_{\mathbb{P}}\big(\log(n)^{-8}\big)\big) - \log\big(h(\boldsymbol{d}_0)\big) = o_{\mathbb{P}}\big(\log(n)^{-8}\big).$$

Hence, $\mathbb{P}\big(\inf_{\Theta_1\backslash\mathcal{M}} L(\boldsymbol{d})\leq 0\big)\to 0$ and $\mathbb{P}\big(\overline{\mathbf{d}}\in\mathcal{M}\big)\to 1$, as $n\to\infty$, follows.

Now, observe that

$$\frac{\partial^2 S(\boldsymbol{d})}{\partial d_r\partial d_s} = \mathrm{tr}\bigg[-\widehat{G}(\boldsymbol{d})^{-1}\frac{\partial\widehat{G}(\boldsymbol{d})}{\partial d_r}\widehat{G}(\boldsymbol{d})^{-1}\frac{\partial\widehat{G}(\boldsymbol{d})}{\partial d_s} + \widehat{G}(\boldsymbol{d})^{-1}\frac{\partial^2\widehat{G}(\boldsymbol{d})}{\partial d_r\partial d_s}\bigg].$$

The derivatives of $\widehat{G}(\boldsymbol{d})$ are given by

$$\frac{\partial\widehat{G}(\boldsymbol{d})}{\partial d_r} = \frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\Big[\varphi(\lambda_j)\mathrm{I}_{(r)}\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}\,'\Big] + \frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\Big[\overline{\varphi(\lambda_j)}\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}\,'\mathrm{I}_{(r)}\Big], \tag{A.21}$$

and

$$\frac{\partial^2\widehat{G}(\boldsymbol{d})}{\partial d_r\partial d_s} = \frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\Big[\varphi(\lambda_j)^2\mathrm{I}_{(r)}\mathrm{I}_{(s)}\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}\,'\Big]+$$
$$+\frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\Big[\big|\overline{\varphi(\lambda_j)}\big|^2\mathrm{I}_{(r)}\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}\,'\mathrm{I}_{(s)}\Big]+$$

$$+\frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\Big[\big|\overline{\varphi(\lambda_j)}\big|^2\mathrm{I}_{(s)}\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}{}'\mathrm{I}_{(r)}\Big]+$$
$$+\frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\Big[\overline{\varphi(\lambda_j)}^{\,2}\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}{}'\mathrm{I}_{(r)}\mathrm{I}_{(s)}\Big],$$

where φ is given by (A.17). For $k=0,1,2$, let

$$\mathcal{R}_k(\boldsymbol{d})=\frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j)^k\mathrm{Re}\Big[\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}{}'\Big]$$
$$\mathcal{I}_k(\boldsymbol{d})=\frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j)^k\mathrm{Im}\Big[\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}{}'\Big],$$

so that, we can write

$$\frac{\partial\widehat{G}(\boldsymbol{d})}{\partial d_r}=\mathrm{I}_{(r)}\mathcal{R}_1(\boldsymbol{d})+\mathcal{R}_1(\boldsymbol{d})\mathrm{I}_{(r)}+\frac{\pi}{2}\Big(\mathrm{I}_{(r)}\mathcal{I}_0(\boldsymbol{d})-\mathcal{I}_0(\boldsymbol{d})\mathrm{I}_{(r)}\Big)+o_{\mathbb{P}}\left(\frac{1}{\log(n)}\right),\tag{A.22}$$

and

$$\frac{\partial^2\widehat{G}(\boldsymbol{d})}{\partial d_r\partial d_s}=\frac{\pi^2}{4}\Big[\mathrm{I}_{(r)}\mathrm{I}_{(s)}\mathcal{R}_0(\boldsymbol{d})+\mathrm{I}_{(r)}\mathcal{R}_0(\boldsymbol{d})\mathrm{I}_{(s)}+\mathrm{I}_{(s)}\mathcal{R}_0(\boldsymbol{d})\mathrm{I}_{(r)}+\mathcal{R}_0(\boldsymbol{d})\mathrm{I}_{(s)}\mathrm{I}_{(r)}\Big]+$$
$$+\pi\Big[\mathrm{I}_{(r)}\mathrm{I}_{(s)}\mathcal{I}_1(\boldsymbol{d})+\mathcal{I}_1(\boldsymbol{d})\mathrm{I}_{(r)}\mathrm{I}_{(s)}\Big]+\mathrm{I}_{(r)}\mathrm{I}_{(s)}\mathcal{R}_2(\boldsymbol{d})+\mathrm{I}_{(r)}\mathcal{R}_2(\boldsymbol{d})\mathrm{I}_{(s)}+$$
$$+\mathrm{I}_{(s)}\mathcal{R}_2(\boldsymbol{d})\mathrm{I}_{(r)}+\mathcal{R}_2(\boldsymbol{d})\mathrm{I}_{(s)}\mathrm{I}_{(r)}+o_{\mathbb{P}}(1).\tag{A.23}$$

The order of the remainder term in (A.22) is obtained as follows. Rewrite the first term on the RHS of (A.21) as

$$\frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\Big[\varphi(\lambda_j)\mathrm{I}_{(r)}\Big(\mathrm{Re}\Big[\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}{}'\Big]+\mathrm{i}\,\mathrm{Im}\Big[\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}{}'\Big]\Big)\Big]=$$
$$=\frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\bigg[\log(\lambda_j)\mathrm{I}_{(r)}\mathrm{Re}\Big[\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}{}'\Big]+\mathrm{i}\log(\lambda_j)\mathrm{I}_{(r)}\mathrm{Im}\Big[\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}{}'\Big]+$$
$$+\mathrm{i}\Big(\frac{\lambda_j-\pi}{2}\Big)\mathrm{I}_{(r)}\mathrm{Re}\Big[\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}{}'\Big]-\Big(\frac{\lambda_j-\pi}{2}\Big)\mathrm{I}_{(r)}\mathrm{Im}\Big[\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}{}'\Big]\bigg]$$
$$=\frac{1}{m}\sum_{j=1}^{m}\bigg[\log(\lambda_j)\mathrm{I}_{(r)}\mathrm{Re}\Big[\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}{}'\Big]-\Big(\frac{\lambda_j-\pi}{2}\Big)\mathrm{I}_{(r)}\mathrm{Im}\Big[\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}{}'\Big]\bigg].$$

By summation by parts, Lemma 4.1 and assumption **B**4

$$\frac{1}{m}\sum_{j=1}^{m}\lambda_j\Big[\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}{}'\Big]\leq\frac{1}{m}\sum_{j=1}^{m-1}\big|\lambda_j-\lambda_{j+1}\big|\bigg\|\sum_{k=1}^{j}\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}{}'\bigg\|_\infty+$$
$$+\frac{\lambda_m}{m}\bigg\|\sum_{k=1}^{j}\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}{}'\bigg\|_\infty$$
$$\leq\frac{1}{m}\frac{m-1}{n}\bigg[O_{\mathbb{P}}\Big((m-1)^{1/2}\log(m-1)+\frac{(m-1)^{\alpha+1}}{n^\alpha}\Big)+O\Big(\frac{1}{m}\Big)\bigg]+$$
$$+O\Big(\frac{1}{n}\Big)\bigg[O_{\mathbb{P}}\Big(m^{1/2}\log(m)+\frac{m^{\alpha+1}}{n^\alpha}\Big)+O\Big(\frac{1}{m}\Big)\bigg]$$
$$=O_{\mathbb{P}}\Big(\frac{m^{1/2}\log(m)}{n}+\frac{m^\alpha}{n^\alpha}\Big)+O\Big(\frac{1}{mn}\Big)\;=\;o_{\mathbb{P}}\Big(\frac{1}{\log(n)}\Big),$$

where the last equality follows from

$$\frac{m^{1/2}\log(m)}{n}\log(n)=\frac{m^{1/2}}{n^{1/2}}\frac{\log(m)}{n^{1/4}}\frac{\log(n)}{n^{1/4}}\longrightarrow 0,$$

as n tends to infinity, by assumption **B**4. The other term in (A.21) is dealt analogously. The remainder term in (A.23) involves

$$\frac{1}{m}\sum_{j=1}^{m}\rho(\lambda_j)\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}\,',$$

for $\rho(\lambda_j)$ proportional to λ_j, λ_j^2 and $\lambda_j\log(\lambda_j)$. The order of the terms proportional to λ_j has already been obtained, while the one proportional to λ_j^2 is dealt analogously since $\lambda_j^2 = O(\lambda_j)$. The term proportional to $\lambda_j\log(\lambda_j)$ is $o_{\mathbb{P}}(1)$ since, by summation by parts, Lemma 4.1 and assumption **B**4,

$$\begin{aligned}\bigg\|\frac{1}{m}\sum_{j=1}^{m}\lambda_j\log(\lambda_j)\Big[\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}\,'\Big]\bigg\|_\infty &\leq\\ \leq \frac{1}{m}\sum_{j=1}^{m-1}\Big|\lambda_j\log(\lambda_j)-\lambda_{j+1}\log(\lambda_{j+1})\Big|&\bigg\|\sum_{k=1}^{j}\Lambda_k(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_k(\boldsymbol{d})^{-1}}\,'\bigg\|_\infty+\\ +\frac{\lambda_m\log(\lambda_m)}{m}&\bigg\|\sum_{j=1}^{m}\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}\,'\bigg\|_\infty\\ =\frac{1}{m}\,O_{\mathbb{P}}\bigg((m-1)^{1/2}\log(m-1)+\frac{(m-1)^{\alpha+1}}{n^\alpha}\bigg)&\,o(1)+o_{\mathbb{P}}(1) \;=\; o_{\mathbb{P}}(1).\end{aligned}$$

In order to finish the proof, it suffices to show that,

$$\mathcal{R}_k(\boldsymbol{d}) = G_0\,\frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j)^k + o_{\mathbb{P}}\big(\log(n)^{k-2}\big) \tag{A.24}$$

and

$$\mathcal{I}_k(\boldsymbol{d}) = o_{\mathbb{P}}\big(\log(n)^{k-2}\big), \tag{A.25}$$

uniformly in $\boldsymbol{d}\in\mathcal{M}$. Indeed, if (A.24) and (A.25) hold, upon defining for a matrix M,

$$T_1(M,r) := \mathrm{I}_{(r)}M + M\mathrm{I}_{(r)}, \qquad T_2(M,r,s) := \mathrm{I}_{(r)}\mathrm{I}_{(s)}M + \mathrm{I}_{(r)}M\mathrm{I}_{(s)} + \mathrm{I}_{(s)}M\mathrm{I}_{(r)} + M\mathrm{I}_{(r)}\mathrm{I}_{(s)}$$

and

$$T_3(M,r,s) := -\mathrm{I}_{(r)}\mathrm{I}_{(s)}M + \mathrm{I}_{(r)}M\mathrm{I}_{(s)} + \mathrm{I}_{(s)}M\mathrm{I}_{(r)} - M\mathrm{I}_{(r)}\mathrm{I}_{(s)},$$

it follows that (cf. Shimotsu, 2007, p.301)

$$\widehat{G}(\overline{\mathbf{d}})^{-1} = G_0^{-1} + o_{\mathbb{P}}\big(\log(n)^{-2}\big), \qquad \frac{\partial\widehat{G}(\overline{\mathbf{d}})}{\partial d_r} = \frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j)T_1(G_0,r) + o_{\mathbb{P}}\big(\log(n)^{-1}\big),$$

$$\text{and}\quad \frac{\partial^2\widehat{G}(\overline{\mathbf{d}})}{\partial d_r\partial d_s} = \frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j)^2T_2(G_0,r,s) + \frac{\pi^2}{4}\,T_3(G_0,r,s) + o_{\mathbb{P}}(1).$$

Since $\mathrm{tr}\Big[G_0^{-1}T_1(G_0,r)G_0^{-1}T_1(G_0,s)\Big] = \mathrm{tr}\big[G_0^{-1}T_2(G_0,r,s)\big]$ and

$$\frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j)^2 - \bigg(\frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j)\bigg)^2 \longrightarrow 1,$$

we obtain

$$\frac{\partial^2 S(\boldsymbol{d})}{\partial d_r\partial d_s} = \mathrm{tr}\bigg[G_0^{-1}T_2(G_0,r,s) + \frac{\pi^2}{4}\,G_0^{-1}T_3(G_0,r,s)\bigg] + o_{\mathbb{P}}(1),$$

from where (A.16) follows. We proceed to show (A.24) and (A.25). For $k=0,1,2$, let

$$\mathcal{F}_k(\boldsymbol{\theta}) := \frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j)^k\Lambda_j(\boldsymbol{\theta})^{-1}G_0\overline{\Lambda_j(\boldsymbol{\theta})^{-1}}\,'.$$

Then, (A.24) and (A.25) follow if

$$\sup_{\boldsymbol{d}\in\mathcal{M}}\left\{\left\|\frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j)^k\Lambda_j(\boldsymbol{d})^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d})^{-1}}\,'-\mathcal{F}_k(\boldsymbol{\theta})\right\|_\infty\right\}=o_{\mathbb{P}}\big(\log(n)^{k-2}\big),\tag{A.26}$$

and

$$\sup_{\boldsymbol{d}\in\mathcal{M}}\left\{\left\|\mathcal{F}_k(\boldsymbol{\theta})-G_0\frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j)^k\right\|_\infty\right\}=o\big(\log(n)^{k-2}\big).\tag{A.27}$$

Following Shimotsu (2007), p.302, notice that, by applying (A.5), we can rewrite (A.26) as

$$\sup_{\boldsymbol{d}\in\mathcal{M}}\left\{\left\|\frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j)^k\Lambda_j(\boldsymbol{\theta})^{-1}\Big[\Lambda_j(\boldsymbol{d}_0)^{-1}f_n(\lambda_j)\overline{\Lambda_j(\boldsymbol{d}_0)^{-1}}\,'-G_0\Big]\overline{\Lambda_j(\boldsymbol{\theta})^{-1}}\,'\right\|_\infty\right\}.$$

Define $b_j(\boldsymbol{\theta};k):=\log(\lambda_j)^k\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(\theta_r-\theta_s)/2}\lambda_j^{\theta_r+\theta_s}$, for $k=0,1,2$. Then, by omitting the supremum, the (r,s)-th element of (A.26) is equal to

$$\begin{aligned}\left|\frac{1}{m}\sum_{j=1}^{m}b_j(\boldsymbol{\theta};k)\Big[\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(d_r^0-d_s^0)/2}\lambda_j^{d_r^0+d_s^0}f_n^{rs}(\lambda_j)-G_0^{rs}\Big]\right|&\le\\ \le\frac{1}{m}\sum_{j=1}^{m-1}\Big|b_j(\boldsymbol{\theta};k)-b_{j+1}(\boldsymbol{\theta};k)\Big|&\left|\sum_{l=1}^{j}\mathrm{e}^{\mathrm{i}(\lambda_l-\pi)(d_r^0-d_s^0)/2}\lambda_l^{d_r^0+d_s^0}f_n^{rs}(\lambda_l)-G_0^{rs}\right|\\ +\frac{b_m(\boldsymbol{\theta};k)}{m}&\left|\sum_{j=1}^{m}\mathrm{e}^{\mathrm{i}(\lambda_j-\pi)(d_r^0-d_s^0)/2}\lambda_j^{d_r^0+d_s^0}f_n^{rs}(\lambda_j)-G_0^{rs}\right|,\end{aligned}\tag{A.28}$$

where the inequality follows from summation by parts. Now, since

$$b_j(\boldsymbol{\theta};k)-b_{j+1}(\boldsymbol{\theta};k)=O\left(\frac{\log(n)^k}{j}\right)\qquad\text{and}\qquad b_m(\boldsymbol{\theta};k)=O\big(\log(n)^k\big),$$

uniformly in $\boldsymbol{\theta}\in\mathcal{M}$, for any $k=0,1,2$, it follows by Lemma 4.1 and Remark 4.1 that (A.28) can be rewritten as

$$\begin{aligned}&O\left(\frac{\log(n)^k}{m}\right)\frac{1}{m}O_{\mathbb{P}}\left(\frac{m^{\alpha+1}}{n^\alpha}+m^{1/2}\log(m)\right)+\\&\qquad+O\big(\log(n)^k\big)\frac{1}{m}O_{\mathbb{P}}\left(\frac{m^{\alpha+1}}{n^\alpha}+m^{1/2}\log(m)\right)=o_{\mathbb{P}}\big(\log(n)^{k-2}\big),\end{aligned}$$

where the last equality follows from assumption **B**4 (see also Remark 4.1), because

$$\log(n)^2\frac{1}{m}O_{\mathbb{P}}\left(\frac{m^{\alpha+1}}{n^\alpha}+m^{1/2}\log(m)\right)=\left[\frac{\log(n)^2}{m^{1/2}\log(m)}+\frac{\log(m)}{m^{1/4}}\frac{\log(n)^2}{m^{1/4}}\right]O_{\mathbb{P}}(1)=o_{\mathbb{P}}(1).$$

The other term is dealt analogously, so that (A.26) follows. As for (A.27), the result follows from the proof of theorem 2, p.302, in Shimotsu (2007) (notice that it does not depend on f_n). This completes the proof. ∎

Proof of Corollary 5.1

The proof follows the same lines as the proof of lemma 1(a) in Shimotsu (2007) p.308 in view of $I_T(\lambda;n)=O\big(I_n(\lambda)\big)$. ∎

Proof of Corollary 5.2

By carefully inspecting the proof of Theorem 4.1, we notice that suffices to show that part (a) and (b) of Lemma 4.3 hold for the result of Theorem 4.1 to hold. Part (a) and (b) of Lemma 4.3 are proven following the same lines as the proof of lemma 1(b1) and lemma 1(b2) in Shimotsu (2007), in view of $I_T(\lambda; n) = O\big(I_n(\lambda)\big)$. ■

References

[1] Chiriac, R. and Voev, V. (2011). "Modeling and Forecasting Multivariate Realized Volatility". *Journal of Applied Econometrics*, **26**(6), 922-947.

[2] Dahlhaus, R. (1983). "Spectral Analysis with Tapered Data". *Journal of Time Series Analysis*, **4**, 163-175.

[3] Fox, R. and Taqqu, M.S. (1986). "Large-Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series". *The Annals of Statistics*, **14**(2), 517-532.

[4] Fryzlewicz, P., Nason, G.P. and von Sachs, R. (2008). "A Wavelet-Fisz Approach to Spectrum Estimation". *Journal of Time Series Analysis*, **29**(5), 868-880.

[5] Giraitis, L. and Surgailis, D. (1990)."A Central Limit Theorem for Quadratic Forms in Strongly Dependent Linear Variables and its Application to Asymptotic Normality of Whittle's Estimate". *Probability Theory and Related Fields*, **86**, 87-104.

[6] Grenander, U. (1951). "On Empirical Spectral Analysis of Stochastic Process". *Arkiv för Matematik*, **1**, 197-277.

[7] Hurvich, C.M. and Beltrão, K.I. (1993). "Asymptotics for the Low-Frequency Ordinates of the Periodogram of a Long-Memory Time Series". *Journal of Time Series Analysis*, **14**(5), 455-472.

[8] Hurvich, C.M. and B.K. Ray (1995). "Estimation of the Memory Parameter for Nonstationary or Noninvertible Fractionally Integrated Processes". *Journal of Time Series Analysis*, **16**(1), 17-42.

[9] Künsch, H. (1987). "Statistical Aspects of Self-Similar Processes". In Prokhorov, Yu. and Sazanov, V.V. (Eds.), *Proceedings of the First World Congress of the Bernoulli Society*. Utrecht: VNU Science Press, 67-74.

[10] Olbermann, B.P., Lopes, S.R.C. and Reisen, V.A. (2006). "Invariance of the First Difference in ARFIMA Models". *Computational Statistics*, **21**(3), 445-461.

[11] Lobato, I.N. (1999). "A Semiparametric Two-Step Estimator in a Multivariate Long Memory Model". *Journal of Econometrics*, **90**, 129-153.

[12] Luceño, A. (1996). "A Fast Likelihood Approximation for Vector General Linear Processes with Long Series: Application to Fractional Differencing". *Biometrika*, **83**(3), 603-614.

[13] Nielsen, F.S. (2011). "Local Whittle Estimation of Multi-Variate Fractionally Integrated Processes". *Journal of Time Series Analysis*, **32**(3), 317-335.

[14] Priestley, M.B. (1981). *Spectral Analysis and Time Series*. London: Academic Press.

[15] Pumi, G. and Lopes, S.R.C (2013). "A Semiparametric Estimator for Long-Range Dependent Multivariate Processes". Under review.

[16] Robinson, P.M. (1995a). "Log-Periodogram Regression of Time Series with Long Range Dependence". *Annals of Statistics*, **23**(3), 1048-1072.

[17] Robinson, P.M. (1995b). "Gaussian Semiparametric Estimation of Long Range Dependence". *Annals of Statistics*, **23**(5), 1630-1661.

[18] Shimotsu, K. (2007). "Gaussian Semiparametric Estimation of Multivariate Fractionally Integrated Processes". *Journal of Econometrics*, **137**, 277-310.

[19] Shimotsu, K. and Phillips, P.C.B. (2005). "Exact Local Whittle Estimation of Fractional Integration". *Annals of Statistics*, **33**(4), 1890-1933.

[20] Sowell, F. (1989). "Maximum Likelihood Estimation of Fractionally Integrated Time Series Models". Working Paper, Carnegie-Mellon University.

[21] Tsay, W-J. (2010). "Maximum Likelihood Estimation of Stationary Multivariate ARFIMA Processes". *Journal of Statistical Computation and Simulation*, **80**(7-8), 729-745.

[22] Velasco, C. (1999). "Gaussian Semiparametric Estimation of Non-Stationary Time Series". *Journal of Time Series Analysis*, **20**, 87-127.

[23] Zygmund, A. (2002). *Trigonometric Series. Vol. I, II.* Cambridge: Cambridge University Press.