Gaussian semiparametric estimation of multivariate fractionally integrated processes

Katsumi Shimotsu*
Department of Economics
University of Essex

November 25, 2003

Abstract

This paper analyzes the semiparametric estimation of multivariate long-range dependent processes. The class of spectral densities considered includes multivariate fractionally integrated processes, which are not covered by the existing literature. This paper also establishes the consistency of the multivariate Gaussian semiparametric estimator, which has not been shown in the other works. Asymptotic normality of the multivariate Gaussian semiparametric estimator is also established, and the proposed estimator is shown to have a smaller limiting variance than the two-step Gaussian semiparametric estimator studied by Lobato (1999). Gaussianity is not assumed in the asymptotic theory.

JEL Classification: C22

1 Introduction

Consider a real-valued covariance stationary q-vector process X_t that is generated by

$$\begin{pmatrix} (1-L)^{d_1} & & 0 \\ & \ddots & \\ 0 & & (1-L)^{d_q} \end{pmatrix} \begin{pmatrix} X_{1t} - EX_{1t} \\ \vdots \\ X_{qt} - EX_{qt} \end{pmatrix} = \begin{pmatrix} u_{1t} \\ \vdots \\ u_{qt} \end{pmatrix}, \quad -\frac{1}{2} < d_1, \dots, d_q < \frac{1}{2}, \tag{1}$$

where $u_t = (u_{1t}, \dots u_{qt})'$ is a covariance stationary process whose diagonal elements are bounded and bounded away from zero at the origin. This is a multivariate extension of a scalar fractionally integrated process, or the so-called $I(d)$ process, and X_{at} exhibits the long-range dependence when $d_a \neq 0$. The long-range dependent processes are used extensively in economics and finance, in particular in modeling certain financial data, such as volatility and trading volume. X_t becomes a multivariate ARFIMA process when u_t is a vector ARMA process, but the specification (1) does not require u_t to be so.

*Currently at Department of Economics, Queen's University. The author thanks the ESRC for research support under Grant R000223986. Simulations were performed in MATLAB.

Fractionally integrated processes are the most widely used long-range dependent time series in economics and econometrics. They have a time domain representation that extends the conventional ARMA models in a natural way. The relationship between the value of the memory parameter and the persistence of a shock are easily understood as the value of the coefficient in the expansion of

$$(1-L)^{-d} = \sum_{k=0}^{\infty} \frac{(d)_k}{k!} L^k, \quad (d)_k = \frac{\Gamma(d+k)}{\Gamma(d)}.$$

Recent applications of fractional integration are found in, e.g., Bollerslev and Wright (2000) and Brunetti and Gilbert (2000). Henry and Zaffaroni (2003) provide a survey of applications of fractional integration and long-range dependence in macroeconomics and finance.

Let $f(\lambda)$ and $f_u(\lambda)$ denote the spectral density of X_t and u_t, respectively, such that

$$E(X_t - EX_t)(X'_{t+k} - EX'_t) = \int_{-\pi}^{\pi} e^{ij\lambda} f(\lambda) d\lambda,$$

and similarly for $f_u(\lambda)$. Let

$$\Phi(\lambda) = \text{diag}\left((1-e^{i\lambda})^{-d_1}, \cdots, (1-e^{i\lambda})^{-d_q}\right) = \text{diag}\left((1-e^{i\lambda})^{-d_a}\right),$$

then the spectral density of X_t is (e.g., Hannan, 1970, p.61)

$$f(\lambda) = \Phi(\lambda) f_u(\lambda) \Phi^*(\lambda).$$

As we shall see shortly, the memory parameter, d, governs the long-run dynamics of the process and the behavior of $f(\lambda)$ around the origin. Therefore, if the interest lies in the long-run dynamics of the process, it is useful to specify the spectral density only locally in the vicinity of the origin and avoid specifying the short-run dynamics of u_t explicitly. Assume $f_u(\lambda)$ satisfies

$$f_u(\lambda) \sim G, \quad \lambda \to 0,$$

where G is real, symmetric, finite, and positive definite. Since

$$(1-e^{i\lambda})^{\alpha} = \lambda^{\alpha} e^{-i\pi\alpha/2}(1+O(\lambda)), \quad \lambda \to 0, \tag{2}$$

(Phillips and Shimotsu, 2003), it follows that

$$f(\lambda) \sim \text{diag}(\lambda^{-d_a} e^{i\pi d_a/2}) G \, \text{diag}(\lambda^{-d_a} e^{-i\pi d_a/2}), \quad \lambda \to 0, \tag{3}$$

and the behavior of $f(\lambda)$ around the origin is governed only by d and G.

When $f(\lambda)$ is specified locally as (3), we can estimate d semiparametrically using the information only on the long-run dynamics of the process. Semiparametric estimation uses the periodograms evaluated at a band that shrinks toward the origin as the sample size tends to infinity. The semiparametric estimators are robust to misspecification of short-run dynamics, because they are agnostic to the behavior of the periodograms away from the origin.

In a univariate case where $f(\lambda) \sim G\lambda^{-2d}$ as $\lambda \to 0$, one attractive semiparametric estimator was proposed by Künsch (1987) and analyzed by Robinson (1995b). The estimator, Gaussian semiparametric estimator (GSE), is based on the maximization of the frequency domain Gaussian likelihood function that is localized to the vicinity of the origin. The GSE has several advantages over the other semiparametric estimators, including efficiency and a weaker distributional assumption. Lobato (1999) analyzed a version of multivariate extension of GSE. It considers a two-step estimation of d, which is based on the first-step univariate estimate of $d_1, \ldots, d_q$ and a Newton-type second step, and shows the asymptotic normality of the two-step estimator.

We consider semiparametric estimation of d when the spectral density has the form (3). The specification (3) extends the specification $f(\lambda) \sim G\lambda^{-2d}$ into the multivariate case. It includes multivariate fractionally integrated processes and is also general enough to accommodate the presence of poles and zeros at frequencies away from the origin. In (3), the memory parameter d appears in λ^{-d_a} and $e^{i\pi d_a/2}$, and hence the estimation of d needs to take both λ^{-d_a} and $e^{i\pi d_a/2}$ into account. This dependency was thought to make the analysis difficult. Consequently, Lobato (1999) considered semiparametric estimation of d from an alternate form of spectral density[1]

$$\widetilde{f}(\lambda) \sim \text{diag}(\lambda^{-d_a}) G \text{diag}(\lambda^{-d_a}), \quad \lambda \to 0. \tag{4}$$

When X_t is generated by a multivariate fractionally integrated process (1), however, it is not clear if an estimator based on the specification (4) provides a valid estimate of d. This is because the off-diagonal elements of $\text{diag}(\lambda^{d_a}) f(\lambda) \text{diag}(\lambda^{d_a})$ have a nonnegligible imaginary part even in the neighborhood of the origin, and $f(\lambda)$ does not belong to the class of spectral densities specified in (4). Indeed, we are not aware of a time domain model of multivariate time series whose spectral density follows (4).

We also prove the consistency of our multivariate GSE. Two-step estimation is partly motivated by its computational ease. However, in view of today's computational power, a maximization of the objective function with respect to q parameters is not likely to cause any practical difficulty. Indeed, the simulation in this study confirms it. Direct maximization of the objective function also dispenses with the numerical differentiation that is necessary for the evaluation of the score function and Hessian. Although the proof of the consistency of univariate GSE by Robinson (1995b) is not directly applicable to the multivariate case, a proper modification of the proof by Robinson (1995b) enables us to handle the nonuniform convergence of the objective function and establish the consistency of the multivariate GSE.

The GSE is shown to have a Gaussian limiting distribution. Intriguingly, its limiting variance is different from that of the GSE analyzed by Lobato (1999), and the GSE based on (3) has a smaller limiting variance than the one based on (4). This gain of efficiency arises because it takes both real and imaginary parts of the spectral density and periodograms into account, and the presence of d in $e^{i\pi d_a/2}$ provides more information about the value of d. In simulations with multivariate fractionally integrated processes, the GSE based on (3) exhibits smaller variance than the GSE based on (4).

The remainder of the paper is organized as follows. Section 2 describes the GSE. Consistency of the GSE is demonstrated in Section 3. Section 4 derives the limit distribution. Section 5 reports some simulation results. Proofs are given in Appendix A in Section 6. Some technical results are collected in Appendix B in Section 7.

[1] The specification (4) is also used in Lobato and Robinson (1998) to construct a nonparametric test for weak dependence. Lobato and Velasco (2000) extend it to analyze the two-step Gaussian semiparametric estimation of multivariate nonstationary long-range dependent processes.

2 Multivariate semiparametric estimation

We consider semiparametric estimation of $d = (d_1, \cdots, d_q)'$, which uses only Fourier frequencies in the neighborhood of the origin and hence is nonparametric with respect to short-run dynamics of the data. Define the discrete Fourier transform (dft) and the periodogram of X_t evaluated at frequency λ as

$$w(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} X_t e^{it\lambda}, \quad I(\lambda) = w(\lambda) w^*(\lambda),$$

where x^* denotes the conjugate transpose of x. For the reason explained in Section 3, it is useful to consider an approximation finer than (3). Since $|1 - e^{i\lambda}| = |2\sin(\lambda/2)|$ and $\arg(1 - e^{i\lambda}) = (\lambda - \pi)/2$ for $0 \le \lambda < \pi$, we have

$$\begin{aligned} (1 - e^{i\lambda})^{\theta} &= (|2\sin(\lambda/2)|)^{\theta} \exp[i(\lambda - \pi)\theta/2] \\ &= \lambda^{\theta} \exp[i(\lambda - \pi)\theta i/2](1 + O(\lambda^2)). \end{aligned}$$

This is merely a refinement of (2), but the smaller error magnitude ($O(\lambda^2)$) will become essential in the analysis in Section 4. Since $f_u(\lambda) \sim G$ as $\lambda \to 0$, we have, for the Fourier frequencies $\lambda_j = 2\pi j/n$ with $j = 1, \ldots, m$ and $m = o(n)$,

$$f(\lambda_j) \sim \Lambda_j(d) G \Lambda_j^*(d), \quad \Lambda_j(d) = \mathrm{diag}(\lambda_j^{-d_a} e^{i(\pi - \lambda_j)d_a/2}).$$

Therefore, the Gaussian log-likelihood function localized to the origin is

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

Using the fact that $\det AB = \det A \det B$ for any complex matrices A and B (Lütkepohl, 1996, p. 48), the first order condition with respect to G is (Lütkepohl, 1996, p. 179)

$$\frac{\partial Q_m(G, d)}{\partial G} = \frac{1}{m} \sum_{j=1}^{m} \left\{ (G')^{-1} - \left[G^{-1} \Lambda_j(d)^{-1} I(\lambda_j) \Lambda_j^*(d)^{-1} G^{-1} \right]' \right\} = 0.$$

Taking its transpose gives $G = m^{-1} \sum_{j=1}^{m} \{\Lambda_j(d)^{-1} I(\lambda_j) \Lambda_j^*(d)^{-1}\}$. Substituting this into $Q_m(G, d)$ in conjunction with the fact that

$$\begin{aligned} &\log\det \Lambda_j(d) + \log\det \Lambda_j^*(d) \\ = \; &\log\det \Lambda_j(d) \Lambda_j^*(d) = \log(\mathrm{diag}(\lambda_j^{-2d_a})) = -2 \sum_{a=1}^{q} d_a \log \lambda_j, \end{aligned}$$

and G is real, we obtain the objective function

$$\begin{aligned} R(d) &= \log\det \widehat{G}(d) - 2 \sum_{a=1}^{q} d_a \frac{1}{m} \sum_{j=1}^{m} \log \lambda_j, \\ \widehat{G}(d) &= \frac{1}{m} \sum_{j=1}^{m} \mathrm{Re}\left[\Lambda_j(d)^{-1} I(\lambda_j) \Lambda_j^*(d)^{-1} \right]. \end{aligned}$$

In the following, we denote the true parameter values by G^0 and d_0. The estimator is defined as

$$\widehat{d} = \arg\min_{d \in \Theta} R(d),$$

where the space of admissible estimates of d^0, Θ, takes the form $\Theta = [\Delta_1, \Delta_2]^q$, with $-1/2 < \Delta_1 < \Delta_2 < 1/2$.

3 Consistency of the estimator

We now introduce the assumptions on m and $f(\lambda)$ necessary for the consistency of the estimator. Let $f_{ab}(\lambda)$ and G^0_{ab} denote the (a,b) th element of $f(\lambda)$ and G^0, respectively.

Assumption 1 *As* $\lambda \to 0+$,

$$f_{ab}(\lambda) = e^{i\pi(d_a-d_b)/2} G^0_{ab} \lambda^{-d^0_a-d^0_b} + o(\lambda^{-d^0_a-d^0_b}), \quad a,b = 1,\dots,q.$$

Assumption 2

$$X_t - EX_t = A(L)\varepsilon_t = \sum_{j=0}^{\infty} A_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} ||A_j||^2 < \infty,$$

where $||\cdot||$ *denotes the supremum norm and* $E(\varepsilon_t|F_{t-1}) = 0$, $E(\varepsilon_t\varepsilon_t'|F_{t-1}) = I_q$ *a.s.*, $t = 0, \pm1, \dots,$ *in which* F_t *is the* σ*-field generated by* ε_s, $s \le t$, *and there exists a scaler random variable* ε *such that* $E\varepsilon^2 < \infty$ *and for all* $\eta > 0$ *and some* $K > 0$, $\Pr(||\varepsilon_t|| > \eta\ ) \le K \Pr(\varepsilon^2 > \eta)$.

Assumption 3 *In a neighborhood* $(0,\delta)$ *of the origin,* $A(\lambda) = \sum_{j=0}^{\infty} A_j e^{ij\lambda}$ *is differentiable and*

$$\frac{\partial}{\partial\lambda} A_a(\lambda) = O(\lambda^{-d^0_a-1}) \quad \text{as } \lambda \to 0+,$$

where $A_a(\lambda)$ *is the ath row of* $A(\lambda)$.

Assumption 4 *As* $n \to \infty$,

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

Assumptions 1-4 are multivariate extensions of Assumptions A1-A4 of Robinson (1995b) and analogous to the ones used in Robinson (1995a) and Lobato (1999). In Assumption 1, replacing $e^{i\pi(d_a-d_b)/2}$ with $e^{i(\pi-\lambda)(d_a-d_b)/2}$ does not make difference because $e^{i\lambda} - 1 = o(1)$. Assumption 3 implies Assumption A3$'$ of Lobato (1999), i.e., $\partial A_a(\lambda)/\partial\lambda = O(\lambda^{-1}||A_a(\lambda)||)$, because $||A_a(\lambda)|| \ge (A_a(\lambda)A^*_a(\lambda)/q)^{1/2} = (2\pi f_{aa}(\lambda)/q)^{1/2}$.

Under these conditions, we may now establish the consistency of $\widehat{d}$.

3.1 Theorem

Let Assumptions 1-4 hold. Then, for $d_0 \in \Theta$, $\widehat{d} \to_p d_0$ *as* $n \to \infty$.

4 Asymptotic normality of the estimator

We introduce some further assumptions that are used in the results of this section. They are analogous to the assumptions in Lobato (1999).

Assumption 1$'$ *For* $\beta \in (0,2]$ *and* $a,b = 1,\dots,q$,

$$f_{ab}(\lambda) - e^{i(\pi-\lambda)(d^0_a-d^0_b)/2} \lambda^{-d^0_a-d^0_b} G^0_{ab} = O(\lambda^{-d^0_a-d^0_b+\beta}) \quad \text{as } \lambda \to 0+.$$

Assumption 2′ *Assumption 2 holds and also for* $a, b, c, d = 1, 2$,

$$E(\varepsilon_{at}\varepsilon_{bt}\varepsilon_{ct}|F_{t-1}) = \mu_{abc} \quad a.s., \quad E(\varepsilon_{at}\varepsilon_{bt}\varepsilon_{ct}\varepsilon_{dt}|F_{t-1}) = \mu_{abcd}, \quad t = 0, \pm 1, \ldots,$$

where $|\mu_{abc}| < \infty$ *and* $|\mu_{abcd}| < \infty$.

Assumption 3′ *Assumption 3 holds.*

Assumption 4′ As $n \to \infty$,

$$\frac{1}{m} + \frac{m^{1+2\beta}(\log m)^2}{n^{2\beta}} + \frac{\log n}{m^{\gamma}} \to 0, \quad \text{for any } \gamma > 0.$$

Assumption 5′ *There exists a finite real matrix* H *and* $\alpha > 0$ *such that*

$$\Lambda_j(d^0)^{-1}A(\lambda_j) = H + O(\lambda_j^{\alpha}).$$

Assumption 1′ does not hold for $\beta > 1$ if we replace $e^{i(\pi-\lambda)(d_a^0-d_b^0)/2}$ with $e^{i\pi(d_a^0-d_b^0)/2}$, because $e^{i\lambda} = 1 + O(\lambda)$. Assumption 4′ is slightly stronger than the assumptions in Robinson (1995b) and Lobato (1999), i.e., $m^{-1} + m^{1+2\beta}n^{-2\beta}(\log m)^2 \to 0$. It is satisfied if $m \sim Cn^{\xi}$ with a finite positive constant C and $0 < \xi < 2\beta/(1+2\beta)$. The third term on the left hand side of Assumption 4′ is necessary in establishing the convergence of the Hessian. Assumption 5′ complements Assumption 1′ in that it controls the degree of approximation of the transfer function by $\Lambda_j(d^0)$. This assumption obviously implies $HH' = 2\pi G^0$ and is satisfied by multivariate ARFIMA models.

4.1 Theorem

Let Assumptions 1′-4′ hold. Then, for $d^0 \in \text{Int}(\Theta)$, *as* $n \to \infty$,

$$m^{1/2}\left(\widehat{d} - d_0\right) \to {}_dN\left(0, \Omega^{-1}\right), \quad \Omega = 2\left[I_q + G^0 \odot (G^0)^{-1} + \frac{\pi^2}{4}\left(G^0 \odot (G^0)^{-1} - I_q\right)\right],$$

$$\widehat{G}(\widehat{d}) \to {}_pG^0,$$

where $\odot$ *denotes the Hadamard product.*

Lobato (1999) analyzes the two-step GSE that uses the objective function based on (4):

$$\widetilde{R}(d) = \log\det\widetilde{G}(d) - 2\sum_{a=1}^{q} d_a \frac{1}{m}\sum_{j=1}^{m}\log\lambda_j, \quad \widetilde{G}(d) = \frac{1}{m}\sum_{j=1}^{m}\text{Re}\left[\text{diag}(\lambda_j^{d_a})I(\lambda_j)\text{diag}(\lambda_j^{d_a})\right],$$

and show that the limiting variance of that estimator is $\Xi = 2[I_q + G^0 \odot (G^0)^{-1}]$. Because $G^0 \odot (G^0)^{-1} - I_q$ is positive semidefinite (Horn and Johnson, 1985, p.475), $\widehat{d}$ has a smaller (in a matrix sense) limiting variance than the two-step estimator analyzed by Lobato (1999), if $G^0 \neq cI_q$ for a positive scalar c. The properties the GSE based on the objective function $\widetilde{R}(d)$ remains unclear when the data are generated by (1). We conjecture it is still consistent, but the limiting variance may depend on d because $[\text{diag}(\lambda^{d_a})f(\lambda)\text{diag}(\lambda^{d_a})]$ depends on d as $\lambda \to 0$.

We compare the diagonal elements of Ω^{-1} and Ξ^{-1} with the asymptotic variance of the univariate GSE ($= 1/4$) when $q = 2$. Note that $(\Omega)^{-1}_{11} = (\Omega)^{-1}_{22}$ and $(\Xi)^{-1}_{11} = (\Xi)^{-1}_{22}$. G^0 is chosen to be

$$G^0 = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \quad \rho = 0.0, 0.2, 0.4, 0.6, 0.8.$$

Table 1 reports $(\Omega)^{-1}_{11}$ and $(\Xi)^{-1}_{11}$ and their ratio to 1/4.

Table 1. Comparison of asymptotic variance

ρ	0.0	0.2	0.4	0.6	0.8
univariate	0.250	0.250	0.250	0.250	0.250
$(\Omega^{-1})_{11}$	0.250	0.234	0.200	0.167	0.142
$(\Xi^{-1})_{11}$	0.250	0.245	0.230	0.205	0.170
$(\Omega^{-1})_{11}/(0.25)$	1.000	0.937	0.801	0.670	0.570
$(\Xi^{-1})_{11}/(0.25)$	1.000	0.980	0.920	0.820	0.680

When $\rho \leq 0.2$, the variance of the three estimators is not substantially different. When $\rho \geq 0.4$, both $(\Omega^{-1})_{11}$ and $(\Xi^{-1})_{11}$ are noticeably smaller than 1/4. As ρ gets larger, they become still smaller, and also the difference between $(\Omega^{-1})_{11}$ and $(\Xi^{-1})_{11}$ increases. Therefore, we may expect a nonnegligible gain in efficiency from estimating the elements of d jointly, and the gain may be substantial, especially when both real and imaginary parts of the spectral density are taken into account.

5 Simulations

This section reports some simulations that were conducted to examine the finite sample performance of the analyzed GSE (hereafter GSE1). We also examine the finite sample properties of the GSE based on the objective function $\widetilde{R}(d)$ (hereafter GSE2). The sample size and band parameter m were chosen to be $n = 500$ and $m = n^{0.65} = 56$. We generate X_t by truncating the infinite order moving average representation of (1):

$$\begin{aligned} X_t &= \begin{pmatrix} (1-L)^{-d_1} & 0 \\ 0 & (1-L)^{-d_2} \end{pmatrix} \left[\begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} I\{t \geq 1\} \right], \\ \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} &\sim iidN\left(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right). \end{aligned}$$

$n + 2000$ observations of X_t were generated, and the first 2000 observations were discarded. The bias, standard deviation, and root mean squared error (RMSE) were computed using 10,000 replications. The value of d was chosen to be $(0.2, -0.2)$, $(0.2, 0.2)$, and $(0.2, 0.4)$. The results do not appear to depend on the value of d. Three values of ρ were used; $\rho = 0$, 0.5, 0.8.

Tables 2-4 show the simulation results of both estimators. Table 2 shows the results for $\rho = 0$. Both GSE1 and GSE2 have little bias for all values of d. The standard deviation and RMSE of GSE1 are slightly higher than those of GSE2. The limiting variance of the two estimators is the same, and the simulation result appears to corroborate it. The bias, standard deviation, and RMSE do not appear to be affected by the value of d. Table 3 shows the results for $\rho = 0.5$. GSE1 has smaller standard deviation and RMSE than GSE2. Again, the performance of the estimators is not substantially affected by the value of d. Table 4 shows the results for $\rho = 0.8$. Interestingly, the standard deviation and RMSE of GSE2 appear to depend on the value of d. This suggests that the limiting variance of GSE2 depends on d when the data are generated by an ARFIMA process. This is not surprising, however, because when X_t follows an ARFIMA process and $\rho \neq 0$, the matrix G in (4) depends on the value of d. Both GSE1 and GSE2 have smaller standard deviations than the case when $\rho = 0.5$. A simulation for a single pair of $(d_{1,}d_{2,}\rho)$ took around 60 minutes with a PC box with a dual 2.0 Ghz CPU running the Linux operating system.

Table 2. Simulation results: $n = 500$, $m = n^{0.65} = 56$

	GSE1			**GSE2**		
	bias	s.d.	RMSE	bias	s.d.	RMSE
			$\rho = 0.0$			
			$(d_1, d_2) = (0.2, -0.2)$			
d_1	-0.0064	0.0789	0.0792	-0.0066	0.0784	0.0787
d_2	-0.0038	0.0777	0.0778	-0.0037	0.0773	0.0774
			$(d_1, d_2) = (0.2, 0.2)$			
d_1	-0.0060	0.0781	0.0783	-0.0061	0.0776	0.0778
d_2	-0.0074	0.0781	0.0785	-0.0075	0.0777	0.0781
			$(d_1, d_2) = (0.2, 0.4)$			
d_1	-0.0062	0.0785	0.0787	-0.0063	0.0780	0.0782
d_2	-0.0020	0.0790	0.0790	-0.0021	0.0786	0.0786

Table 3. Simulation results: $n = 500$, $m = n^{0.65} = 56$

	GSE1			**GSE2**		
	bias	s.d.	RMSE	bias	s.d.	RMSE
			$\rho = 0.5$			
			$(d_1, d_2) = (0.2, -0.2)$			
d_1	-0.0043	0.0672	0.0674	-0.0037	0.0752	0.0753
d_2	-0.0007	0.0665	0.0665	-0.0001	0.0747	0.0747
			$(d_1, d_2) = (0.2, 0.2)$			
d_1	-0.0059	0.0667	0.0670	-0.0067	0.0728	0.0731
d_2	-0.0055	0.0665	0.0667	-0.0070	0.0730	0.0733
			$(d_1, d_2) = (0.2, 0.4)$			
d_1	-0.0034	0.0670	0.0671	-0.0054	0.0740	0.0742
d_2	-0.0016	0.0673	0.0673	-0.0008	0.0744	0.0744

Table 4. Simulation results: $n = 500$, $m = n^{0.65} = 56$

	GSE1			GSE2		
	bias	s.d.	RMSE	bias	s.d.	RMSE
			$\rho = 0.8$			
			$(d_1, d_2) = (0.2, -0.2)$			
d_1	0.0022	0.0597	0.0598	0.0056	0.0721	0.0723
d_2	0.0056	0.0599	0.0601	0.0093	0.0721	0.0727
			$(d_1, d_2) = (0.2, 0.2)$			
d_1	-0.0052	0.0587	0.0589	-0.0070	0.0644	0.0647
d_2	-0.0047	0.0585	0.0586	-0.0067	0.0644	0.0647
			$(d_1, d_2) = (0.2, 0.4)$			
d_1	0.0001	0.0595	0.0595	-0.0013	0.0678	0.0678
d_2	0.0004	0.0597	0.0597	0.0034	0.0682	0.0683

6 Appendix A: Proofs

In this and the following section, C denotes a generic constant such that $C \in (1, \infty)$ unless specified otherwise, and it may take different values in different places.

6.1 Proof of Theorem 3.1

Define $\theta = (\theta_1, \cdots, \theta_q)' = d - d^0$ and $S(d) = R(d) - R(d^0)$. Fix $1/2 > \delta > 0$, and define $\overline{N}_\delta = \{d : ||d - d^0|| \geq \delta\}$, where $||\cdot||$ denotes the supremum norm. For arbitrary small $\Delta > 0$, define $\Theta_1 = \{\theta : \theta \in [-1/2 + \Delta, 1/2]^q\}$ and $\Theta_2 = \Theta \backslash \Theta_1$, possibly empty. Without loss of generality, assume $\Delta < 1/4$. Then we have (c.f. Robinson, 1995b, p. 1634)

$$\begin{aligned} \Pr\left(||\widehat{d} - d^0|| > \delta\right) &\leq \Pr\left(\inf_{\overline{N}_\delta \cap \Theta} S(d) \leq 0\right) \\ &\leq \Pr\left(\inf_{\overline{N}_\delta \cap \Theta_1} S(d) \leq 0\right) + \Pr\left(\inf_{\Theta_2} S(d) \leq 0\right). \end{aligned} \tag{5}$$

For the first probability on the right of (5), rewrite $S(d)$ as

$$\begin{aligned} S(d) &= \log\det\widehat{G}(d) - \log\det\widehat{G}(d^0) - 2\sum_{a=1}^{q}\theta_a \frac{1}{m}\sum_{j=1}^{m}\log\lambda_j \\ &= \log\det\widehat{G}(d) + \log\left(\frac{2\pi m}{n}\right)^{-2(\theta_1+\cdots+\theta_q)} - \log\det\widehat{G}(d^0) \\ &\quad -2\sum_{a=1}^{q}\theta_a\left(\frac{1}{m}\sum_{j=1}^{m}\log j - \log m\right) \\ &= S_1(d) - S_1(d_0) + S_2(d), \end{aligned}$$

where

$$\begin{aligned} S_1(d) &= \log A(d) - \log B(d), \\ A(d) &= \left(\frac{2\pi m}{n}\right)^{-2(\theta_1+\cdots+\theta_q)} \det\widehat{G}(d), \quad B(d) = \prod_{a=1}^{q}(2\theta_a + 1)^{-1}\det G^0, \\ S_2(d) &= -2\sum_{a=1}^{q}\theta_a\left(\frac{1}{m}\sum_{j=1}^{m}\log j - \log m\right) - \sum_{a=1}^{q}\log(2\theta_a + 1). \end{aligned}$$

Since $m^{-1}\sum_1^m \log j - \log m + 1 = O(m^{-1}\log m)$ (see, e.g. Robinson, 1995b, Lemma 2), we have

$$S_2(d) = \sum_{a=1}^{q} [2\theta_a - \log(2\theta_a + 1)] + O(m^{-1}\log m).$$

Because $x - \log(x+1)$ achieves a unique global minimum on $(-1,\infty)$ at $x=0$ and $x - \log(x+1) \geq x^2/6$ for $0 \leq |x| < 1$, for all sufficiently large n

$$\inf_{\overline{N}_\delta \cap \Theta_1} S_2(d) \geq \delta^2/6.$$

For $S_1(d)$, if there exists nonrandom $\Xi(d)$ such that

$$\text{(i)} \quad \sup_{\Theta_1} |A(d) - \Xi(d)| = o_p(1), \quad \text{(ii)} \quad \Xi(d) \geq B(d), \quad \text{(iii)} \quad \Xi(d^0) = B(d^0), \quad (6)$$

as $n \to \infty$, then we have

$$\begin{aligned} S_1(d) &= \log(\Xi(d) + o_p(1)) - \log B(d) = \log\left(1 + \frac{\Xi(d) - B(d) + o_p(1)}{B(d)}\right), \\ S_1(d^0) &= \log\left(1 + o_p(1)/B(d^0)\right), \end{aligned}$$

uniformly in Θ_1. Then $\Pr(\inf_{\overline{N}_\delta \cap \Theta_1} S_1(d) - S_1(d^0) \leq -\delta^2/12) \to 0$ follows because $|\log(1+x)| \leq 2|x|$ for $|x| \leq 1/2$ and $\inf_{\Theta_1} B(d) > 0$. Thus $\Pr(\inf_{\overline{N}_\delta \cap \Theta_1} S(d) \leq 0) \to 0$ follows.

We proceed to show (6). For (i), recall that $\Lambda_j(d)^{-1} = \text{diag}(e^{\frac{1}{2}(\lambda_j - \pi)d_a i}\lambda_j^{d_a})$ and

$$\Lambda_j(d)^{-1} = \Lambda_j(d - d^0)^{-1}\Lambda_j(d^0)^{-1} = \Lambda_j(\theta)^{-1}\Lambda_j(d^0)^{-1}.$$

It follows that

$$\begin{aligned} A(d) &= \left(\frac{2\pi m}{n}\right)^{-2(\theta_1 + \cdots + \theta_q)} \\ &\quad \times \det\left\{\frac{1}{m}\sum_{j=1}^{m} \text{Re}\left[\Lambda_j(\theta)^{-1}\Lambda_j(d^0)^{-1}I(\lambda_j)\Lambda_j^*(d^0)^{-1}\Lambda_j^*(\theta)^{-1}\right]\right\} \\ &= \det\left\{\frac{1}{m}\sum_{j=1}^{m} \text{Re}\left[M_j(\theta)\Lambda_j(d^0)^{-1}I(\lambda_j)\Lambda_j^*(d^0)^{-1}M_j^*(\theta)\right]\right\}, \qquad (7) \end{aligned}$$

where $M_j(\theta) = \text{diag}(e^{i(\lambda_j - \pi)\theta_a/2}(j/m)^{\theta_a})$. Hereafter let I_j denote $I(\lambda_j)$ and w_{aj} denote $w_a(\lambda_j)$, the ath element of $w(\lambda_j)$. Observe that the (a,b)th element of the inside of $\det\{\cdot\}$ in (7) is

$$\frac{1}{m}\sum_{j=1}^{m} \text{Re}\left[e^{i(\lambda_j - \pi)(\theta_a - \theta_b)/2}\left(\frac{j}{m}\right)^{\theta_a + \theta_b} e^{i(\lambda_j - \pi)(d_a^0 - d_b^0)/2}\lambda_j^{d_a^0 + d_b^0} w_{aj} w_{bj}^*\right].$$

Summation by parts (Robinson, 1995b, p. 1636) and Lemma 7.1 give, uniformly in (a,b),

$$\sup_{\Theta_1}\left|\frac{1}{m}\sum_{j=1}^{m} e^{i(\lambda_j - \pi)(\theta_a - \theta_b)i}\left(\frac{j}{m}\right)^{\theta_a + \theta_b}\left(e^{i(\lambda_j - \pi)(d_a^0 - d_b^0)/2}\lambda_j^{d_a^0 + d_b^0} w_{aj} w_{bj}^* - G_{ab}^0\right)\right|$$

$$\leq \frac{1}{m}\sum_{r=1}^{m-1}\sup_{\Theta_1}\left|\left(\frac{r}{m}\right)^{\theta_a+\theta_b}e^{i(\lambda_r-\pi)(\theta_a-\theta_b)/2}-\left(\frac{r+1}{m}\right)^{\theta_a+\theta_b}e^{i(\lambda_{r+1}-\pi)(\theta_a-\theta_b)/2}\right|$$

$$\times\left|\sum_{j=1}^{r}\left(e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0}w_{aj}w_{bj}^*-G_{ab}^0\right)\right|$$

$$+\left|\frac{1}{m}\sum_{j=1}^{m}\left(e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0}w_{aj}w_{bj}^*-G_{ab}^0\right)\right|$$

$$\leq 6\sum_{r=1}^{m-1}\left(\frac{r}{m}\right)^{2\Delta}\frac{1}{r^2}\left|\sum_{j=1}^{r}\left(e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0}w_{aj}w_{bj}^*-G_{ab}^0\right)\right|$$

$$+\frac{1}{m}\left|\sum_{j=1}^{m-1}\left(e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0}w_{aj}w_{bj}^*-G_{ab}^0\right)\right|=o_p(1). \tag{8}$$

It follows that

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

uniformly in Θ_1. Define $\mathcal{E}(\theta)$ and $M_\infty(\theta)$ to be matrices whose (a,b) elements are $e^{-\pi(\theta_a-\theta_b)i/2}$ and $\int_0^1 x^{\theta_a+\theta_b}dx$, respectively. From Lemma 2 of Robinson (1995b), we have

$$\sup_{C\geq\gamma\geq\varepsilon}\left|\frac{\gamma}{m}\sum_{1}^{m}\left(\frac{j}{m}\right)^{\gamma-1}-1\right|=O\left(\frac{1}{m^{\varepsilon}}\right)\quad\text{as } m\to\infty,$$

for $\varepsilon\in(0,1]$ and $C\in(\varepsilon,\infty)$. Hence, in view of $e^{i(\lambda-\pi)(\theta_a-\theta_b)/2}=e^{-i\pi(\theta_a-\theta_b)/2}+O(\lambda)$, we have

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

and (i) of (6) follows with

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

because the determinant is a continuous function of each element and $\mathcal{E}(\theta), M_\infty(\theta)$, and G^0 are finite for $\theta\in\Theta_1$.

For (ii) and (iii) of (6), since we can rewrite $\mathcal{E}(\theta)=\xi\xi^*$ with $\xi=(e^{-i\pi\theta_1/2},\cdots,e^{-i\pi\theta_q/2})$,

$$\mathrm{Re}\left[\mathcal{E}(\theta)\right]=\mathrm{Re}\left(\xi\xi^*\right)=\mathrm{Re}\left[\xi\right]\left(\mathrm{Re}\left[\xi\right]\right)'+\mathrm{Im}\left[\xi\right]\left(\mathrm{Im}\left[\xi\right]\right)', \tag{9}$$

and it follows that $\mathrm{Re}\left[\mathcal{E}(\theta)\right]$ is positive semidefinite. Since $M_\infty(\theta)$ and G^0 are positive semidefinite, $\mathrm{Re}\left[\mathcal{E}(\theta)\right]\odot M_\infty(\theta)$ is also positive semidefinite (Lütkepohl, 1996, p.152). It follows from Oppenheim's inequality (Lütkepohl, 1996, p.56) that

$$\Xi(d)\geq\prod_{a=1}^{q}\left(\mathrm{Re}\left[\mathcal{E}(\theta)\right]\odot M_\infty(\theta)\right)_{aa}\det(G^0)=\prod_{a=1}^{q}\left[M_\infty(\theta)\right]_{aa}\left(\det G^0\right)=B(d),$$

giving the second part of (6). (iii) follows because $\Xi(d^0)=\det(M_\infty(0)\odot G^0)=B(d^0)$, since all elements of $\mathcal{E}(0)$ are one.

We move to bound the second probability in (5). Observe that

$$\begin{aligned}
S(d) &= \log\det\widehat{G}(d) - \log\det\widehat{G}(d^0) - 2\sum_{a=1}^{q}\theta_a\frac{1}{m}\sum_{j=1}^{m}\log\lambda_j \\
&= \log\det\frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\left[\Lambda_j(\theta)^{-1}\Lambda_j(d^0)^{-1}I_j\Lambda_j^*(d^0)^{-1}\Lambda_j^*(\theta)^{-1}\right] \\
&\quad -2\sum_{a=1}^{q}\theta_a\frac{1}{m}\sum_{j=1}^{m}\log\lambda_j - \log\det\widehat{G}(d^0) \\
&= \log\det\widehat{D}(d) - \log\det\widehat{D}(d^0),
\end{aligned}$$

where

$$\begin{aligned}
\widehat{D}(d) &= \frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\left[P_j(\theta)\Lambda_j(d^0)^{-1}I_j\Lambda_j^*(d^0)^{-1}P_j^*(\theta)\right], \\
P_j(\theta) &= \mathrm{diag}(e^{i(\lambda_j-\pi)\theta_a/2}(j/p)^{\theta_a}), \\
p &= \exp\left(m^{-1}\sum_{1}^{m}\log j\right) \sim m/e, \quad \text{as } m\to\infty.
\end{aligned}$$

Since $\log x$ is a monotone increasing function of x, $\Pr(\inf_{\Theta_2} S(d) \le 0) \to 0$ follows if

$$\Pr(\inf_{\Theta_2}\det\widehat{D}(d) - \det\widehat{D}(d^0) \le 0) \to 0 \quad \text{as } n\to\infty. \tag{10}$$

For a q-vector W_j, we can write down each summand of $\widehat{D}(d)$ as

$$\begin{aligned}
&\mathrm{Re}[P_j(\theta)\Lambda_j(d^0)^{-1}I_j\Lambda_j^*(d^0)^{-1}P_j^*(\theta)] \\
&= \mathrm{Re}[W_jW_j^*] = \mathrm{Re}[W_j](\mathrm{Re}[W_j])' + \mathrm{Im}[W_j](\mathrm{Im}[W_j])',
\end{aligned}$$

which is positive semidefinite. Thus $\widehat{D}(d)$ is a sum of m positive semidefinite matrices. Define

$$\widehat{D}_\kappa(d) = \frac{1}{m}\sum_{j=[\kappa m]}^{m}\mathrm{Re}\left[P_j(\theta)\Lambda_j(d^0)^{-1}I_j\Lambda_j^*(d^0)^{-1}P_j^*(\theta)\right].$$

Then, it follows from Lütkepohl (1996, p.55) that, for any $\kappa\in(0,1)$,

$$\det\widehat{D}(d) \ge \det\widehat{D}_\kappa(d). \tag{11}$$

Define

$$K_\kappa(d) = \frac{1}{m}\sum_{j=[\kappa m]}^{m}\mathrm{Re}\left[\mathrm{diag}\left(e^{i(\lambda_j-\pi)\theta_a/2}(j/p)^{\theta_a}\right)G^0\ \mathrm{diag}\left(e^{-i(\lambda_j-\pi)\theta_a/2}(j/p)^{\theta_a}\right)\right].$$

The (a,b)th element of $\widehat{D}_\kappa(d) - K_\kappa(d)$ is

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

From summation by parts and Lemma 7.1, this is bounded by, uniformly in $\theta \in \Theta_2$,

$$C(e+o(1))\sum_{r=[\kappa m]}^{m-1}\left(\frac{r}{m}\right)^{-2}\frac{1}{r^2}\left|\sum_{j=[\kappa m]}^{r}\left(e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0}w_{aj}w_{bj}^*-G_{ab}^0\right)\right|$$
$$+\frac{e+o(1)}{m}\left|\sum_{j=[\kappa m]}^{m}\left(e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0}w_{aj}w_{bj}^*-G_{ab}^0\right)\right|=o_p(1).$$

It follows that, for any $\kappa \in (0,1)$,

$$\sup_{\Theta_2}\left|\det \widehat{D}_\kappa(d)-\det K_\kappa(d)\right|=o_p(1), \quad \text{as } n\to\infty$$

We proceed to derive the lower bound of $K_\kappa(d)$ for $d\in\Theta_2$. Rewrite $K_\kappa(d)$ as

$$K_\kappa(d)=\mathrm{Re}\left[\mathcal{E}\left(\theta\right)\right]\odot M_m^\kappa(\theta)\odot G^0,$$

where a positive semidefinite matrix $M_m^\kappa(\theta)$ is defined as

$$M_m^\kappa(\theta)=\frac{1}{m}\sum_{j=[\kappa m]}^{m}\mathrm{Re}\left[Z_jZ_j^*\right],\quad Z_j=\left(e^{i\lambda_j\theta_1/2}(j/p)^{\theta_1},\cdots,e^{i\lambda_j\theta_q/2}(j/p)^{\theta_q}\right)'.$$

Fix $\varepsilon\in(0,0.1)$. Then, in view of (9), Oppenheim's inequality, and Lemmas 7.4 and 7.5, there exists $\bar{\kappa}\in(0,1/4)$ such that, for sufficiently large m and all $\kappa\in(0,\bar{\kappa})$,

$$\begin{aligned}\inf_{\Theta_2}\det K_\kappa(d) &\geq \det G^0\inf_{\Theta_2}\prod_{a=1}^{q}\frac{1}{m}\sum_{j=[\kappa m]}^{m}\left(\frac{j}{p}\right)^{2\theta_a}\\ &\geq \det G^0(1+2\varepsilon)(1-\kappa^{2\Delta})^{q-1}+o(1).\end{aligned}$$

Choose κ sufficiently small so that $(1+2\varepsilon)(1-\kappa^{2\Delta})^{q-1}\geq 1+\varepsilon$. It follows that

$$\inf_{\Theta_2}\det\widehat{D}_\kappa(d)=\inf_{\Theta_2}\det K_\kappa(d)+o_p(1)\geq\det G^0(1+\varepsilon)+o_p(1).$$

From the results for $d\in\Theta_1$, we have $\det\widehat{D}(d^0)\to_p\det G^0$ as $n\to\infty$. Therefore,

$$\Pr(\inf_{\Theta_2}\det\widehat{D}_\kappa(d)-\det\widehat{D}(d^0)\leq 0)\to 0\quad\text{as } n\to\infty,$$

and (10) follows in view of (11), completing the proof. ■

6.2 Proof of Theorem 4.1

Theorem 3.1 holds under the current conditions and implies that with probability approaching to one, as $n\to\infty$, $\widehat{d}$ satisfies

$$0=\left.\frac{\mathrm{d}R(d)}{\mathrm{d}d}\right|_{\widehat{d}}=\left.\frac{\mathrm{d}R(d)}{\mathrm{d}d}\right|_{d^0}+\left(\left.\frac{\mathrm{d}^2R(d)}{\mathrm{d}d\mathrm{d}d'}\right|_{\widetilde{d}}\right)(\widehat{d}-d^0).$$

where $||\widetilde{d}-d^0||\leq||\widehat{d}-d^0||$. $\widehat{d}$ has the stated limiting distribution if, for any $q\times 1$ vector η, as $n\to\infty$,

$$\eta'\sqrt{m}\left.\frac{\mathrm{d}R(d)}{\mathrm{d}d}\right|_{d^0}=\sum_{a=1}^{q}\eta_a\sqrt{m}\left.\frac{\partial R(d)}{\partial d_a}\right|_{d^0}\to_d N(0,\eta'\Omega\eta), \tag{12}$$

$$\left.\frac{\mathrm{d}^2R(d)}{\mathrm{d}d\mathrm{d}d'}\right|_{\widetilde{d}}\to{}_p\Omega,\quad\Omega=2\left[I_q+G^0\odot(G^0)^{-1}+\frac{\pi^2}{4}\left(G^0\odot(G^0)^{-1}-I_q\right)\right] \tag{13}$$

6.2.1 Score vector approximation

First we show (12). Observe that

$$\sqrt{m}\frac{\partial R(d)}{\partial d_a} = -\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\log\lambda_j + \operatorname{tr}\left[\widehat{G}(d)^{-1}\sqrt{m}\frac{\partial\widehat{G}(d)}{\partial d_a}\right].$$

Let i_a be a $q\times q$ matrix whose ath diagonal element is one and all other elements are zero, and let Λ_j^0 denote $\Lambda_j(d^0)$ in the following. It follows that

$$\begin{aligned}
\sqrt{m}\left.\frac{\partial\widehat{G}(d)}{\partial d_a}\right|_{d^0} &= \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\operatorname{Re}\left[\left(\log\lambda_j + \frac{\lambda_j-\pi}{2}i\right)(\Lambda_j^0)^{-1}i_a I_j(\Lambda_j^{0*})^{-1}\right] \\
&\quad + \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\operatorname{Re}\left[\left(\log\lambda_j - \frac{\lambda_{j-\pi}}{2}i\right)(\Lambda_j^0)^{-1}I_j i_a(\Lambda_j^{0*})^{-1}\right] \\
&= \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\log\lambda_j\operatorname{Re}\left[(\Lambda_j^0)^{-1}(i_aI_j + I_ji_a)(\Lambda_j^{0*})^{-1}\right] \\
&\quad + \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\frac{\lambda_j-\pi}{2}\operatorname{Im}\left[(\Lambda_j^0)^{-1}(i_aI_j - I_ji_a)(\Lambda_j^{0*})^{-1}\right], \\
&= H_{1a} + H_{2a}.
\end{aligned}$$

Therefore, $\sum_{a=1}^{q}\eta_a\sqrt{m}(\partial R(d))/(\partial d_a)|_{d^0}$ is equal to

$$\begin{aligned}
&\sum_{a=1}^{q}\eta_a\left\{-\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\log\lambda_j + \operatorname{tr}\left[\widehat{G}\left(d^0\right)^{-1}H_{1a}\right]\right\} + \sum_{a=1}^{q}\eta_a\left\{\operatorname{tr}\left[\widehat{G}\left(d^0\right)^{-1}H_{2a}\right]\right\} \\
= \;& R_1 + R_2.
\end{aligned}$$

We proceed to find an approximation of R_1 and R_2. For R_1, define

$$\widetilde{G}_1 = \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\log\lambda_j\operatorname{Re}\left[(\Lambda_j^0)^{-1}I_j(\Lambda_j^{0*})^{-1}\right],$$

so that $H_{1a} = i_a\widetilde{G}_1 + \widetilde{G}_1 i_a$. It follows that, with $\nu_j = \log\lambda_j - m^{-1}\sum_1^m\log\lambda_j = \log j - m^{-1}\sum_1^m\log j = O(\log m)$,

$$\begin{aligned}
&-\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\log\lambda_j + \operatorname{tr}\left[\widehat{G}\left(d^0\right)^{-1}H_{1a}\right] \\
=\;& 2\sqrt{m}\operatorname{tr}\left[\widehat{G}^{-1}(d^0)\left[\widetilde{G}_1 - \frac{1}{m}\sum_{1}^{m}\log\lambda_j\widehat{G}(d^0)\right]i_a\right] \\
=\;& 2\operatorname{tr}\left[\widehat{G}^{-1}(d^0)\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\operatorname{Re}\left[(\Lambda_j^0)^{-1}I_j(\Lambda_j^{0*})^{-1}\right]i_a\right] \\
=\;& (g^a + o_p(1))\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left\{\operatorname{Re}\left[(\Lambda_j^0)^{-1}I_j(\Lambda_j^{0*})^{-1}\right]\right\}_a, \qquad (14)
\end{aligned}$$

where g^a is the ath row of $(G^0)^{-1}$ and $\{A\}_a$ denotes the ath column of matrix A. Observe that

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j(\Lambda_j^0)^{-1}I_j(\Lambda_j^{0*})^{-1}$$

$$= \quad \frac{1}{\sqrt{m}} \sum_{j=1}^{m} \nu_j \left[(\Lambda_j^0)^{-1} A(\lambda_j) I_{\varepsilon j} A^*(\lambda_j) (\Lambda_j^{0*})^{-1} - G^0 \right] + o_p(1) = O_p(1), \quad (15)$$

where the first equality follows from summation by parts, Lemma 7.1 (b1), and $\sum_1^m \nu_j = 0$, and the second equality follows from $E[(\Lambda_j^0)^{-1} A(\lambda_j) I_{\varepsilon j} A^*(\lambda_j)(\Lambda_j^{0*})^{-1}] - G^0 = O(j^\beta n^{-\beta})$, $\mathrm{Cov}(I_{\varepsilon j}, I_{\varepsilon k}) = O(1)$ if $j = k$ and $O(n^{-1})$ if $j \neq k$, and $\sum_1^m \nu_j^2 = O(m)$. It follows that

$$R_1 = \frac{2}{\sqrt{m}} \sum_{a=1}^{q} \eta_a \sum_{j=1}^{m} \nu_j \left(g^a \left\{ \mathrm{Re} \left[(\Lambda_j^0)^{-1} A(\lambda_j) I_{\varepsilon j} A^*(\lambda_j) (\Lambda_j^{0*})^{-1} \right] \right\}_a - 1 \right) + o_p(1).$$

The first term is equal to

$$\frac{2}{\sqrt{m}} \sum_{a=1}^{q} \eta_a \sum_{j=1}^{m} \nu_j \left(g^a \left\{ \mathrm{Re} \left[(\Lambda_j^0)^{-1} A(\lambda_j) \left(\frac{1}{2\pi n} \sum_{t=1}^{n} \varepsilon_t \varepsilon_t' \right) A^*(\lambda_j) (\Lambda_j^{0*})^{-1} \right] \right\}_a - 1 \right)$$
$$+ \frac{2}{\sqrt{m}} \sum_{a=1}^{q} \eta_a \sum_{j=1}^{m} \nu_j \left(g^a \left\{ \mathrm{Re} \left[(\Lambda_j^0)^{-1} A(\lambda_j) \left(\frac{1}{2\pi n} \sum_{t \neq s}^{n} \sum \varepsilon_t \varepsilon_s' e^{i(t-s)\lambda_j} \right) A^*(\lambda_j) (\Lambda_j^{0*})^{-1} \right] \right\}_a \right).$$

The first part is $o_p(1)$ because $E||n^{-1} \sum_{t=1}^n \varepsilon_t \varepsilon_t' - I_p|| = O(n^{-1/2})$ and Assumption $1'$ imply that

$$\begin{aligned}
& \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \nu_j \left(g^a \left\{ (\Lambda_j^0)^{-1} A(\lambda_j) \left(\frac{1}{2\pi n} \sum_{t=1}^{n} \varepsilon_t \varepsilon_t' \right) A^*(\lambda_j) (\Lambda_j^{0*})^{-1} \right\}_a - 1 \right) \\
= \; & \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \nu_j \left(g^a \left\{ (\Lambda_j^0)^{-1} A(\lambda_j) \left(\frac{1}{2\pi n} \sum_{t=1}^{n} \varepsilon_t \varepsilon_t' - \frac{I_p}{2\pi} \right) A^*(\lambda_j) (\Lambda_j^{0*})^{-1} \right\}_a - 1 \right) \\
& + \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \nu_j \left(\left\{ g^a (\Lambda_j^0)^{-1} \frac{A(\lambda_j) A^*(\lambda_j)}{2\pi} (\Lambda_j^{0*})^{-1} \right\}_a - 1 \right) \\
= \; & O_p(m^{1/2} n^{-1/2} \log m) + O(m^{\beta+1/2} n^{-\beta} \log m). \qquad (16)
\end{aligned}$$

The second part can be rewritten as

$$\sum_{t=1}^{n} \varepsilon_t' \sum_{s=1}^{t-1} \mathrm{Re}[\Phi_{t-s}] \varepsilon_s, \quad \Phi_t = (\pi \sqrt{m} n)^{-1} \sum_{j=1}^{m} \nu_j \left[\Omega_j e^{-it\lambda_j} + \Omega_j' e^{it)\lambda_j} \right], \quad (17)$$

where Ω_j is defined as

$$\Omega_j = \sum_{a=1}^{q} \eta_a \left\{ A^*(\lambda_j) (\Lambda_j^{0*})^{-1} \right\}_a g^a (\Lambda_j^0)^{-1} A(\lambda_j). \quad (18)$$

Rewrite $\sum_{t=1}^n \varepsilon_t' \sum_{s=1}^{t-1} \mathrm{Re}\left[\Phi_{t-s}\right] \varepsilon_s$ as

$$\sum_{t=1}^{n} \varepsilon_t' \sum_{s=1}^{t-1} \left\{ \frac{1}{\pi \sqrt{m} n} \sum_{j=1}^{m} \nu_j \, \mathrm{Re} \left[\Omega_j + \Omega_j' \right] \cos(t-s)\lambda_j \right\} \varepsilon_s$$
$$+ \sum_{t=1}^{n} \varepsilon_t' \sum_{s=1}^{t-1} \left\{ \frac{1}{\pi \sqrt{m} n} \sum_{j=1}^{m} \nu_j \, \mathrm{Im} \left[\Omega_j - \Omega_j' \right] \sin(t-s)\lambda_j \right\} \varepsilon_s.$$

The second term is $o_p(1)$ because its second moment is equal to

$$\frac{1}{\pi^2 mn^2}\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\sum_{j=1}^{m}\nu_j^2 \mathrm{tr}\left\{\left(\mathrm{Im}\left[\Omega_j-\Omega_j'\right]\right)'\mathrm{Im}\left[\Omega_j-\Omega_j'\right]\right\}\sin^2(s\lambda_j)$$
$$+\frac{1}{\pi^2 mn^2}\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\sum\sum_{j\neq k}^{m}\nu_j\nu_k \mathrm{tr}\left\{\left(\mathrm{Im}\left[\Omega_j-\Omega_j'\right]\right)'\mathrm{Im}\left[\Omega_k-\Omega_k'\right]\right\}\sin(s\lambda_j)\sin(s\lambda_k). \tag{19}$$

This is $o(1)$ from $\mathrm{Im}[\Omega_j]\to 0$ and Lemma 7.7 (c) and (d). Therefore, we can rewrite R_1 as

$$R_1=\sum_{t=1}^{n}\varepsilon_t'\sum_{s=1}^{t-1}\Theta_{t-s}\varepsilon_s+o_p(1);\qquad \Theta_s=\frac{1}{\pi\sqrt{mn}}\sum_{j=1}^{m}\nu_j\,\mathrm{Re}\left[\Omega_j+\Omega_j'\right]\cos(s\lambda_j).$$

We move to R_2. An argument similar to (15) and (16) gives

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}(\lambda_j-\pi)(\Lambda_j^0)^{-1}I_j(\Lambda_j^{0*})^{-1}=-\frac{\pi}{\sqrt{m}}\sum_{j=1}^{m}(\Lambda_j^0)^{-1}A(\lambda_j)I_{\varepsilon j}A^*(\lambda_j)(\Lambda_j^{0*})^{-1}+o_p(1)=O_p(1),$$

and

$$\frac{2}{\sqrt{m}}\sum_{a=1}^{q}\eta_a\sum_{j=1}^{m}\nu_j\left(g^a\left\{\mathrm{Im}\left[(\Lambda_j^0)^{-1}A(\lambda_j)\left(\frac{1}{2\pi n}\sum_{t=1}^{n}\varepsilon_t\varepsilon_t'\right)A^*(\lambda_j)(\Lambda_j^{0*})^{-1}\right]\right\}_a\right)=o_p(1).$$

Therefore, R_2 is equal to

$$\frac{\pi}{2}\frac{2}{\sqrt{m}}\sum_{a=1}^{q}\eta_a\sum_{j=1}^{m}\left(g^a\left\{\mathrm{Im}\left[(\Lambda_j^0)^{-1}A(\lambda_j)\left(\frac{1}{2\pi n}\sum\sum_{t\neq s}\varepsilon_t\varepsilon_s'e^{i(t-s)\lambda_j}\right)A^*(\lambda_j)(\Lambda_j^{0*})^{-1}\right]\right\}_a\right)+o_p(1)$$
$$=\frac{\pi}{2}\sum_{t=1}^{n}\varepsilon_t'\sum_{s=1}^{t-1}\mathrm{Im}[\widetilde{\Phi}_{t-s}]\varepsilon_s+o_p(1),\quad \widetilde{\Phi}_t=(\pi\sqrt{mn})^{-1}\sum_{j=1}^{m}\left[\Omega_je^{-it\lambda_j}+\Omega_j'e^{it)\lambda_j}\right].$$

where Ω_j is defined in (18). Rewrite $(\pi/2)\sum_{t=1}^{n}\varepsilon_t'\sum_{s=1}^{t-1}\mathrm{Im}[\widetilde{\Phi}_{t-s}]\varepsilon_s$ as

$$\frac{\pi}{2}\sum_{t=1}^{n}\varepsilon_t'\sum_{s=1}^{t-1}\left\{\frac{1}{\pi\sqrt{mn}}\sum_{j=1}^{m}\mathrm{Re}\left[-\Omega_j+\Omega_j'\right]\sin(t-s)\lambda_j\right\}\varepsilon_s$$
$$+\frac{\pi}{2}\sum_{t=1}^{n}\varepsilon_t'\sum_{s=1}^{t-1}\left\{\frac{1}{\pi\sqrt{mn}}\sum_{j=1}^{m}\mathrm{Im}\left[\Omega_j+\Omega_j'\right]\cos(t-s)\lambda_j\right\}\varepsilon_s,$$

The second term is $o_p(1)$ from an decomposition similar to (19), $\mathrm{Im}[\Omega_j]\to 0$ and Lemma 7.7 (a) and (b). Hence, we can rewrite R_2 as

$$R_2=\sum_{t=1}^{n}\varepsilon_t'\sum_{s=1}^{t-1}\widetilde{\Theta}_{t-s}\varepsilon_s+o_p(1);\qquad \widetilde{\Theta}_s=\frac{1}{\pi\sqrt{mn}}\sum_{j=1}^{m}\mathrm{Re}\left[-\Omega_j+\Omega_j'\right]\sin(s\lambda_j).$$

It follows that, with $z_1=0$,

$$\sum_{a=1}^{q}\eta_a\sqrt{m}\frac{\partial R(d)}{\partial d_a}\bigg|_{d^0}=\sum_{t=1}^{n}z_t+o_p(1),\quad z_t=\varepsilon_t'\sum_{s=1}^{t-1}\left[\Theta_{t-s}+\widetilde{\Theta}_{t-s}\right]\varepsilon_s.$$

By a standard martingale CLT, (12) follows if

$$\sum_{t=1}^{n} E(z_t^2|F_{t-1}) - \sum_{a=1}^{q}\sum_{b=1}^{q}\eta_a\eta_b\Omega_{ab} \quad \to \quad {}_p0, \tag{20}$$

$$\sum_{t=1}^{n} E(z_t^2 I(|z_t| \quad > \quad \delta)) \to 0 \quad \text{for all } \delta > 0. \tag{21}$$

Following the argument in Lobato (1999, pp. 149-51), we obtain $||\Theta_s||, ||\widetilde{\Theta}_s|| = O(n^{-1}m^{1/2}\log m)$ for $1 \leq s \leq n/2$ and $||\Theta_s||, ||\widetilde{\Theta}_s|| = O(m^{-1/2}s^{-1}\log m)$, and Assumption 1 implies that $\Omega_j = O(1)$. Therefore, Lemmas 2 and 3 in Lobato (1999) hold for Θ_s and $\widetilde{\Theta}_s$ defined in this paper, and Lemma 4 of Lobato holds for Ω_j defined in (18). Hence, we can apply the arguments in Lobato (1999, pp. 142-43) to show that (21) holds. For (20), from the results in Lobato (1999, p.142 and Lemmas 2 and 3), we have

$$\sum_{t=1}^{n} E(z_t^2|F_{t-1}) = \sum_{t=2}^{n}\sum_{s=1}^{t-1}\left[\Theta_{t-s} + \widetilde{\Theta}_{t-s}\right]'\left[\Theta_{t-s} + \widetilde{\Theta}_{t-s}\right] + o_p(1).$$

Now

$$\begin{aligned}
&\sum_{t=2}^{n}\sum_{s=1}^{t-1}\left[\Theta'_{t-s}\Theta_{t-s} + \widetilde{\Theta}'_{t-s}\widetilde{\Theta}_{t-s}\right] \\
= \ & \frac{1}{\pi^2 mn^2}\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\sum_{j=1}^{m}\nu_j^2 \mathrm{tr}\left\{\left(\mathrm{Re}\left[\Omega_j + \Omega_j'\right]\right)'\mathrm{Re}\left[\Omega_j + \Omega_j'\right]\right\}\cos^2(s\lambda_j) \\
&+ \frac{1}{\pi^2 mn^2}\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\sum\sum_{j\neq k}^{m}\nu_j\nu_k \mathrm{tr}\left\{\left(\mathrm{Re}\left[\Omega_j + \Omega_j'\right]\right)'\mathrm{Re}\left[\Omega_k + \Omega_k'\right]\right\}\cos(s\lambda_j)\cos(s\lambda_k) \\
&+ \frac{1}{\pi^2 mn^2}\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\sum_{j=1}^{m}\mathrm{tr}\left\{\left(\mathrm{Re}\left[-\Omega_j + \Omega_j'\right]\right)'\mathrm{Re}\left[-\Omega_j + \Omega_j'\right]\right\}\sin^2(s\lambda_j) \\
&+ \frac{1}{\pi^2 mn^2}\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\sum\sum_{j\neq k}^{m}\mathrm{tr}\left\{\left(\mathrm{Re}\left[-\Omega_j + \Omega_j'\right]\right)'\mathrm{Re}\left[-\Omega_k + \Omega_k'\right]\right\}\sin(s\lambda_j)\sin(s\lambda_k).
\end{aligned}$$

The second and fourth terms are $o(1)$ from $\Omega_j = O(1)$ and Lemma 7.7 (b) and (d). For the first and third terms, observe that

$$\left\{\begin{array}{l} \mathrm{tr}\left\{(4\pi^2)^{-1}\mathrm{Re}\left[\Omega_j'\right]\mathrm{Re}\left[\Omega_j\right]\right\} \to 2\sum_{a=1}^{q}\sum_{b=1}^{q}\eta_a\eta_b G^0_{ab}(G^0)^{-1}_{ba}, \\ \mathrm{tr}\left\{(4\pi^2)^{-1}\mathrm{Re}\left[\Omega_j\right]\mathrm{Re}\left[\Omega_j\right]\right\} \to 2\sum_{a=1}^{q}\eta_a^2, \end{array}\right. \tag{22}$$

as $\lambda_j \to 0$. It follows that

$$\mathrm{tr}\left[(4\pi^2)^{-1}\left(\mathrm{Re}\left[\Omega_j + \Omega_j'\right]\right)'\mathrm{Re}\left[\Omega_j + \Omega_j'\right]\right] \quad \to \quad 2\sum_{a=1}^{q}\sum_{b=1}^{q}\eta_a\eta_b G^0_{ab}(G^0)^{-1}_{ba} + 2\sum_{a=1}^{q}\eta_a^2,$$

$$\mathrm{tr}\left[(4\pi^2)^{-1}\left(\mathrm{Re}\left[-\Omega_j + \Omega_j'\right]\right)'\mathrm{Re}\left[-\Omega_j + \Omega_j'\right]\right] \quad \to \quad 2\sum_{a=1}^{q}\sum_{b=1}^{q}\eta_a\eta_b G^0_{ab}(G^0)^{-1}_{ba} - 2\sum_{a=1}^{q}\eta_a^2,$$

as $\lambda_j \to 0$, and the sum of the first and third terms converge to $\sum_{a=1}^q \sum_{b=1}^q \eta_a \eta_b \Omega_{ab}$ in view of Lemma 7.7 (a) and (c). Finally,

$$
\begin{aligned}
&\sum_{t=2}^{n}\sum_{s=1}^{t-1}\Theta_{t-s}'\widetilde{\Theta}_{t-s} \\
= \quad & \frac{1}{\pi^2 mn^2}\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\sum_{j=1}^{m}\nu_j \mathrm{tr}\left\{\left(\mathrm{Re}\left[\Omega_j+\Omega_j'\right]\right)'\mathrm{Re}\left[-\Omega_j+\Omega_j'\right]\right\}\cos(s\lambda_j)\sin(s\lambda_j) \\
& +\frac{1}{\pi^2 mn^2}\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\sum\sum_{j\neq k}^{m}\nu_j \mathrm{tr}\left\{\left(\mathrm{Re}\left[\Omega_j+\Omega_j'\right]\right)'\mathrm{Re}\left[-\Omega_k+\Omega_k'\right]\right\}\cos(s\lambda_j)\sin(s\lambda_k).
\end{aligned}
$$

The first term on the right is zero because $\mathrm{tr}\{\mathrm{Re}[\Omega_j+\Omega_j']\}'\{\mathrm{Re}[-\Omega_j+\Omega_j]\}=0$. Since $\Omega_j=\sum_{a=1}^q \eta_a\{H'\}_a g^a H + O(m^\alpha n^{-\alpha})$, the second term on the right is equal to

$$
\begin{aligned}
& \frac{1}{\pi^2 mn^2}\sum\sum_{j\neq k}^{m}\nu_j O(m^\alpha n^{-\alpha})\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\cos(s\lambda_j)\sin(s\lambda_k) \\
= \quad & O\left(m^{\alpha-1}n^{-\alpha}\log m\sum\sum_{j\neq k}^{m}\left(\frac{1}{j+k}+\frac{1}{|j-k|}\right)\right) = O\left(m^\alpha n^{-\alpha}(\log m)^2\right)=o(1),
\end{aligned}
$$

from Lemma 7.7 (d). Therefore, $\sum_{t=2}^n\sum_{s=1}^{t-1}[\Theta_{t-s}+\widetilde{\Theta}_{t-s}]'[\Theta_{t-s}+\widetilde{\Theta}_{t-s}] \to \sum_{a=1}^q\sum_{b=1}^q \eta_a\eta_b\Omega_{ab}$ and (20) follows.

6.2.2 Hessian approximation

Define $\theta = d - d^0$. Fix $\varepsilon > 0$ and let $M = \{d : (\log n)^4||d-d^0|| < \varepsilon\} = \{\theta : (\log n)^4||\theta|| < \varepsilon\}$. First, we show $\Pr(\widetilde{d} \notin M) \to 0$ as $n \to \infty$. Using the notations in the proof of Theorem 3.1, $\inf_{\Theta_1\backslash M} S_2(d)$ is bounded as

$$\inf_{\Theta_1\backslash M} S_2(d) \geq \varepsilon^2(\log n)^8/6.$$

By applying Lemma 7.1 (b2) to (8), we strengthen (i) of (6) to

$$\sup_{\Theta_1}|A(d)-\Xi(d)| = O_p(m^\beta n^{-\beta} + m^{-2\Delta}\log m + m^{-2/3}\log m + m^{-1/2} + mn^{-1}).$$

It follows that, uniformly in Θ_1,

$$
\begin{aligned}
S_1(d) &= \log\left(1+\frac{\Xi(d)-B(d)+o_p((\log n)^{-8})}{B(d)}\right), \\
S_1(d^0) &= \log\left(1+o_p((\log n)^{-8})/B(d^0)\right),
\end{aligned}
$$

and since $\Xi(d)-B(d)\geq 0$ and $\inf_{\Theta_1} B(d) > 0$ we obtain

$$\Pr\left(\inf_{\Theta_1\backslash M}\left[S_1(d)-S_1(d^0)+S_2(d)\right]\leq 0\right)\to 0, \quad \text{as } n\to\infty.$$

Therefore, $\Pr(\widetilde{d} \notin M) \to 0$ as $n \to \infty$ follows.

Observe that

$$\frac{\partial^2 R(d)}{\partial d_a \partial d_b} = \mathrm{tr}\left[-\widehat{G}^{-1}(d)\frac{\partial \widehat{G}(d)}{\partial d_a}\widehat{G}^{-1}(d)\frac{\partial \widehat{G}(d)}{\partial d_b} + \widehat{G}^{-1}(d)\frac{\partial^2 \widehat{G}(d)}{\partial d_a \partial d_b}\right]. \quad (23)$$

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

$$\begin{aligned}\frac{\partial \widehat{G}(d)}{\partial d_a} &= \frac{1}{m}\sum_{j=1}^{m}\operatorname{Re}\left[\left(\log\lambda_j + i\frac{\lambda_j-\pi}{2}\right) i_a\Lambda_j(d)^{-1}I_j\Lambda_j^*(d)^{-1}\right] \\ &+\frac{1}{m}\sum_{j=1}^{m}\operatorname{Re}\left[\left(\log\lambda_j - i\frac{\lambda_j-\pi}{2}\right)\Lambda_j(d)^{-1}I_j\Lambda_j^*(d)^{-1}i_a\right],\end{aligned}$$

and

$$\begin{aligned}\frac{\partial^2 \widehat{G}(d)}{\partial d_a\partial d_b} &= \frac{1}{m}\sum_{j=1}^{m}\operatorname{Re}\left[\left(\log\lambda_j + i\frac{\lambda_j-\pi}{2}\right)^2 i_a i_b\Lambda_j(d)^{-1}I_j\Lambda_j^*(d)^{-1}\right] \\ &+\frac{1}{m}\sum_{j=1}^{m}\operatorname{Re}\left[\left((\log\lambda_j)^2 + \frac{(\lambda_j-\pi)^2}{4}\right) i_a\Lambda_j(d)^{-1}I_j\Lambda_j^*(d)^{-1}i_b\right] \\ &+\frac{1}{m}\sum_{j=1}^{m}\operatorname{Re}\left[\left((\log\lambda_j)^2 + \frac{(\lambda_j-\pi)^2}{4}\right) i_b\Lambda_j(d)^{-1}I_j\Lambda_j^*(d)^{-1}i_a\right] \\ &+\frac{1}{m}\sum_{j=1}^{m}\operatorname{Re}\left[\left(\log\lambda_j - i\frac{\lambda_j-\pi}{2}\right)^2 \Lambda_j(d)^{-1}I_j\Lambda_j^*(d)^{-1}i_a i_b\right].\end{aligned}$$

Define, for $k=0,1,2$,

$$\begin{aligned}\widehat{G}_k(d) &= m^{-1}\textstyle\sum_{j=1}^{m}(\log\lambda_j)^k\operatorname{Re}\left[\Lambda_j(d)^{-1}I_j\Lambda_j^*(d)^{-1}\right], \\ \overline{G}_k(d) &= m^{-1}\textstyle\sum_{j=1}^{m}(\log\lambda_j)^k\operatorname{Im}\left[\Lambda_j(d)^{-1}I_j\Lambda_j^*(d)^{-1}\right],\end{aligned}$$

Then it follows that

$$\begin{aligned}\frac{\partial \widehat{G}(d)}{\partial d_a} &= i_a\widehat{G}_1(d) + \widehat{G}_1(d)i_a + (\pi/2)i_a\overline{G}_0(d) - (\pi/2)\overline{G}_0(d)i_a + o_p((\log n)^{-1}), \\ \frac{\partial^2 \widehat{G}(d)}{\partial d_a\partial d_b} &= i_a i_b\widehat{G}_2(d) + i_a\widehat{G}_2(d)i_b + i_b\widehat{G}_2(d)i_a + \widehat{G}_2(d)i_a i_b \\ &+(\pi^2/4)\left[-i_a i_b\widehat{G}(d) + i_a\widehat{G}(d)i_b + i_b\widehat{G}(d)i_a - \widehat{G}(d)i_a i_b\right] \\ &+\pi i_a i_b\overline{G}_1(d) - \pi\overline{G}_1(d)i_a i_b + o_p(1),\end{aligned}$$

where the order of the reminder terms follows from summation by parts, $\sum_{j=1}^{r}\Lambda_j(d)^{-1}I_j\Lambda_j^*(d)^{-1} = O_p(r)$, and Assumption 4′. We proceed to show, uniformly in $d\in M$,

$$\widehat{G}_k(d) = G^0 m^{-1}\textstyle\sum_{j=1}^{m}(\log\lambda_j)^k + o_p((\log n)^{k-2}), \quad \overline{G}_k(d) = o_p((\log n)^{k-2}). \qquad (24)$$

The assumption $m^{-\gamma}\log n \to 0$ is necessary here, because the terms with $\overline{G}_1(d)$ do not cancel out even if we take the trace of $\widehat{G}^{-1}(d)(\partial^2\widehat{G}(d))/(\partial d_a\partial d_b)$. Define

$$F_k(d) = m^{-1}\textstyle\sum_{j=1}^{m}(\log\lambda_j)^k\Lambda_j(\theta)^{-1}G^0\Lambda_j^*(\theta)^{-1},$$

then (24) follows if

$$\begin{aligned}\sup_{d\in M}\left\|m^{-1}\textstyle\sum_{j=1}^{m}(\log\lambda_j)^k\left[\Lambda_j(d)^{-1}I_j\Lambda_j^*(d)^{-1}\right] - F_k(d)\right\| &= o_p((\log n)^{k-2}), \quad (25) \\ \sup_{d\in M}\left\|F_k(d) - G^0 m^{-1}\textstyle\sum_{j=1}^{m}(\log\lambda_j)^k\right\| &= o((\log n)^{k-2}). \quad (26)\end{aligned}$$

We show (25). The (a,b)th element of the left hand side of (25) is equal to

$$m^{-1}\sum_{j=1}^{m}(\log\lambda_j)^k e^{i(\lambda_j-\pi)(\theta_a-\theta_b)/2}\lambda_j^{\theta_a+\theta_b}\left[e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0}w_{aj}w_{bj}^*-G_{ab}^0\right].$$

Define $b_{nj}(\theta)=(\log\lambda_j)^k e^{i(\lambda_j-\pi)(\theta_a-\theta_b)/2}\lambda_j^{\theta_a+\theta_b}$ then it follows from the summation by parts that the above is equal to

$$m^{-1}\sum_{r=1}^{m-1}\left[b_{nr}(\theta)-b_{n,r+1}(\theta)\right]\sum_{j=1}^{r}\left[e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0}w_{aj}w_{bj}^*-G_{ab}^0\right] \quad (27)$$
$$+m^{-1}b_{nm}(\theta)\sum_{j=1}^{m}\left[e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0}w_{aj}w_{bj}^*-G_{ab}^0\right].$$

Since $(\log\lambda_r)^k = O((\log n)^k)$, $(\log\lambda_r)^k-(\log\lambda_{r+1})^k = O(r^{-1})$, $\lambda_{r+1}^{\theta_a+\theta_b} = O(1)$, $\lambda_r^{\theta_a+\theta_b}-\lambda_{r+1}^{\theta_a+\theta_b}=O(r^{-1})$, $e^{i(\lambda_j-\pi)(\theta_a-\theta_b)/2}=O(1)$, and $e^{i(\lambda_r-\pi)(\theta_a-\theta_b)/2}-e^{i(\lambda_{r+1}-\pi)(\theta_a-\theta_b)/2}=O(r^{-1})$ uniformly in $\theta\in M$, we obtain

$$b_{nr}(\theta)-b_{n,r+1}(\theta)=O((\log n)^k r^{-1}),\quad b_{nm}=O((\log n)^k).$$

In conjunction with Lemma 7.1 (b2), we have

$$\begin{aligned}(27) &= O_p\left((\log n)^k m^{-1}\sum_{r=1}^{m}\left(r^{\beta}n^{-\beta}+r^{-2/3}(\log m)^{2/3}+r^{-1}\log m+r^{-1/2}\right)\right)\\ &= O_p\left((\log n)^k\left(m^{\beta}n^{-\beta}+m^{-2/3}(\log m)^{2/3}+m^{-1}(\log m)^2+m^{-1/2}\right)\right)\\ &= o_p\left((\log n)^{k-2}\right),\end{aligned}$$

giving (25).

We move to the proof of (26). The (a,b)th element of the left hand side of (26) is equal to

$$m^{-1}\sum_{j=1}^{m}(\log\lambda_j)^k\left[e^{i(\lambda_j-\pi)(\theta_a-\theta_b)/2}\lambda_j^{\theta_a+\theta_b}-1\right]G_{ab}^0.$$

Since, for $\theta\in M$ and $0<\lambda_j\le 1$,

$$|\lambda_j^{\theta_a+\theta_b}-1|/|\theta_a+\theta_b|\le|\log\lambda_j|n^{|\theta_a|+|\theta_b|}\le(\log n)n^{1/\log n}\le C\log n, \quad (28)$$

it follows that

$$\begin{aligned}e^{i(\lambda_j-\pi)(\theta_a-\theta_b)/2}\lambda_j^{\theta_a+\theta_b}-1 &= (e^{i(\lambda_j-\pi)(\theta_a-\theta_b)/2}-1)\lambda_j^{\theta_a+\theta_b}+(\lambda_j^{\theta_a+\theta_b}-1)\\ &\le C\left(|\theta_a|+|\theta_b|\right)+C(|\theta_a|+|\theta_b|)\log n=O((\log n)^{-3}).\end{aligned}$$

Therefore,

$$m^{-1}\sum_{j=1}^{m}(\log\lambda_j)^k\left[e^{i(\lambda_j-\pi)(\theta_a-\theta_b)/2}\lambda_j^{\theta_a+\theta_b}-1\right]G_{ab}^0=o\left((\log n)^{k-2}\right),$$

giving (26).

Define $G_{1a}^0=i_aG^0+G^0i_a$, $G_{2ab}^0=i_ai_bG^0+i_aG^0i_b+i_bG^0i_a+G^0i_ai_b$, and $G_{3ab}^0=-i_ai_bG^0+i_aG^0i_b+i_bG^0i_a-G^0i_ai_b$. It follows from (24) that

$$\begin{aligned}&\widehat{G}^{-1}(\widetilde{d})(\partial\widehat{G}(\widetilde{d})/\partial d_a)\widehat{G}^{-1}(\widetilde{d})(\partial\widehat{G}(\widetilde{d})/\partial d_b)\\ =\ &\left[G^0+o_p((\log n)^{-2})\right]^{-1}\left[m^{-1}\sum_{j=1}^{m}(\log\lambda_j)G_{1a}^0+o_p((\log n)^{-1})\right]\\ &\times\left[G^0+o_p((\log n)^{-2})\right]^{-1}\left[m^{-1}\sum_{j=1}^{m}(\log\lambda_j)G_{1b}^0+o_p((\log n)^{-1})\right]\\ =\ &\left[m^{-1}\sum_{j=1}^{m}(\log\lambda_j)\right]^2\left[(G^0)^{-1}G_{1a}^0+(G^0)^{-1}G_{1b}^0\right]+o_p(1),\end{aligned}$$

and

$$
\begin{aligned}
& \widehat{G}^{-1}(\widetilde{d})(\partial^2\widehat{G}(\widetilde{d})/\partial d_a \partial d_b) \\
= & \left[G^0 + o_p((\log n)^{-2})\right]^{-1}\left[m^{-1}\textstyle\sum_{j=1}^m (\log\lambda_j)^2 G^0_{2ab} + (\pi^2/4)G^0_{3ab} + o_p(1)\right] \\
= & m^{-1}\textstyle\sum_{j=1}^m (\log\lambda_j)^2 (G^0)^{-1}G^0_{2ab} + (\pi^2/4)(G^0)^{-1}G^0_{3ab} + o_p(1).
\end{aligned}
$$

Since $\mathrm{tr}[(G^0)^{-1}G^0_{1a}(G^0)^{-1}G^0_{1b}] = \mathrm{tr}[(G^0)^{-1}G^0_{2ab}]$ and $m^{-1}\sum_{j=1}^m(\log\lambda_j)^2 - [m^{-1}\sum_{j=1}^m(\log\lambda_j)]^2 \to 1$, we obtain

$$
\frac{\partial^2 R(\widetilde{d})}{\partial d_a \partial d_b} = \mathrm{tr}\left[(G^0)^{-1}G^0_{2ab} + (\pi^2/4)(G^0)^{-1}G^0_{3ab}\right] + o_p(1),
$$

and (13) follows.

$\widehat{G}(\widehat{d}) \to_p G^0$ follows from (24) and $\Pr(\widehat{d} \notin M) \to 0$ as $n \to \infty$, completing the proof. ■

7 Appendix B: technical lemmas

Lemmas 7.3 and 7.4 are from Shimotsu and Phillips (2003, Section 5). They are given for the convenience of readers and are to be removed from the final version.

7.1 Lemma

(a) Under the assumptions of Theorem 3.1, as $n \to \infty$,

$$
\max_{a,b}\sum_{j=s}^{r}\left(e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0} w_{aj}w_{bj}^* - G^0_{ab}\right) = o_p(r) + O_p(r^{1/2}\log m), \quad 1 \le s \le r \le m.
$$

(b) Under the assumptions of Theorem 4.1, as $n \to \infty$,

$$
\begin{aligned}
(b1) \qquad & \max_{a,b}\sum_{j=s}^{r} e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0}\left(w_{aj}w_{bj}^* - A_a(\lambda_j)I_{\varepsilon j}A_b^*(\lambda_j)\right) \\
= \; & O_p(r^{1/3}(\log r)^{2/3} + \log r + r^{1/2}n^{-1/4}), \quad 1 \le s \le r \le m, \\
(b2) \qquad & \max_{a,b}\sum_{j=s}^{r}\left(e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0} w_{aj}w_{bj}^* - G^0_{ab}\right) \\
= \; & O_p(r^{\beta+1}n^{-\beta} + r^{1/3}(\log r)^{2/3} + \log r + r^{1/2}), \quad 1 \le s \le r \le m.
\end{aligned}
$$

7.2 Proof

Decompose the term inside the summation as $H_{1j} + H_{2j} + H_{3j}$, where

$$
\begin{aligned}
H_{1j} &= e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0}\left[w_{aj}w_{bj}^* - A_a(\lambda_j)I_{\varepsilon j}A_b^*(\lambda_j)\right] \\
H_{2j} &= e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0}\left[A_a(\lambda_j)I_{\varepsilon j}A_b^*(\lambda_j) - f_{ab}(\lambda_j)\right] \\
H_{3j} &= e^{i(\lambda_j-\pi)(d_a^0-d_b^0)/2}\lambda_j^{d_a^0+d_b^0} f_{ab}(\lambda_j) - G^0_{ab},
\end{aligned}
$$

where $A_a(\lambda_j)$ is the ath row of $A(\lambda_j) = \sum_{k=0}^{\infty} A_k e^{ik\lambda_j}$ and $A_b^*(\lambda_j)$ is the bth column of $A^*(\lambda_j)$. We prove part (a) first. Assumption 1 implies that, for any $\eta > 0$, n can be chosen such that

$$\max_{a,b} \left| e^{i(\lambda_j - \pi)(d_a^0 - d_b^0)/2} \lambda_j^{d_a^0 + d_b^0} f_{ab}(\lambda_j) - G_{ab}^0 \right| \le \eta, \quad j = 1, \ldots, m,$$

and $\max_{a,b} \sum_{j=s}^{r} |H_{3j}| = o(r)$ follows. For the contribution from H_{1j}, from the proof of Theorem 2 of Robinson (1995a) (also see Robinson (1995b) p. 1673) we have

$$\begin{cases} EI_j = f_j\{1 + O(j^{-1}\log(j+1))\}, \\ Ew_{aj}w_{\varepsilon j}^* = A_a(\lambda_j)/2\pi + O(j^{-1}\log(j+1)\lambda_j^{-d_a}), \qquad j = 1, \ldots, m. \\ EI_{\varepsilon j} = I_n/2\pi + O(j^{-1}\log(j+1)), \end{cases} \tag{29}$$

Rewrite H_{1j} as

$$e^{i(\lambda_j - \pi)(d_a^0 - d_b^0)/2} \lambda_j^{d_a^0 + d_b^0} \left\{ \left[w_{aj} - A_a(\lambda_j) w_{\varepsilon j} \right] w_{bj}^* + A_a(\lambda_j) w_{\varepsilon j} \left[w_{bj}^* - w_{\varepsilon j}^* A_b^*(\lambda_j) \right] \right\}. \tag{30}$$

The Cauchy-Schwartz inequality gives

$$E \left| \lambda_j^{d_a^0 + d_b^0} \left[w_{aj} - A_a(\lambda_j) w_{\varepsilon j} \right] w_{bj}^* \right| \le \left(E\lambda_j^{2d_a^0} |w_{aj} - A_a(\lambda_j) w_{\varepsilon j}|^2 \right)^{1/2} \left(E\lambda_j^{2d_b^0} w_{bj} w_{bj}^* \right)^{1/2}. \tag{31}$$

From (29), $A_a(\lambda_j)A_a^*(\lambda_j)/2\pi = f_{aa}(\lambda_j)$, and $\lambda_j^{2d_a^0} f_{aa}(\lambda_j) \sim G_{aa}^0$, we have

$$\begin{aligned} E\lambda_j^{2d_a^0} |w_{aj} - A_a(\lambda_j) w_{\varepsilon j}|^2 &= 2\lambda_j^{2d_a^0} f_{aa}(\lambda_j)\{1 + O(j^{-1}\log(j+1))\} \\ &\quad -2\lambda_j^{2d_a^0} f_{aa}(\lambda_j)\{1 + O(j^{-1}\log(j+1))\} \\ &= O(j^{-1}\log(j+1)). \end{aligned}$$

$E\lambda_j^{2d_b^0} w_{bj} w_{bj}^* = O(1)$ follows from (29), hence (31) is $O(j^{-1/2}\log(j+1))$. The second term of (30) is bounded in the same manner, and we obtain $\max_{a,b} \sum_{j=s}^{r} H_{1j} = O_p(r^{1/2}\log m)$.

For the contribution from H_{2j}, as in Lobato (1999, p.148) use $I_{\varepsilon j} = (2\pi n)^{-1}(\sum_{t=1}^{n} \varepsilon_t \varepsilon_t' + \sum\sum_{s \ne t} \varepsilon_s \varepsilon_t' e^{i(s-t)\lambda_j})$ to rewrite $\sum_{j=s}^{r} H_{2j}$ as

$$e^{i(\lambda_j - \pi)(d_a^0 - d_b^0)/2} \frac{1}{2\pi} \sum_{j=s}^{r} \lambda_j^{d_a^0 + d_b^0} A_a(\lambda_j) \frac{1}{n} \sum_{t=1}^{n} \left(\varepsilon_t \varepsilon_t' - I_q \right) A_b^*(\lambda_j) \tag{32}$$

$$+ e^{i(\lambda_j - \pi)(d_a^0 - d_b^0)/2} \frac{1}{2\pi} \sum_{j=s}^{r} \lambda_j^{d_a^0 + d_b^0} A_a(\lambda_j) \left(\frac{1}{n} \sum\sum_{s \ne t} \varepsilon_s \varepsilon_t' e^{i(s-t)\lambda_j} \right) A_b^*(\lambda_j). \tag{33}$$

(32) is $o_p(r)$ because

$$\begin{aligned} &\left| \sum_{j=s}^{r} \lambda_j^{d_a^0 + d_b^0} A_a(\lambda_j) \frac{1}{n} \sum_{t=1}^{n} \left(\varepsilon_t \varepsilon_t' - I_q \right) A_b^*(\lambda_j) \right| \\ = &\left| \operatorname{tr} \left(\frac{1}{n} \sum_{t=1}^{n} \left(\varepsilon_t \varepsilon_t' - I_q \right) \sum_{j=s}^{r} \lambda_j^{d_a^0 + d_b^0} A_b^*(\lambda_j) A_a(\lambda_j) \right) \right| \\ \le &\; C \left\| \frac{1}{n} \sum_{t=1}^{n} \left(\varepsilon_t \varepsilon_t' - I_q \right) \right\| \left\| \sum_{j=s}^{r} \lambda_j^{d_a^0 + d_b^0} \left\| A_b^*(\lambda_j) A_a(\lambda_j) \right\| \right\|, \end{aligned}$$

$n^{-1}\sum_1^n(\varepsilon_t\varepsilon_t' - I_q) \to_p 0$ from Theorem 1 of Heyde and Senata (1972), and

$$\|A_b^*(\lambda_j)A_a(\lambda_j)\| \leq \|A_b^*(\lambda_j)A_a(\lambda_j)\|_2 \leq \|A_b(\lambda_j)\|_2 \|A_a(\lambda_j)\|_2 = O(\lambda_j^{-d_a^0-d_b^0}).$$

by Cauchy-Schwartz inequality, where $||\cdot||_2$ denotes the Euclidean norm. For (33), note that (33) is bounded by

$$\frac{1}{2\pi}\sum_{j=s}^{r}\lambda_j^{d_a^0+d_b^0}A_a(\lambda_j)\left(\frac{1}{n}\sum\sum_{s\neq t}\varepsilon_s\varepsilon_t' e^{i(s-t)\lambda_j}\right)A_b^*(\lambda_j) = \sum\sum_{s\neq t}\varepsilon_t'\Xi_{t-s}\varepsilon_s,$$

where

$$\Xi_{t-s} = \frac{1}{2\pi n}\sum_{j=s}^{r}\lambda_j^{d_a^0+d_b^0}A_a(\lambda_j)e^{i(s-t)\lambda_j}A_b^*(\lambda_j).$$

$\sum\sum_{s\neq t}\varepsilon_t'\Xi_{t-s}\varepsilon_s$ has mean zero and variance $\sum\sum_{s\neq t}\text{vec}'\Xi_{t-s}\text{vec}'\Xi_{t-s}\text{vec}'\Xi_{t-s}' = O(r)$ in view of the arguments in Lobato (1999, p.148). Therefore, $\sum\sum_{s\neq t}\varepsilon_t'\Xi_{t-s}\varepsilon_s$ is $O_p(r^{1/2})$, giving part (a).

For part (b), (b1) holds because $\max_{a,b}\sum_{j=s}^{r}H_{1j} = O_p(r^{1/3}(\log r)^{2/3} + \log r + r^{1/2}n^{-1/4})$, which follows from applying the proof of (C.2) in Lobato (1999). For (b2), we have $\max_{a,b}\sum_{j=s}^{r}H_{2j} = O_p(r^{1/2})$ because (32)= $O_p(r^{1/2})$ since $n^{-1}\sum_{t=1}^{n}(\varepsilon_t\varepsilon_t' - I_q) = O_p(n^{-1/2})$ from Assumption 2′ and (33)= $O_p(r^{1/2})$ still holds. Assumption 1′ implies $\max_{a,b}\sum_{j=s}^{r}|H_{3j}| = O(r^{\beta+1}n^{-\beta})$, giving (b2). ■

7.3 Lemma (Shimotsu and Phillips, 2003)

For $\kappa \in (0,1)$, as $m \to \infty$,

$$(a) \qquad \sup_{-C\leq\gamma\leq C}\left|\frac{1}{m}\sum_{j=[\kappa m]}^{m}\left(\frac{j}{m}\right)^{\gamma} - \int_{\kappa}^{1}x^{\gamma}dx\right| = O\left(m^{-1}\right),$$

$$(b) \qquad \begin{array}{c}\sup_{-C\leq\gamma\leq C}|m^{-1}\sum_{j=[\kappa m]}^{m}(j/m)^{\gamma}| = O(1), \\ \liminf_{m\to\infty}\inf_{-C\leq\gamma\leq C}|m^{-1}\sum_{j=[\kappa m]}^{m}(j/m)^{\gamma}| > \varepsilon > 0.\end{array}$$

7.4 Lemma (Shimotsu and Phillips, 2003)

For $p \sim m/e$ as $m \to \infty$, $\varepsilon \in (0, 0.1)$, and $\Delta \in (0, 1/(2e))$, there exists $\bar{\kappa} \in (0, 1/4)$ such that, for sufficiently large m and all fixed $\kappa \in (0, \bar{\kappa})$,

$$(a) \inf_{-C\leq\gamma\leq-1+2\Delta}\frac{1}{m}\sum_{j=[\kappa m]}^{m}\left(\frac{j}{p}\right)^{\gamma} \geq 1+2\varepsilon, \quad (b) \inf_{1\leq\gamma\leq C}\frac{1}{m}\sum_{j=[\kappa m]}^{m}\left(\frac{j}{p}\right)^{\gamma} \geq 1+2\varepsilon.$$

7.5 Lemma

For $p \sim m/e$ as $m \to \infty$, $\varepsilon \in (0, 0.1)$, $\Delta \in (0, 1/(2e))$, and $\kappa \in (0, 1/4)$, we have, for sufficiently large m,

$$\inf_{-1+2\Delta\leq\gamma\leq 1}\frac{1}{m}\sum_{\kappa m}^{m}\left(\frac{j}{p}\right)^{\gamma} \geq 1 - \kappa^{2\Delta} + o(1).$$

7.6 Proof

It follows from Lemma 7.3 that

$$\frac{1}{m}\sum_{\kappa m}^{m}\left(\frac{j}{p}\right)^{\gamma}=\left(\frac{m}{p}\right)^{\gamma}\frac{1}{m}\sum_{\kappa m}^{m}\left(\frac{j}{m}\right)^{\gamma}=e^{\gamma}\int_{\kappa}^{1}x^{\gamma}dx+o(1)=\frac{e^{\gamma}(1-\kappa^{\gamma+1})}{\gamma+1}+o(1).$$

The stated result follows because $e^{\gamma}/(\gamma+1)\geq 1$ for $\gamma\in[-1+2\Delta,1]$. ■

7.7 Lemma

For $j,k=1,\ldots,m$ with $m=O(n)$, as $n\to\infty$,

(a) $\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\cos^2(s\lambda_j)=(1/4)n^2+o(n^2)$,
(b) $\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\cos(s\lambda_j)\cos(s\lambda_k)=O(n),\quad j\neq k$,
(c) $\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\sin^2(s\lambda_j)=(1/4)n^2+o(n^2)$,
(d) $\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\sin(s\lambda_j)\sin(s\lambda_k)=O(n),\quad j\neq k$,
(e) $\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\cos(s\lambda_j)\sin(s\lambda_k)=O(n^2(j+k)^{-1}+n^2|j-k|^{-1}),\quad j\neq k$.

7.8 Proof

Robinson (1995b, p. 1645) shows that $\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\cos^2(s\lambda_j)=(n-1)^2/4$, $\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\cos(s\lambda_j)=-n/2$, and $\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\cos(s\lambda_j)\cos(s\lambda_k)=-n/2$ for $j,k=1,\ldots,m<\frac{1}{2}n$, $j\neq k$, giving parts (a) and (b). Part (c) follows from

$$\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\sin^2(s\lambda_j)=\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\left\{1-\cos^2(s\lambda_j)\right\}=\frac{(n-1)(n-2)}{2}-\frac{(n-1)^2}{4}=\frac{n^2(1+o(1))}{4}.$$

Part (d) follows from

$$2\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\sin(s\lambda_j)\sin(s\lambda_k)=\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\{\cos(s\lambda_{j-k})-\cos(s\lambda_{j+k})\}=O(n).$$

For Part (e), first observe that $2\cos(s\lambda_j)\sin(s\lambda_k)=\sin(s\lambda_{j+k})-\sin(s\lambda_{j-k})$. Robinson (1995b, p. 1645) shows $\sum_{s=1}^{r}\sin(s\lambda)=[\cos(\lambda/2)-\cos((r+1/2)\lambda)]/(2\sin(\lambda/2))$ for $\lambda\neq 0$, mod(2π). The stated result follows from $1/(2\sin(\lambda/2))=O(|\lambda|^{-1})$. ■

References

[1] Bollerslev, T., and Wright, J. H. (2000), "Semiparametric Estimation of Long-Memory Volatility Dependencies: The Role of High-Frequency Data," *Journal of Econometrics* **98**, 81-106.

[2] Brunetti, C., and Gilbert, C. L. (2000). "Bivariate FIGARCH and Fractional Cointegration," *Journal of Empirical Finance* **7**, 509-530.

[3] Hannan, E. J. (1970). *Multiple Time Series.* Wiley.

[4] Henry, M. and P. Zaffaroni (2003). "The long range dependence paradigm for macroeconomics and finance." In *Theory and Applications of Long-Range Dependence*, P. Doukhan, G. Oppenheim and M. Taqqu (eds.), Birkhäuser.

[5] Heyde, C. C., and Senata, E. (1972), "Estimation Theory for Growth and Immigration Rates in a Multiplicative Process," *Journal of Applied Probability* **9**, 235-256.

[6] Horn, R. A. and C. R. Johnson (1985). *Matrix Analysis*, Cambridge University Press.

[7] Künsch, H. (1987). Statistical aspects of self-similar processes. In *Proceedings of the First World Congress of the Bernoulli Society* (Yu. Prokhorov and V. V. Sazanov, eds.) **1** 67-74. VNU Science Press, Utrecht.

[8] Lobato, I. N. (1999). "A semiparametric two-step estimator in a multivariate long memory model." *Journal of Econometrics* **90**: 129-53.

[9] Lobato, I. N. and P. M. Robinson (1998). "A nonparametric test for I(0)." *Review of Economic Studies* **65**: 475-495.

[10] Lobato, I. N. and C. Velasco (2000). "Long memory in stock-market trading volume." *Journal of Business & Economic Statistics* **18**: 410-427.

[11] Lütkepohl, H. (1996). *Handbook of Matrices*. Wiley.

[12] Robinson, P. M. (1995a). "Log-periodogram regression of time series with long range dependence." *Annals of Statistics* **23**: 1048-1072.

[13] Robinson, P. M. (1995b). "Gaussian semiparametric estimation of long range dependence." *Annals of Statistics* **23**: 1630-1661.

[14] Phillips, P. C. B. and K. Shimotsu (2003). Local Whittle estimation in nonstationary and unit root cases. Cowles Foundation Discussion Paper #1266, Yale University.

[15] Shimotsu, K. and P. C. B. Phillips (2003). Exact local Whittle estimation of fractional integration. Cowles Foundation Discussion Paper #1367, Yale University.