Semiparametric Estimation of Fractional Cointegrating Subspaces

Willa W. Chen [*] Clifford M. Hurvich [†]

November 19, 2004

Abstract:

We consider a common components model for multivariate fractional cointegration, in which the $s \geq 1$ components have different memory parameters. The cointegrating rank is allowed to exceed 1. The true cointegrating vectors can be decomposed into orthogonal fractional cointegrating subspaces such that vectors from distinct subspaces yield cointegrating errors with distinct memory parameters, denoted by d_k, for $k = 1, \ldots, s$. We estimate each cointegrating subspace separately using appropriate sets of eigenvectors of an averaged periodogram matrix of tapered, differenced observations. The averaging uses the first m Fourier frequencies, with m fixed. We will show that any vector in the k'th estimated cointegrating subspace is, with high probability, close to the k'th true cointegrating subspace, in the sense that the angle between the estimated cointegrating vector and the true cointegrating subspace converges in probability to zero. This angle is $O_p(n^{-\alpha_k})$, where n is the sample size and α_k is the shortest distance between the memory parameters corresponding to the given and adjacent subspaces. We show that the cointegrating residuals corresponding to an estimated cointegrating vector can be used to obtain a consistent and asymptotically normal estimate of the memory parameter for the given cointegrating subspace, using a univariate Gaussian semiparametric estimator with a bandwidth that tends to ∞ more slowly than n. We also show how these memory parameter estimates can be used to test for fractional cointegration and to consistently identify the cointegrating subspaces.

Keywords: Fractional cointegration, long memory, tapering, periodogram.

1 Introduction

Fractional cointegration has been the subject of much recent attention. See, for example, the work of Robinson (1994), Robinson and Marinucci (2001), Marinucci and Robinson (2002), Chen and Hurvich (2003a). All of these papers assume either that the observed series is bivariate or that the cointegrating rank is 1. Arguably the most interesting case from an econometric point of view is the situation where the series is multivariate and has cointegrating rank which may exceed 1. This situation was covered by

[*]Department of Statistics, Texas A&M University, College Station, Texas 77843, USA

[†]New York University, 44 W. 4'th Street, New York NY 10012, USA

The authors thank three referees for comments that substantially improved the paper, and they especially thank the Associate Editor for a very detailed reading of the paper and for providing extremely useful suggestions

Robinson and Yajima (2002), who considered methods of determining the cointegrating rank, and also by Chen and Hurvich (2003b), who focused on estimation of the space of cointegrating vectors.

Chen and Hurvich (2003b) studied the properties of eigenvectors of an averaged periodogram matrix of differenced, tapered observations, averaging over the first m Fourier frequencies, where m is held fixed as the sample size grows. They showed that the eigenvectors corresponding to the r smallest eigenvalues (where r is the cointegrating rank) lie close to the space of true cointegrating vectors with high probability. They also presented an empirical analysis of fractional cointegration in US interest rates for bonds of seven different maturities. They found evidence that the cointegrating rank was greater than one, and furthermore, that the memory parameter of the cointegrating errors may take on a variety of values that differ substantially if cointegrating vectors corresponding to substantially different eigenvalues are used. This last finding, while of apparent interest from an econometric point of view, could not be explained directly from the theoretical results presented in Chen and Hurvich (2003b), since they did not attempt in their theory to separate the space of cointegrating vectors into subspaces yielding different memory parameters.

The goals of the present paper are to exhibit a model that allows us to highlight these subspaces, to show that the subspaces and their corresponding memory parameters can be estimated individually, and to show how to use the residual-based Gaussian semiparametric estimates of the memory parameters to consistently identify the cointegrating subspaces and to test for fractional cointegration. By contrast, Chen and Hurvich (2003b) did not consider either testing for cointegration or estimation of the degree of cointegration.

We first present in Section 2 a semiparametric common components model in which the components have different memory parameters, while the entries of the observed multivariate series have just one common memory parameter. Next, we show that the space of cointegrating vectors can be decomposed into a direct sum of orthogonal *cointegrating subspaces* such that vectors from distinct subspaces yield cointegrating errors with distinct memory parameters.

We show in Section 5 that each of these cointegrating subspaces can be separately estimated using sets of eigenvectors of the averaged periodogram matrix. Since m is held fixed, we are able to obtain a rate of convergence for the estimated cointegrating vectors that depends only on the difference between the memory parameters in the given and adjacent subspaces, and is not hampered by the rate of increase of m as in other related work (cf. Robinson and Marinucci 2001, in the bivariate case).

To each true cointegrating subspace, there corresponds an estimated cointegrating subspace spanned by an orthonormal set of eigenvectors of the averaged periodogram matrix, where membership in the set is determined by a partitioning of the sorted observed eigenvalues into contiguous groups, of sizes that match the dimensions of the corresponding true cointegrating subspaces. We show in Section 4 that the eigenvalues for the k'th estimated cointegrating subspace are $O_p(n^{2d_k})$, where n is the sample size, and d_k is the memory parameter of the cointegrating error for the k'th true cointegrating subspace. This result, and further refinements of it, plays a key role in our subsequent theory.

We will show in Theorem 1 that any vector in the k'th estimated cointegrating subspace is, with high probability, close to the k'th true cointegrating subspace, in the sense that the norm of the sine of the angle between these two subspaces converges in probability to zero. The norm of the sine of this angle is $O_p(n^{-\alpha_k})$, where α_k is the shortest distance between the memory parameters corresponding to the given and adjacent subspaces. This implies that the sine of the angle between any vector in the k'th estimated cointegrating subspace and the k'th true cointegrating subspace is $O_p(n^{-\alpha_k})$. (We provide more details on the notion of the sine of the angle between subspaces, and also the sine of the angle between a vector and a subspace, in Section 5). This convergence rate, which improves as α_k increases, is at least as fast as the rates obtained for existing semiparametric estimators of cointegrating vectors in the bivariate case (see, *eg.*, Robinson and Marinucci 2001 and the discussion in Chen and Hurvich (2003a)). Furthermore, we show in Lemma 29 that the normalized eigenvectors of the averaged periodogram matrix converge in distribution to random vectors that lie in the corresponding cointegrating subspace.

We then show in Section 6 that the cointegrating residuals corresponding to an estimated cointegrating vector can be used to obtain a consistent and asymptotically normal estimate of the memory parameter for the given cointegrating subspace, using a univariate Gaussian semiparametric estimator with a bandwidth that tends to ∞ more slowly than n.

In Section 7, we propose and justify a test for fractional cointegration in the current context. In Section 8 we provide a procedure for consistently identifying the cointegrating subspaces, i.e., for determining the number of subspaces and their dimensions.

2 A Fractional Common Components Model

Suppose that the original data are a $q \times 1$ time series such that the $p-1^{th}$ differences $\{y_t\}$ are weakly stationary with a common memory parameter $d_0 \in (-p+1/2, 1/2)$, where $p \geq 1$ is a fixed integer. The use of $p-1^{th}$ differences converts any additive polynomial trend of order $p-1$ in the original series into an additive constant. The value of this constant is irrelevant for our purposes since the estimators considered here are functions of the discrete Fourier transform at nonzero Fourier frequencies. We can therefore take the mean of $\{y_t\}$ to be zero, without loss of generality, and our estimators are invariant to polynomial trends of order $p-1$ in the original series.

In order to guarantee that the cointegrating relationships in the stochastic component of the levels are preserved in the differences, we apply a taper to the differences, that is, we multiply the differences by a sequence of constants prior to Fourier transformation. This prevents detrimental leakage effects due to potential overdifferencing, and allows us to obtain uniform results over a wide range of memory parameters. A convenient family of tapers for use on the differences, and which we will use here, was given in Hurvich and Chen (2000). The exact form of the taper is given below.

The fractional common components model for the $(q \times 1)$ series $\{y_t\}$ with cointegrating rank r $(1 \leq r < q)$, and s cointegrating subspaces $(1 \leq s \leq r)$, is given by

$$y_t = \mathbf{A}_0 u_t^{(0)} + \mathbf{A}_1 u_t^{(1)} + \cdots + \mathbf{A}_s u_t^{(s)} \quad , \tag{1}$$

where $\mathbf{A}_k$ $(0 \leq k \leq s)$ are $q \times a_k$ full-rank matrices with $a_0 = q - r$ and $a_1 + \cdots + a_s = r$ such that all columns of $\mathbf{A}_0, \ldots, \mathbf{A}_s$ are linearly independent, $\{u_t^{(k)}\}$ $k = 0, \ldots, s$, are $(a_k \times 1)$ processes with memory parameters $\{d_k\}_{k=0}^{s}$ with $-p + 1/2 < d_s < \cdots < d_0 < 1/2$. Equation (1) can be written as

$$y_t = \mathbf{A} z_t, \tag{2}$$

where $z_t = \mathrm{vec}\left(u_t^{(0)}, \ldots, u_t^{(s)}\right)$ and $\mathbf{A} = \begin{bmatrix} \mathbf{A}_0 & \cdots & \mathbf{A}_s \end{bmatrix}$. We will make additional assumptions on $\{z_t\}$ in Section 3. These assumptions guarantee that $\{z_t\}$ is not cointegrated. The methodology presented in this paper does not require either r or s to be known.

Remark 1 *Our assumption that all entries of* $\{y_t\}$ *have memory parameter* d *implies that all rows of* A_0 *are nonzero. The model (1), without the assumption that all entries of* $\{y_t\}$ *have a common memory parameter, could also be entertained, though we do not pursue this here, and would then include the model considered by Robinson and Yajima (2002).*

Next, we exhibit the cointegrating subspaces. For any matrix $\mathbf{A}$, let $\mathcal{M}(\mathbf{A})$ denote the column space of $\mathbf{A}$, and let $\mathcal{M}^{\perp}(\mathbf{A})$ denote the orthogonal complement of $\mathbf{A}$. Note that, for $k = 1, \ldots, s$,

$$\mathcal{M}^{\perp}\left(\mathbf{A}_0, \ldots, \mathbf{A}_k\right) \subset \mathcal{M}^{\perp}\left(\mathbf{A}_0, \ldots, \mathbf{A}_{k-1}\right).$$

Let $\mathcal{B}_0 = \mathcal{M}\left(\mathbf{A}_0\right)$, and $\mathcal{B}_k$, $k = 1, \ldots, s$, be the subspace such that

$$\mathcal{M}^{\perp}\left(\mathbf{A}_0, \ldots, \mathbf{A}_{k-1}\right) = \mathcal{M}^{\perp}\left(\mathbf{A}_0, \ldots, \mathbf{A}_k\right) \oplus \mathcal{B}_k,$$

and $\mathcal{B}_k \perp \mathcal{M}^{\perp}\left(\mathbf{A}_0, \ldots, \mathbf{A}_k\right)$. Hence a vector $\beta \in \mathcal{B}_k$, $k \in \{1, \ldots, s\}$, satisfies $\beta' \mathbf{A}_\ell = 0$, $\ell = 0, \ldots, k-1$ and $\beta' \mathbf{A}_k \neq 0$. Also $\mathcal{B}_j \perp \mathcal{B}_k$ for $j \neq k$, $(j,k) \in \{0, \ldots, s\}$, and

$$\mathbb{R}^q = \mathcal{B}_0 \oplus \mathcal{B}_1 \oplus \cdots \oplus \mathcal{B}_s \quad . \tag{3}$$

It can be seen from (1) and the preceding discussion that any nonzero vector $\beta \in \mathcal{B}_k$ with $k \in \{1, \ldots, s\}$ produces a cointegrating error series $\{\beta' y_t\}$ with memory parameter d_k. Thus, $\mathcal{B}_1, \ldots, \mathcal{B}_s$ are the cointegrating subspaces. The space $\mathcal{B}_0$, on the other hand, is the space spanned by the non-cointegrating vectors in $\mathbb{R}^q$. Equation (3) shows that $\mathbb{R}^q$ may be written as a direct sum of the space of non-cointegrating vectors and the space of cointegrating vectors, and that the latter space may be further decomposed into a direct sum of cointegrating subspaces.

3 Assumptions

Here, we specify a linear model for the series $z_t = \operatorname{vec}\left(u_t^{(0)}, \ldots, u_t^{(s)}\right)$. As stated in the previous section, we assume that $\{u_t^{(k)}\}$ $k = 0, \ldots, s$, are $(a_k \times 1)$ processes with memory parameters $\{d_k\}_{k=0}^{s}$ with $-p + 1/2 < d_s < \cdots < d_0 < 1/2$. Define $N_0 = \{1, \ldots, a_0\}$ and $N_k = \{(a_0 + \cdots + a_{k-1}) + 1, \ldots, (a_0 + \cdots + a_k)\}$ for $k = 1, \ldots, s$.

Let ψ_k be a sequence of $q \times q$ matrices such that

$$\psi_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\omega} \mathbf{\Psi}\left(\omega\right) d\omega \quad ,$$

where for each $\omega \in [-\pi, \pi]$, $\mathbf{\Psi}(\omega)$ is a complex-valued matrix such that $\mathbf{\Psi}\left(-\omega\right) = \overline{\mathbf{\Psi}\left(\omega\right)}$ and ψ_0 is an identity matrix.

Define the $q \times 1$ vector process $\{z_t\}$ as

$$z_t = \sum_{k=-\infty}^{\infty} \psi_k \varepsilon_{t-k} \quad , \tag{4}$$

where $\{\varepsilon_t = (\varepsilon_{t,1}, \dots, \varepsilon_{t,q})'\} \sim iid\,(\mathbf{0}, 2\pi\mathbf{\Sigma})$, $\mathbf{\Sigma}$ is a symmetric positive definite matrix with entries σ_{ab}, $a, b \in \{1, \dots, q\}$ and $E\,\|\varepsilon_t\|^4 < \infty$, where $\|\cdot\|$ denotes the Euclidean norm. The spectral density matrix of $\{z_t\}$ is

$$\mathbf{f}(\omega) = \mathbf{\Psi}(\omega)\,\mathbf{\Sigma}\mathbf{\Psi}^*(\omega) \quad , \quad \omega \in [-\pi, \pi] \quad ,$$

where the superscript $*$ denotes conjugate transposition. We further assume that for $\omega \in [-\pi, \pi]$, the (a, b)'th entry of $\mathbf{\Psi}(\omega)$ is given by

$$\Psi_{ab}(\omega) = \left(1 - e^{-i\omega}\right)^{-d_{ab}} \tau_{ab}(\omega)\, e^{i\phi_{ab}(\omega)} \tag{5}$$

where $d_{aa} = d_k$ for $a \in N_k$, and $d_{ab} \leq \min(d_k, d_h)$ for $a \in N_k$, $b \in N_h$, $b \neq a$ $(k, h = 0, \dots, s)$, and for all $a, b \in \{1, \dots, q\}$, $\tau_{ab}(\cdot)$ are positive even real-valued functions, $\phi_{ab}(\cdot)$ are odd real-valued functions, all continuously differentiable in an interval containing zero. It follows from (5) that the first derivatives of $\Psi_{ab}(\omega)$ satisfy

$$\Psi'_{ab}(\omega) = O\left(\left|\Psi_{aa}(\omega)\Psi_{bb}(\omega)\right|^{1/2} |\omega|^{-1}\right) \quad . \tag{6}$$

In keeping with (5), we assume that we can write the spectral density matrix of $\{z_t\}$ as

$$\mathbf{f}(\omega) = \mathbf{\Upsilon}(\omega)\mathbf{f}^{\dagger}(\omega)\,\mathbf{\Upsilon}^*(\omega) \quad , \tag{7}$$

where $\mathbf{\Upsilon}(\omega) = \mathrm{diag}\left\{\left(1 - e^{-i\omega}\right)^{-d_0}, \dots, \left(1 - e^{-i\omega}\right)^{-d_0}, \dots, \left(1 - e^{-i\omega}\right)^{-d_s}, \dots, \left(1 - e^{-i\omega}\right)^{-d_s}\right\}$, i.e, the a'th diagonal entry is $\left(1 - e^{-i\omega}\right)^{-d_k}$ for all $a \in N_k$ $(k = 0, \dots, s)$, and

$$\mathbf{f}^{\dagger}(\omega) = \mathbf{\Psi}^{\dagger *}(\omega)\,\Sigma\mathbf{\Psi}^{\dagger}(\omega)\,, \tag{8}$$

is positive definite, Hermitian, continuous at zero frequency, and therefore real-valued at zero frequency. Thus, $\{z_t\}$ is not fractionally cointegrated. (See Robinson and Marinucci, 1998).

4 The Averaged Periodogram Matrix and its Eigenvalues

For any vector sequence of observations $\{\xi_t\}_{t=1}^n$, define the tapered discrete Fourier transform by

$$J_\xi(\omega_j) = \frac{1}{\sqrt{2\pi \sum \left|h_t^{p-1}\right|^2}} \sum_{t=1}^{n} h_t^{p-1} \xi_t e^{i\omega_j t} \quad ,$$

where $\omega_j = 2\pi j/n$ is the j'th Fourier frequency, and $\{h_t\}$ is the complex-valued taper of Hurvich and Chen (2000),

$$h_t = 0.5\left(1 - e^{i2\pi t/n}\right) \quad , \quad t = 1, \dots, n \quad .$$

Note that $p = 1$ yields the no-tapering case. Next, define the tapered cross-periodogram matrix of two vector sequences $\{\xi_t\}_{t=1}^n$ and $\{\zeta_t\}_{t=1}^n$ by

$$I_{\xi\zeta}(\omega_j) = J_\xi(\omega_j) J_\zeta^*(\omega_j) \quad .$$

We will work with the (real part of the) averaged periodogram matrix of a sample of n observations $\{y_t\}_{t=1}^n$,

$$I_m = \sum_{j=1}^{m} \mathrm{Re}\left\{I_{yy}(\omega_j)\right\} \quad ,$$

where m is a fixed positive integer, $m > q+3$.

Denote $I_m(\xi_t, \zeta_t) = \sum_{j=1}^{m} \mathrm{Re}\{I_{\xi\zeta}(\omega_j)\}$. We first focus on the asymptotic distribution of $I_m(z_t, z_t)$. Define the function (for $x \in \mathbb{R}$)

$$\Delta_p(x) = \begin{pmatrix} 2p-2 \\ p-1 \end{pmatrix}^{-1/2} \sum_{k=0}^{p-1} \begin{pmatrix} p-1 \\ k \end{pmatrix} (-1)^k \Delta(x + 2\pi k) \quad ,$$

where

$$\Delta(x) = \frac{1}{\sqrt{2\pi}} \frac{e^{ix}-1}{ix} \quad .$$

Now, define

$$\upsilon_j(x) = \frac{1}{2}\left[\overline{\Delta_p(-x+2\pi j)} + \Delta_p(x+2\pi j)\right],$$
$$\nu_j(x) = \frac{i}{2}\left[\overline{\Delta_p(-x+2\pi j)} - \Delta_p(x+2\pi j)\right].$$

Define the Hermitian positive definite $q \times q$ matrix-valued measure $\mathbf{G}_0$ on $\mathbb{R}$ by

$$\mathbf{G}_0(dx) = \mathbf{\Pi}(x)\, \mathbf{f}^{\dagger}(0) \mathbf{\Pi}^*(x)\, dx \tag{9}$$

for $x > 0$ and $\mathbf{G}_0(-dx) = \overline{\mathbf{G}_0(dx)}$ where

$$\mathbf{\Pi}(x) = \mathrm{diag}\left(e^{-i\pi d_0/2}|x|^{-d_0}, \ldots, e^{-i\pi d_0/2}|x|^{-d_0}, \ldots, e^{-i\pi d_s/2}|x|^{-d_s}, \ldots, \ldots, e^{-i\pi d_s/2}|x|^{-d_s}\right) \quad .$$

Let $\mathbf{U}_n$ and $\mathbf{V}_n$ be $q \times m$ matrices given by

$$\mathbf{U}_n = \mathbf{d}_n^{-1} \mathrm{Re}(J_{z,1}, \ldots, J_{z,m}) \quad \text{and} \quad \mathbf{V}_n = \mathbf{d}_n^{-1} \mathrm{Im}(J_{z,1}, \ldots, J_{z,m}). \tag{10}$$

Lemma 1 *Let $\mathbf{d}_n$ be a $(q \times q)$ diagonal matrix with i^{th} diagonal entry $n^{d_k}, i \in N_k$, $k = 0, \ldots, s$ and $\mathbf{Q}_n = \mathbf{d}_n^{-1} I_m(z_t, z_t)\, \mathbf{d}_n^{-1} = (\mathbf{U}_n, \mathbf{V}_n)(\mathbf{U}_n, \mathbf{V}_n)'$. If $m \geq q$, then*

$$\mathbf{Q}_n \xrightarrow{D} \mathbf{U}\mathbf{U}' + \mathbf{V}\mathbf{V}',$$

where $\mathbf{U} = (U_1, \ldots, U_m)$ and $\mathbf{V} = (V_1, \ldots, V_m)$, U_j, V_k are $q \times 1$ vectors, and $\mathrm{vec}(\mathbf{U}, \mathbf{V})$ is a $2mq$-variate normal random variable with zero mean, and covariance matrix $\mathbf{\Xi}$ determined by

$$\mathbb{E}(U_j U_k') = \int_{\mathbb{R}} \upsilon_j(x) \overline{\upsilon_k(x)} \mathbf{G}_0(dx),$$
$$\mathbb{E}(V_j V_k') = \int_{\mathbb{R}} \nu_j(x) \overline{\nu_k(x)} \mathbf{G}_0(dx),$$
$$\mathbb{E}(U_j V_k') = \int_{\mathbb{R}} \upsilon_j(x) \overline{\nu_k(x)} \mathbf{G}_0(dx).$$

Furthermore, $\mathbf{U}\mathbf{U}' + \mathbf{V}\mathbf{V}'$ is positive definite and has distinct eigenvalues with probability 1.

Proof. The proof is identical to the proof of Lemma 1, Corollary 1 and 2 of Chen and Hurvich (2003b).□

We next derive upper and lower bounds for the eigenvalues of $I_m(y_t, y_t)$. We will use the notation $\lambda_j(\cdot)$ for the j'th eigenvalue of a given Hermitian matrix, $\lambda_j(\cdot) \geq \lambda_{j+1}(\cdot)$. Also we let $\lambda_j = \lambda_j(I_m(y_t, y_t))$. We have the following lemma.

Lemma 2 $\lambda_j = O_p\left(n^{2d_k}\right)$, *for* $j \in N_k$, $k = 0, \ldots, s$.

In the case $k \geq 1$, the upper bound in Lemma 2 strengthens Lemma 4 of Chen and Hurvich (2003b).

Lemma 3 *Let* $j_k^* = \max\{j : j \in N_k\}$ *and* $\mathbf{Q}_n^{(k)}$ *be the leading* $j_k^* \times j_k^*$ *principal submatrix of* $\mathbf{Q}_n$, *for* $k = 0, \ldots, s$. *Then*

$$n^{-2du_k} \lambda_{j_k^*} \geq c_k \lambda_{j_k^*}\left(\mathbf{Q}_n^{(k)}\right) \xrightarrow{D} \eta_{j_k^*}^{(k)},$$

where $c_k > 0$ *and* $\eta_{j_k^*}^{(k)}$ *is a random variable that has no mass at* 0.

5 Estimation of the Cointegrating Subspaces

Let $\mathbf{X}(\cdot) = \begin{bmatrix} \chi_1(\cdot) & \ldots & \chi_q(\cdot) \end{bmatrix}$, an orthogonal matrix such that $\chi_j(\cdot)$ is the eigenvector corresponding to the j'th largest eigenvalue $\lambda_j(\cdot)$ of a given symmetric $q \times q$ matrix, and let $\mathbf{X}_k(\cdot)$ be a matrix with columns $\chi_j(\cdot), j \in N_k$, for $k = 0, \ldots, s$. Also we let $\chi_j = \chi_j(I_m(y_t, y_t))$, $\mathbf{X} = \mathbf{X}(I_m(y_t, y_t))$ and $\mathbf{X}_k = \mathbf{X}_k(I_m(y_t, y_t))$. For $k = 0, 1, \ldots, s$, let $\mathbf{B}_k$ be a $q \times a_k$ matrix with orthonormal columns such that $\mathcal{M}(\mathbf{B}_k) = \mathcal{B}_k$ and let $\mathbf{B} = \begin{bmatrix} \mathbf{B}_0 & \cdots & \mathbf{B}_s \end{bmatrix}$. Since $\mathbf{B}'\mathbf{B} = \mathbf{I}$, it follows that for any $q \times q$ matrix $\mathbf{P}$, $\mathbf{B}'\mathbf{P}\mathbf{B}$ is similar to $\mathbf{P}$, i.e., $\lambda_j(\mathbf{P}) = \lambda_j(\mathbf{B}'\mathbf{P}\mathbf{B})$ and $\chi_j(\mathbf{P}) = \mathbf{B}'\chi_j(\mathbf{B}'\mathbf{P}\mathbf{B})$.

Define

$$\mathbf{\Phi} = \mathbf{B}' I_m(y_t, y_t) \mathbf{B} \quad ,$$

and partition $\mathbf{\Phi}$ into $(s+1)^2$ blocks, such that the (k, ℓ) block $\mathbf{\Phi}_{k\ell}$ has dimension $(a_k \times a_\ell)$, for $k, \ell = 0, \ldots, s$. Define $\mathbf{\Phi}_D = \text{diag}[\mathbf{\Phi}_{00}, \ldots, \mathbf{\Phi}_{ss}]$, and $\Delta\mathbf{\Phi} = \mathbf{\Phi} - \mathbf{\Phi}_D$, so that

$$\mathbf{\Phi} = \mathbf{\Phi}_D + \Delta\mathbf{\Phi} \quad .$$

We have

$$I_m(y_t, y_t) = \mathbf{B}\mathbf{\Phi}\mathbf{B}' = \mathbf{B}\mathbf{\Phi}_D\mathbf{B}' + \mathbf{B}\Delta\mathbf{\Phi}\mathbf{B}' =: \mathbf{H} + \Delta\mathbf{H} \quad ,$$

so we can think of $I_m(y_t, y_t)$ as a perturbed version of $\mathbf{H}$. Using results by Barlow and Slapničar (2000) on perturbation theory for eigenvalues and eigenvectors of nonrandom Hermitian matrices, we will show in Lemma 4 that the k'th estimated cointegrating subspace $\mathcal{M}(\mathbf{X}_k)$ is close to $\mathcal{M}(\mathbf{X}_k(\mathbf{H}))$ in the sense that the norm of the sine of the angle between the two subspaces converges to 0 in probability.

Let $\Theta(\cdot, \cdot)$ denote the matrix of canonical angles between two subspaces of the same dimension (see, e.g., Stewart and Sun 1990, p.43). The notion of the sine of the angle between two subspaces of the same dimension is given in Davis and Kahan (1970). For simplicity, suppose that $\mathbf{S}$ and $\mathbf{T}$ are both real $q \times a$ matrices $(q > a)$ with orthonormal columns. Then the orthogonal projector into $\mathcal{M}(\mathbf{T})$ is given by $\mathbf{T}\mathbf{T}'$, and the projector into the orthogonal complement $\mathcal{M}^{\perp}(\mathbf{T})$ of $\mathcal{M}(\mathbf{T})$ is given by $\mathbf{I}-\mathbf{TT}'$, where $\mathbf{I}$ is a $q\times q$ identity matrix. The sine of the angle between $\mathcal{M}(\mathbf{S})$ and $\mathcal{M}(\mathbf{T})$ is an $a\times a$ matrix defined in Davis and Kahan (1970), and denoted by $\sin\Theta(\mathcal{M}(\mathbf{S}),\mathcal{M}(\mathbf{T}))$. It follows from Davis and Kahan (1970, page 10) that $\|\sin\Theta(\mathcal{M}(\mathbf{S}),\mathcal{M}(\mathbf{T}))\|_F=\|(\mathbf{I}-\mathbf{TT}')\mathbf{SS}'\|_F$ where $\|\cdot\|_F$ is the Frobenius norm. It follows from Stewart and Sun (1990, Corollary 5.4, p. 43) that

$$\|\sin\Theta(\mathcal{M}(\mathbf{S}),\mathcal{M}(\mathbf{T}))\|_F=\left\|\left(\mathbf{T}^{\perp}\right)'\mathbf{S}\right\|_F \tag{11}$$

where $\mathbf{T}^{\perp}$ is a matrix with orthonormal columns spanning $\mathcal{M}^{\perp}(\mathbf{T})$, so that $\left\|\left(\mathbf{T}^{\perp}\right)'\mathbf{S}\right\|_F$ is the square root of the sum of the squared lengths of the residuals from the orthogonal projections of the columns of $\mathbf{S}$ on the space $\mathcal{M}(\mathbf{T})$.

For any nonzero vector $x\in\mathcal{M}(\mathbf{S})$, the sine of the angle between x and the subspace $\mathcal{M}(\mathbf{T})$ is a real number defined as

$$\sin\theta(x,\mathcal{M}(\mathbf{T}))=\frac{\|(\mathbf{I}-\mathbf{TT}')\,x\|}{\|x\|}.$$

See Wedin (1983, p. 274). It then follows from (11) that

$$\max_{x\in\mathcal{M}(\mathbf{S})}|\sin\theta(x,\mathcal{M}(\mathbf{T}))|\le\left\|\left(\mathbf{T}^{\perp}\right)'\mathbf{S}\right\|_F.$$

In Lemma 5, we show that under the additional assumption that the process is Gaussian, $\mathcal{M}(\mathbf{X}_k(\mathbf{H}))$ is equal to $\mathcal{B}_k$ with probability approaching one, for $k=0,\ldots,s$. Lemmas 4 and 5 taken together imply our Theorem 1, stating that if the process is Gaussian, the k'th estimated cointegrating subspace $\mathcal{M}(\mathbf{X}_k)$ is close to the corresponding true cointegrating subspace $\mathcal{B}_k$, in the sense that $\|\sin\Theta\{\mathcal{M}(\mathbf{X}_k),\mathcal{B}_k\}\|_F=O_p(n^{-\alpha_k})$, where α_k is the shortest distance between the memory parameters corresponding to the given and adjacent subspaces, i.e.,

$$\alpha_k=\begin{cases} d_0-d_1 & k=0,\\ \min\{(d_{k-1}-d_k),(d_k-d_{k+1})\} & k=1,\ldots,s-1\\ d_{s-1}-d_s & k=s\end{cases}.$$

Lemma 4 *The sine of the angle between $\mathcal{M}(\mathbf{X}_k)$ and $\mathcal{M}(\mathbf{X}_k(\mathbf{H}))$ satisfies*

$$\|\sin\Theta\{\mathcal{M}(\mathbf{X}_k(\mathbf{H})),\mathcal{M}(\mathbf{X}_k)\}\|_F=O_p\left(n^{-\alpha_k}\right).$$

The following Gaussianity assumption is sufficient for obtaining a rate at which $P(\mathcal{M}(\mathbf{X}_k(\mathbf{H}))\neq\mathcal{B}_k)$ converges to zero. More specifically, the assumption allows us to bound the inverse second moment of eigenvalues of $\mathbf{Q}_n$. We believe that such bounds, and therefore Lemma 5, hold without the Gaussianity assumption, but we will not pursue this here.

Assumption 1 *The process $\{\varepsilon_t\}$ in (4) is Gaussian.*

Lemma 5 *Under Assumption 1, $P(\mathcal{M}(\mathbf{X}_k(\mathbf{H}))\neq\mathcal{B}_k)=O\left(n^{-2\alpha_k}\right)$, $k=0,\ldots,s$.*

The following theorem is a corollary of Lemmas 4 and 5.

Theorem 1 *Under Assumption 1,*

$$\left\| \sin\Theta\left\{\mathcal{M}\left(\mathbf{X}_k\right), \mathcal{B}_k\right\}\right\|_F = O_p(n^{-\alpha_k}) \quad , k = 0, \ldots, s.$$

6 Estimation of the Memory Parameters Using Cointegrating Residuals

Let b be any $q \times 1$ vector in $\mathcal{M}\left(\mathbf{X}_k\right)$ with length one, where $k \in \{0, \ldots, s\}$ is fixed but not necessarily known. For example, if b is a unit eigenvector of the averaged periodogram matrix corresponding to one of the sorted eigenvalues of this matrix, then there exists some $k \in \{0, \ldots, s\}$ such that $b \in \mathcal{M}\left(\mathbf{X}_k\right)$, though k is unknown since it depends on the unknown s and $a_0, \ldots, a_s$. We then use this vector b to construct the residual process $\{v_t\}$, where

$$v_t := b' y_t = b' \mathbf{A}_0 u_t^{(0)} + b' \mathbf{A}_1 u_t^{(1)} + \cdots + b' \mathbf{A}_k u_t^{(k)} + \cdots + b' \mathbf{A}_s u_t^{(s)}. \tag{12}$$

The periodogram of $\{v_t\}$ is

$$I_{vv}\left(\omega_j\right) = b' \mathbf{A} I_{zz}\left(\omega_j\right) \mathbf{A}' b.$$

We consider the Gaussian semiparametric estimator (GSE; see Kunsch, 1987, Robinson, 1995b) for d_k based on $\{v_t\}$,

$$\hat{d}_k = \arg\min_{d \in \Theta} R\left(d\right) = \log \widehat{G}\left(d\right) - 2d\left(\frac{1}{m_n}\sum_{j=1}^{m_n} \log \omega_{\tilde{j}}\right), \tag{13}$$

where $\Theta = \left[\Delta_1, \Delta_2\right]$, $-p + 0.5 < \Delta_1 < \Delta_2 < 0.5$, $\omega_{\tilde{j}} = 2\pi\tilde{j}/n$, $\tilde{j} = j + (p-1)/2$, and

$$\widehat{G}\left(d\right) = \frac{1}{m_n}\sum_{j=1}^{m_n} \frac{I_{vv}(\omega_j)}{\omega_{\tilde{j}}^{-2d}} = \frac{1}{m_n}\sum_{j=1}^{m_n} \frac{b' \mathbf{A} I_{zz}(\omega_j) \mathbf{A}' b}{\omega_{\tilde{j}}^{-2d}} \quad .$$

Here, we use slightly shifted Fourier frequencies $\omega_{\tilde{j}}$ to parallel corresponding shifts inherent in our tapering scheme and thereby reduce finite-sample bias, as was also done in Hurvich and Chen (2000).

The two theorems below establish the consistency and the limiting distribution of the $\hat{d}_k$, under some additional conditions on the transfer function , $a^*(\omega) = \tau_{ab}\left(\omega\right) e^{i\phi_{ab}(\omega)}$; see (5). Following Hurvich *et al* (2002), we define a smoothness class for transfer functions as follows. For $\mu > 1$ and $1 < \rho \leq 2$, let $\mathcal{L}^*\left(\mu, \rho\right)$ be the set of continuously differentiable functions u on $[-\pi, \pi]$ such that for all x, y with $|x| \in (0, \pi]$, $|y| \in (0, \pi]$,

$$\frac{\max_{0 \leq z \leq \pi} |u(z)|}{\min_{0 \leq z \leq \pi} |u(z)|} \leq \mu,$$

$$\frac{|u(x) - u(y)|}{\min_{0 \leq z \leq \pi} |u(z)|} \leq \mu \frac{|y - x|}{\min(|x|, |y|)},$$

and

$$\frac{|u'(x) - u'(y)|}{\min_{0 \leq z \leq \pi} |u(z)|} \leq \mu \frac{|y - x|^{(\rho - 1)}}{[\min(|x|, |y|)]^{\rho}}.$$

It follows from the discussion in Hurvich, *et al* (2002) that if a^* is the transfer function of a stationary and invertible autoregressive moving average process, or of a stationary and invertible fractional Gaussian noise, then $a^* \in \mathcal{L}^*(\mu, \rho)$ for some μ, with $\rho = 2$.

We now state an assumption on a^*.

Assumption 2 *$a^* \in \mathcal{L}^*(\mu, \rho)$ for some $\mu > 1$, and some $\rho \in (1, 2]$.*

Note that this assumption is global in that it pertains to the behavior of a^* at all frequencies. By contrast, our estimation of the $\hat{d}_k$ is based on frequencies in a shrinking neighborhood around zero. It seems plausible, then, that a local version of Assumption 2 would suffice for our purposes, though we do not pursue this here.

The following standard assumption is needed to establish the consistency of $\hat{d}_k$.

Assumption 3a. As $n \to \infty$,

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

Theorem 2 *Under Assumptions 1, 2 and 3a, for $k \in \{0, \ldots, s\}$, $\hat{d}_k \xrightarrow{p} d_k$.*

The next assumption is used for establishing the asymptotic normality of $m_n^{1/2}\left(\hat{d}_k - d_k\right)$, for the particular fixed value of k under consideration.

Assumption 3b. (i) If $k \in \{1, \ldots, s\}$, $d_{k-1} - d_k > 1/2$. (ii) If $k \in \{0, \ldots, s-1\}$, as $n \to \infty$,

$$\frac{1}{m_n} + \frac{m_n^{1+2(d_k - d_{k+1})} \log^2 m_n}{n^{2(d_k - d_{k+1})}} \to 0.$$

Note that part (i) is vacuous is $k = 0$, and part (ii) is vacuous if $k = s$. Assumption 3b may be compared with the assumptions in Theorems 2 and 4 of Velasco (2003), which he required for residual-based estimators of the memory parameters of a bivariate fractionally cointegrated system.

To present the asymptotic variance of $\hat{d}_k$, we define

$$\Phi_p = \frac{\Gamma(4p-3)\Gamma^4(p)}{\Gamma^4(2p-1)} \quad .$$

Theorem 3 *Under Assumptions 1,2 and 3b, for $k \in \{0, \ldots, s\}$,*

$$m_n^{1/2}\left(\hat{d}_k - d_k\right) \xrightarrow{D} N\left(0, \Phi_p/4\right).$$

Note that in Theorem 3, the limiting distribution of $m_n^{1/2}\left(\hat{d}_k - d_k\right)$ has mean zero. This asymptotic unbiasedness is ensured by Assumption 3b, which places strong restrictions on the separation between the memory parameters and also places a potentially stringent upper bound on the bandwidth m_n. A much weaker and indeed more standard assumption involving only m_n is the following.

Assumption 3c. As $n \to \infty$,

$$\frac{1}{m_n} + \frac{m_n^{1+2\rho} \log^2 m_n}{n^{2\rho}} \to 0.$$

If we account for the asymptotic bias, which can be determined from Corollary 10, and use Assumption 3c, we obtain the following result.

Corollary 1 *Under Assumptions 1, 2, and 3c, for* $k \in \{0, \ldots, s\}$,

$$m_n^{1/2} \left(\hat{d}_k - d_k - \mu_n \right) \xrightarrow{D} N\left(0, \Phi_p/4\right),$$

where $\mu_n = O_p\left(m_n^{d_k - d_{k-1}} + \omega_{m_n}^{d_k - d_{k+1}}\right)$, *the* $O_p\left(m_n^{d_k - d_{k-1}}\right)$ *term is vacuous if* $k = 0$, *and the* $O_p\left(\omega_{m_n}^{d_k - d_{k+1}}\right)$ *term is vacuous if* $k = s$.

Theorems 2, 3, and Corollary 1 pertain to a GSE estimator based on the residual series $\{b' y_t\}$, where b is any vector in $\mathcal{M}(\mathbf{X}_k)$. In practice, b will typically be an eigenvector of the averaged periodogram matrix corresponding to one of the sorted eigenvalues of this matrix, so that b is indeed in $\mathcal{M}(\mathbf{X}_k)$ for some k, but k is unknown. See the discussion at the beginning of this section. Here, we present some results on estimators of memory parameters based on the residual series constructed from an eigenvector corresponding to a particular sorted eigenvalue of the averaged periodogram matrix. Let w_t be a $q \times 1$ residual series, $w_t = \mathbf{X}' y_t$, and $\hat{d} = \left(\hat{d}_{11}, \ldots, \hat{d}_{qq}\right)'$ be the vector of univariate GSE estimates of $d = (d_{11}, \ldots, d_{qq})'$ based on w_t. First note that by Lemma 29 and the remark that follows it, $\mathbf{X} \xrightarrow{D} \mathring{\mathbf{X}}(\mathbf{H})$, where $\mathring{\mathbf{X}}(\mathbf{H})$ is a continuous function of $\mathbf{U}$ and $\mathbf{V}$ in Lemma 1. We will need the following assumption for our results.

Assumption 3d. (i) For all $k \in \{0, \ldots, s\}$, $\alpha_k > 1/2$. (ii) As $n \to \infty$,

$$\frac{1}{m_n} + \frac{m_n^{1+2\xi} \log^2 m_n}{n^{2\xi}} \to 0$$

where $\xi = \min\{\min_k \alpha_k, \rho\}$.

Theorem 4 *Under Assumptions 1, 2 and 3d,*

$$m_n^{1/2} \left(\hat{d} - d \right) \xrightarrow{D} N\left(0, \frac{\Phi_p}{4} \left(diag\mathbf{\Omega}\right)^{-1} \circ \mathbf{\Omega} \circ \mathbf{\Omega} \circ \left(diag\mathbf{\Omega}\right)^{-1}\right),$$

where

$$\mathbf{\Omega} = \mathbb{E}\left(\mathring{\mathbf{X}}' \mathbf{A} \mathbf{f}^{\dagger}(0) \mathbf{A}' \mathring{\mathbf{X}}\right).$$

***Remark* 2** *Simulation results not shown here reveal that the small-sample bias is reduced and the variance is stabilized if the GSE estimators omit the first* $m + p - 1$ *frequencies. This does not affect the validity* *of Theorem 4. The proof of Lemma 28 provides some motivation for this omission. Note that if no frequencies are omitted, then the first $m + p - 1$ frequencies are used twice: once for estimating the cointegrating vector, and once for estimating the memory parameter. If the frequencies are omitted, the finite-sample approximation to the variance in Hurvich and Chen (2000) is quite accurate.*

Theorem 4 yields the following result on the asymptotic distribution of $m_n^{1/2}\left(\hat{d}_{aa} - \hat{d}_{bb} - (d_{aa} - d_{bb})\right)$ under conditions that ensure asymptotic unbiasedness.

Corollary 2 *Under the assumptions of Theorem 4, for $a, b \in \{1, \ldots, q\}$, $a \neq b$,*

$$m_n^{1/2}\left(\hat{d}_{aa} - \hat{d}_{bb} - (d_{aa} - d_{bb})\right) \xrightarrow{D} N\left(0, \frac{\Phi_p}{2}\left(1 - \frac{\Omega_{ab}^2}{\Omega_{aa}\Omega_{bb}}\right)\right).$$

Next, we modify Corollary 2 to include a bias term, thereby allowing for weaker assumptions.

Corollary 3 *If $a \in N_k$, $b \in N_h$, for $k, h \in \{0, \ldots, s\}$ then under the assumptions of Corollary 1,*

$$m_n^{1/2}\left(\hat{d}_{aa} - \hat{d}_{bb} - (d_{aa} - d_{bb}) - \widetilde{\mu}_n\right) \xrightarrow{D} N\left(0, \frac{\Phi_p}{2}\left(1 - \frac{\Omega_{ab}^2}{\Omega_{aa}\Omega_{bb}}\right)\right)$$

where

$$\widetilde{\mu}_n = O_p\left(m_n^{d_k - d_{k-1}} + m_n^{d_h - d_{h-1}} + \omega_{m_n}^{d_k - d_{k+1}} + \omega_{m_n}^{d_h - d_{h+1}}\right).$$

7 Testing for Fractional Cointegration

In model (1), used throughout the paper thus far, we have assumed that $s \geq 1$, so that cointegration exists. Here, we expand model (1) to include the case of no cointegration ($s = 0$, or equivalently, $r = 0$), that is,

$$y_t = \mathbf{A}_0 u_t^{(0)} \tag{14}$$

where $\mathbf{A}_0$ is $q \times q$ with linearly independent columns, and all entries of $u_t^{(0)}$ have memory parameter d_0.

In practice, it is of interest to test for the presence of fractional cointegration. Such a test was proposed by Marinucci and Robinson (2001, pp. 236-237), following from an idea originally suggested in a different context by Hausman (1978), using a comparison of two estimates of d_0, one based on a multivariate Gaussian semiparametric estimator (see Lobato 1999) using $\{y_t\}_{t=1}^n$ with an imposed restriction that all entries have the same memory parameter, and the other estimator based on a univariate Gaussian semiparametric estimator of d_0 using (say) the first entry $\{y_{1,t}\}$ of $\{y_t\}$. It seems possible to use this idea together with differencing and tapering to yield a test for fractional integration in the current context, though we do not pursue this here. We focus instead on residual based methods, in which estimated memory parameters based on the various cointegrating residual series are compared. In a bivariate context, Velasco (2003) has considered properties of semiparametric memory parameter estimates based on cointegrating residuals under certain assumptions on the rate of convergence of the semiparametric estimator of the cointegrating parameters. However, he did not present a test for cointegration since his assumptions ruled out the no-cointegration case.

For our GSE estimators $\hat{d}$ based on cointegrating residuals, we have the following extensions of Theorem 4 and Corollary 2 to the no-cointegration case (14).

Corollary 4 *Under Assumptions 1, 2 and 3c, if there is no cointegration,*

$$m_n^{1/2}\left(\hat{d}-d\right)\xrightarrow{D}N\left(0,\frac{\Phi_p}{4}\left(diag\mathbf{\Omega}\right)^{-1}\circ\mathbf{\Omega}\circ\mathbf{\Omega}\circ\left(diag\mathbf{\Omega}\right)^{-1}\right),$$

where

$$\mathbf{\Omega}=\mathbb{E}\left(\mathring{\mathbf{X}}'\mathbf{A}\mathbf{f}^{\dagger}\left(0\right)\mathbf{A}'\mathring{\mathbf{X}}\right).$$

Corollary 5 *Under Assumptions 1, 2 and 3c, if there is no cointegration, for* $a,b\in\{1,\ldots,q\}$,

$$m_n^{1/2}\left(\hat{d}_{aa}-\hat{d}_{bb}\right)\xrightarrow{D}N\left(0,\frac{\Phi_p}{2}\left(1-\frac{\Omega_{ab}^2}{\Omega_{aa}\Omega_{bb}}\right)\right).$$

Corollary 5 and Corollary 3 justify a conservative hypothesis test for the null hypothesis of no cointegration based on the test statistic $T_n=m_n^{1/2}(\hat{d}_{11}-\hat{d}_{qq})$, whereby for a nominal level α test the null hypothesis is rejected in favor of the cointegration alternative hypothesis if and only if $T_n>(\Phi_p/2)^{1/2}z_{\alpha/2}$. Here, a bandwidth m_n satisfying Assumption 3c should be used. The test is conservative since $(\Phi_p/2)$ is an upper bound for the asymptotic variance of T_n.

In the next section, we provide a procedure for consistently identifying the cointegrating subspaces, assuming that there is cointegration, i.e., we will assume model (1) with $s>0$. In practice, before applying the procedure, we recommend pre-testing for cointegration using the test we have just described.

8 Identification of the Cointegrating Subspaces

Given data from model (1), assumed to possess fractional cointegration, the number $s>0$ of cointegrating subspaces and their dimensions $a_1,\ldots,a_s$ as well as the dimension a_0 of the non-cointegrating space will be unknown in general. Here, we provide a procedure for consistently identifying s, $a_0,\ldots,a_s$, under Assumption 1, which we make throughout this section. The procedure is based on the GSE estimates $\hat{d}_{11},\ldots,\hat{d}_{qq}$ formed from cointegrating residuals described in Section 6. The primary drawback of the procedure is that it requires the user to specify a lower bound on the minimum separation between the memory parameters. This minimum separation will typically also be unknown in practice, so the procedure we will describe here is not completely satisfactory. However, we note that such lower bounds on the minimum separation arise implicitly or explicitly in other works on semiparametric fractional cointegration. (See Robinson and Yajima 2002 Assumption D, and Velasco 2003, Theorems 2 and 4). In the current context, the need for such a lower bound is due to the nonstandard term $\tilde{\mu}_n$ appearing in Corollary 3. This term increases as the separation of the relevant memory parameters decreases.

Now, suppose that $s>0$ and let $\delta^*>0$ be the minimum separation between the memory parameters, $\delta^*=\min(d_0-d_1,\ldots,d_{s-1}-d_s)$. If $\delta^*>1/2$, the procedure is straightforward, as we will explain later.

If $0 < \delta^* < 1/2$, fix values of $\delta \in (0, \delta^*)$ and $\epsilon \in (1/2 - \delta, 1/2)$. This can be done provided that we have (or can correctly guess) a lower bound $\delta \in (0, 1/2)$ for δ^*. We now work with the pairs of GSE estimators $\hat{d}_{jj}, \hat{d}_{j+1,j+1}$, $j = 1, \ldots, q-1$, where the bandwidth m_n satisfies Assumption 3c. Fix a value of $C > 0$. For each $j \in \{1, \ldots, q-1\}$, we declare that $d_{jj} - d_{j+1,j+1} \neq 0$ if and only if

$$\hat{d}_{jj} - \hat{d}_{j+1,j+1} > C m_n^{-1/2+\epsilon} \quad .$$

We show here that for each j this procedure makes the correct decision as to whether or not $d_{jj} = d_{j+1,j+1}$ with a probability that tends to 1 as n increases. Our assumptions imply that for each $k \in \{0, \ldots, s\}$ $d_k - d_{k+1} > \delta$ and $d_k - d_{k-1} < -\delta$ whenever the left hand sides of these inequalities are well-defined, so the remainder term in Corollary 3 may be written as

$$\tilde{\mu}_n = O_p[m_n^{-\delta} + (m_n/n)^{\delta}] \quad ,$$

and Corollary 3 implies that for all $j \in \{1, \ldots, q-1\}$,

$$\hat{d}_{jj} - \hat{d}_{j+1,j+1} = d_{jj} - d_{j+1,j+1} + O_p[m_n^{-1/2} + m_n^{-\delta} + (m_n/n)^{\delta}] \quad .$$

Thus,

$$P\{\hat{d}_{jj} - \hat{d}_{j+1,j+1} > C m_n^{-1/2+\epsilon}\} = P\{m_n^{1/2-\epsilon} \hat{d}_{jj} - \hat{d}_{j+1,j+1} > C\}$$
$$= P\{m_n^{1/2-\epsilon}(d_{jj} - d_{j+1,j+1}) + O_p[m_n^{-\epsilon} + m_n^{1/2-\epsilon-\delta} + (m_n/n)^{1/2-\epsilon+\delta}] > C\} \quad .$$

Our assumptions imply that $1/2 - \epsilon - \delta < 0$ and $1/2 - \epsilon + \delta > 0$. It follows that if $d_{jj} = d_{j+1,j+1}$ then

$$P\{\hat{d}_{jj} - \hat{d}_{j+1,j+1} > C m_n^{-1/2+\epsilon}\} \to 0 \quad .$$

On the other hand, if $d_{jj} \neq d_{j+1,j+1}$ then

$$P\{\hat{d}_{jj} - \hat{d}_{j+1,j+1} > C m_n^{-1/2+\epsilon}\} = P\{m_n^{1/2-\epsilon}(d_{jj} - d_{j+1,j+1}) + o_p(m_n^{1/2-\epsilon}) > C\} \to 1 \quad .$$

In view of the above discussion, we have the following procedure for identifying the cointegrating subspaces. The procedure is a formalization of the simple idea that we can set the group boundaries at the points where the estimates of the memory parameters differ by a sufficient amount. First, we estimate s by $\hat{s}$, the number values of $j \in \{1, \ldots, q-1\}$ such that $\hat{d}_{jj} - \hat{d}_{j+1,j+1} > C m_n^{-1/2+\epsilon}$. Then $\hat{s} \to s$ almost surely, as long as $s > 0$, as we are assuming here. If $\hat{s} = 0$ then the procedure terminates without identifying any of the cointegrating spaces, but we will not dwell on this scenario since our assumptions imply that $P\{\hat{s} = 0\} \to 0$. Henceforth, we assume that $\hat{s} \geq 1$. We estimate a_0 by $\hat{a}_0 = \min\{j \in \{1, \ldots, q-1\} : \hat{d}_{jj} - \hat{d}_{j+1,j+1} > C m_n^{-1/2+\epsilon}\}$. Then $\hat{a}_0 \to a_0$ almost surely. If $\hat{s} = 1$ we set $\hat{a}_1 = q - \hat{a}_0$ and the procedure terminates. If $\hat{s} > 1$ then for each $k \in \{1, \ldots, \hat{s}-1\}$ we estimate a_k by

$$\hat{a}_k = \min \{j \in \{1, \ldots, q - (\hat{a}_0 + \cdots + \hat{a}_{k-1}) - 1\} : \hat{d}_{j+\hat{a}_0+\cdots+\hat{a}_{k-1}, j+\hat{a}_0+\cdots+\hat{a}_{k-1}} - \hat{d}_{j+\hat{a}_0+\cdots+\hat{a}_{k-1}+1, j+\hat{a}_0+\cdots+\hat{a}_{k-1}+1} > C m_n^{-1/2+\epsilon}\Big\}$$

and we set $\hat{a}_s = q - (\hat{a}_0 + \ldots + \hat{a}_{s-1})$. From the discussion above, it follows that $\hat{s}$, $\hat{a}_0, \ldots, \hat{a}_{\hat{s}}$ are consistent estimators of s, $a_0, \ldots, a_s$, respectively.

Finally, we discuss a modified version of the above procedure for the case $\delta^* > 1/2$. In this case, we compare the GSE estimators $\hat{d}_{jj}$ and $\hat{d}_{j+1,j+1}$ for $j = 1, \ldots, q$, using a bandwidth m_n satisfying

Assumption 3d, part (ii), with $\xi = \min\{\delta^*, \rho\}$. Fix an $\epsilon \in (0, 1/2)$ and a $C > 0$. Then, proceeding as above, for each $j \in \{1, \ldots, q-1\}$, we declare that $d_{jj} - d_{j+1,j+1} \neq 0$ if and only if $\hat{d}_{jj} - \hat{d}_{j+1,j+1} > Cm_n^{-1/2+\epsilon}$. The justification for this procedure follows a simplified version of the argument given for the case $\delta^* < 1/2$, but here we can use Corollary 2, which contains no bias term, instead of the more complicated Corollary 3.

9 Appendix

9.1 Proofs For Section 4

Proof of Lemma 2: For $k = 0, 1, \ldots, s$, let $\mathbf{B}_k$ be a $q \times a_k$ matrix with orthonormal columns such that $\mathcal{M}(\mathbf{B}_k) = \mathcal{B}_k$ and let $\mathbf{B} = \begin{bmatrix} \mathbf{B}_0 & \cdots & \mathbf{B}_s \end{bmatrix}$. Let $\mathbf{\Phi}= \mathbf{B}' I_m(y_t, y_t)\mathbf{B}$. Since $\mathbf{B}'\mathbf{B} = \mathbf{I}$, we have $\lambda_j(\mathbf{\Phi}) = \lambda_j$. Let $z_t^{(k)} = \left(u_t^{(k)}, \ldots, u_t^{(s)}\right), k = 1, \ldots s$, and $z_t^{(0)} = z_t$. Let $\mathbf{A}^{(k)} = \begin{bmatrix} \mathbf{A}_k & \cdots & \mathbf{A}_s \end{bmatrix}$, $k = 0, 1, \ldots, s$. We first partition $\mathbf{\Phi}$ into $(s+1) \times (s+1)$ blocks, such that the (k, ℓ) block has dimension $(a_k \times a_\ell)$. Note that

$$\mathbf{\Phi} = \mathbf{B}'\mathbf{A} I_m(z_t, z_t)\mathbf{A}'\mathbf{B},$$

where $\mathbf{B}'\mathbf{A}$ is an upper triangular block matrix. We have

$$\begin{aligned} \mathbf{\Phi}_{k\ell} &= \mathbf{B}_k'\mathbf{A}^{(k)} I_m\left(z_t^{(k)}, z_t^{(\ell)}\right)\mathbf{A}^{(\ell)\prime}\mathbf{B}_\ell, \quad \text{for } k \le \ell,\ k, \ell = 0, 1, \ldots, s \quad , \\ \mathbf{\Phi}_{\ell k} &= \mathbf{\Phi}_{k\ell}'. \end{aligned} \tag{15}$$

Fix a value of $k \in \{0, \ldots, s\}$. Note that by Lemma 1, all the elements in the kth block, $\mathbf{\Phi}_{kk}$, are $O_p\left(n^{2d_k}\right)$. Now

$$\sum_{j \in N_k} \lambda_j \le \sum_{j \in N_k \cup \cdots \cup N_s} \lambda_j \le \sum_{v=k}^{s} \operatorname{tr}\{\mathbf{\Phi}_{vv}\} = O_p\left(n^{2d_k}\right).$$

See, for example, Theorem 14 of Magnus and Neudecker (1999, p. 211). We have $\lambda_j = O_p\left(n^{2d_k}\right)$, for $j \in N_k$.□

Proof of Lemma 3: We construct another similar matrix for $I_m(y_t, y_t)$. Let $\mathcal{C}_s = \mathcal{M}(\mathbf{A}_s)$, and $\mathcal{C}_k$, $k = 0, \ldots, s-1$, be the subspaces such that

$$\mathcal{M}^\perp(\mathbf{A}_{k+1}, \ldots, \mathbf{A}_s) = \mathcal{M}^\perp(\mathbf{A}_k, \ldots, \mathbf{A}_s) \oplus \mathcal{C}_k,$$

and $\mathcal{C}_k \perp \mathcal{M}^\perp(\mathbf{A}_k, \ldots, \mathbf{A}_s)$. Hence a vector $\alpha \in \mathcal{C}_k$ satisfies $\alpha'\mathbf{A}_\ell = 0$, $\ell = k+1, \ldots, s$ and $\alpha'\mathbf{A}_k \neq 0$. Also $\mathcal{C}_j \perp \mathcal{C}_k$ for $j \neq k$. For $k \in \{0, \ldots, s\}$, let $\mathbf{C}_k$ be a $q \times a_k$ matrix with orthonormal columns such that $\mathcal{M}(\mathbf{C}_k) = \mathcal{C}_k$ and let $\mathbf{C} = \begin{bmatrix} \mathbf{C}_0 & \cdots & \mathbf{C}_s \end{bmatrix}$. Since $\mathbf{C}'\mathbf{C} = \mathbf{I}$, $\mathbf{\Omega} = \mathbf{C}' I_m(y_t, y_t)\mathbf{C}$ is similar to $I_m(y_t, y_t)$.

From Lemma 1, $\mathbf{Q}_n^{(k)}$ converges in distribution to a matrix that is positive definite with probability one. Since an eigenvalue of a matrix is a continuous function of the entries of the matrix, we conclude that $\lambda_{j_k^*}\left(\mathbf{Q}_n^{(k)}\right)$, the smallest eigenvalue of $\mathbf{Q}_n^{(k)}$, converges in distribution to a random variable that has no mass at zero. The proof can be completed by showing that $n^{-2du_k}\lambda_{j_k^*} \ge c_k \lambda_{j_k^*}\left(\mathbf{Q}_n^{(k)}\right)$. Let

$\widetilde{z}_t^{(k)} = \left(x_t, u_t^{(1)}, \ldots, u_t^{(k)}\right), k = 1, \ldots s$, and $\widetilde{z}_t^{(0)} = x_t$. Let $\widetilde{\mathbf{d}}_n^{(k)}$ be the leading $j_k^* \times j_k^*$ principal submatrix of $\mathbf{d}_n$. By Corollary 2.2.1 of Anderson and Das Gupta (1963),

$$\lambda_{j_k^*}\left(I_m\left(\widetilde{z}_t^{(k)}, \widetilde{z}_t^{(k)}\right)\right) = \lambda_{j_k^*}\left(\widetilde{\mathbf{d}}_n^{(k)}\mathbf{Q}_n^{(k)}\widetilde{\mathbf{d}}_n^{(k)}\right) \geq \lambda_{j_k^*}\left(\widetilde{\mathbf{d}}_n^{(k)}\right)\lambda_{j_k^*}\left(\mathbf{Q}_n^{(k)}\right)\lambda_{j_k^*}\left(\widetilde{\mathbf{d}}_n^{(k)}\right) = n^{2d_k}\lambda_{j_k^*}\left(\mathbf{Q}_n^{(k)}\right). \quad (16)$$

We have $\lambda_j(\mathbf{\Omega}) = \lambda_j$. Note that

$$\mathbf{\Omega} = \mathbf{C}'\mathbf{A}I_m(z_t, z_t)\mathbf{A}'\mathbf{C},$$

where $\mathbf{C}'\mathbf{A}$ is a lower triangular block matrix. Let $\mathbf{P}^{(k)} = \widetilde{\mathbf{C}}^{(k)\prime}\widetilde{\mathbf{A}}^{(k)}$ where $\widetilde{\mathbf{A}}^{(k)} = \begin{bmatrix} \mathbf{A}_0 & \cdots & \mathbf{A}_k \end{bmatrix}$, $k = 0, 1, \ldots, s$ and $\widetilde{\mathbf{C}}^{(k)}$ is defined similarly. Let $\mathbf{\Omega}^{(k)}$ denote the leading $j_k^* \times j_k^*$ principal submatrix of $\mathbf{\Omega}$. We have

$$\mathbf{\Omega}^{(k)} = \mathbf{P}^{(k)}I_m\left(\widetilde{z}_t^{(k)}, \widetilde{z}_t^{(k)}\right)\mathbf{P}^{(k)\prime}.$$

Notice that $\mathbf{P}^{(k)}$ and $I_m\left(\widetilde{z}_t^{(k)}, \widetilde{z}_t^{(k)}\right)$ are square matrices with the same dimension. Since $\lambda_{j_k^*}\left(\mathbf{\Omega}^{(k)}\right)$ and $\lambda_{j_k^*}\left(I_m\left(\widetilde{z}_t^{(k)}, \widetilde{z}_t^{(k)}\right)\right)$ are the smallest eigenvalues of $\mathbf{\Omega}^{(k)}$ and $I_m\left(\widetilde{z}_t^{(k)}, \widetilde{z}_t^{(k)}\right)$ respectively,

$$\begin{aligned}
\lambda_{j_k^*}\left(\mathbf{\Omega}^{(k)}\right) &= \min_\alpha \frac{\alpha'\mathbf{\Omega}^{(k)}\alpha}{\alpha'\alpha} = \min_\alpha\left\{\frac{\alpha'\mathbf{P}^{(k)}I_m\left(\widetilde{z}_t^{(k)}, \widetilde{z}_t^{(k)}\right)\mathbf{P}^{(k)\prime}\alpha}{\alpha'\mathbf{P}^{(k)}\mathbf{P}^{(k)\prime}\alpha}\frac{\alpha'\mathbf{P}^{(k)}\mathbf{P}^{(k)\prime}\alpha}{\alpha'\alpha}\right\} \\
&\geq \min_\alpha\left\{\frac{\alpha'\mathbf{P}^{(k)}I_m\left(\widetilde{z}_t^{(k)}, \widetilde{z}_t^{(k)}\right)\mathbf{P}^{(k)\prime}\alpha}{\alpha'\mathbf{P}^{(k)}\mathbf{P}^{(k)\prime}\alpha}\right\}\min_\alpha\left\{\frac{\alpha'\mathbf{P}^{(k)}\mathbf{P}^{(k)\prime}\alpha}{\alpha'\alpha}\right\} \\
&\geq \lambda_{j_k^*}\left(I_m\left(\widetilde{z}_t^{(k)}, \widetilde{z}_t^{(k)}\right)\right)\lambda_{j_k^*}\left(\mathbf{P}^{(k)}\mathbf{P}^{(k)\prime}\right) \\
&= c_k\lambda_{j_k^*}\left(I_m\left(\widetilde{z}_t^{(k)}, \widetilde{z}_t^{(k)}\right)\right).
\end{aligned}$$

By the Sturmian Separation Theorem (Rao 1973, p. 64, or Theorem 12, Magnus and Neudecker 1999, p. 210), the above equation and Equation (16),

$$\lambda_{j_k^*} \geq \lambda_{j_k^*}\left(\mathbf{\Omega}^{(k)}\right) \geq c_k\,\lambda_{j_k^*}\left(I_m\left(\widetilde{z}_t^{(k)}, \widetilde{z}_t^{(k)}\right)\right) \geq c_k n^{2d_k}\lambda_{j_k^*}\left(\mathbf{Q}_n^{(k)}\right).$$

Hence $n^{-2du_k}\lambda_{j_k^*} \geq c_k\lambda_{j_k^*}\left(\mathbf{Q}_n^{(k)}\right)$. □

9.2 Proofs For Section 5

Proof of Lemma 4: Since $\mathbf{X}_k^\perp(\mathbf{H}) = \begin{bmatrix} \mathbf{X}_0(\mathbf{H}) & \cdots & \mathbf{X}_{k-1}(\mathbf{H}) & \mathbf{X}_{k+1}(\mathbf{H}) & \cdots & \mathbf{X}_q(\mathbf{H}) \end{bmatrix}$, we have

$$\begin{aligned}
\left\|\sin\Theta\left\{\mathcal{M}\left(\mathbf{X}_k(\mathbf{H})\right), \mathcal{M}\left(\mathbf{X}_k\right)\right\}\right\|_F &\leq \left\|\left(\mathbf{X}_k^\perp(\mathbf{H})\right)^*\mathbf{X}_k\right\|_F \leq \sum_{\ell=0,\ell\neq k}^{s}\left\|\left(\mathbf{X}_\ell(\mathbf{H})\right)^*\mathbf{X}_k\right\|_F \\
&= O_p\left(\max_{\ell\neq k} n^{-|d_k - d_\ell|}\right) = O_p\left(n^{-\alpha_k}\right),
\end{aligned}$$

by Lemma 7.□

Proof of Lemma 5: For $k = 1, \ldots, s-1$, we have

$$P\left(\mathcal{M}\left(\mathbf{X}_k\left(\mathbf{H}\right)\right) = \mathcal{B}_k\right) = P\left(\left\{\mathcal{M}\mathbf{X}_k\left(\mathbf{H}\right) \cap \oplus_{\ell \leq k-1}\mathcal{B}_\ell = \mathbf{0}\right\} \cap \left\{\mathcal{M}\mathbf{X}_k\left(\mathbf{H}\right) \cap \oplus_{\ell \geq k+1}\mathcal{B}_\ell = \mathbf{0}\right\}\right).$$

Hence

$$\begin{aligned} P\left(\mathcal{M}\left(\mathbf{X}_k\left(\mathbf{H}\right)\right) \neq \mathcal{B}_k\right) &= P\left(\left\{\mathcal{M}\mathbf{X}_k\left(\mathbf{H}\right) \cap \oplus_{\ell \leq k-1}\mathcal{B}_\ell \neq \mathbf{0}\right\} \cup \left\{\mathcal{M}\mathbf{X}_k\left(\mathbf{H}\right) \cap \oplus_{\ell \geq k+1}\mathcal{B}_\ell \neq \mathbf{0}\right\}\right) \\ &\leq P\left(\mathcal{M}\mathbf{X}_k\left(\mathbf{H}\right) \cap \oplus_{\ell \leq k-1}\mathcal{B}_\ell \neq \mathbf{0}\right) + P(\mathcal{M}\mathbf{X}_k\left(\mathbf{H}\right) \cap \oplus_{\ell \geq k+1}\mathcal{B}_\ell \neq \mathbf{0}) \\ &= O\left(n^{-2d_{k-1}+2d_k} + n^{-2d_k+2d_{k+1}}\right), \end{aligned}$$

by Lemma 11. Similarly,

$$P\left(\mathcal{M}\left(\mathbf{X}_0\left(\mathbf{H}\right)\right) \neq \mathcal{B}_0\right) = O_p\left(n^{-2d_0+2d_1}\right)$$

and

$$P\left(\mathcal{M}\left(\mathbf{X}_s\left(\mathbf{H}\right)\right) \neq \mathcal{B}_s\right) = O_p\left(n^{-2d_{s-1}+2d_s}\right).$$

We have completed the proof.□

We will need the following lemma for the proof of Lemma 7.

Lemma 6 *Let* $\mathbf{K} = \text{diag}\left(\mathbf{B}_0'\mathbf{A}_0, \ldots, \mathbf{B}_s'\mathbf{A}_s\right)$, *then*

$$\mathbf{d}_n^{-1}\mathbf{\Phi}\mathbf{d}_n^{-1} \xrightarrow{D} \mathbf{K}\left(\mathbf{U}\mathbf{U}' + \mathbf{V}\mathbf{V}'\right)\mathbf{K}',$$

where $\mathbf{d}_n, \mathbf{U}$ *and* $\mathbf{V}$ *are defined as in Lemma 1.*

Proof. We write $\mathbf{\Phi} = \mathbf{K}I_m\left(z_t, z_t\right)\mathbf{K} + \mathbf{R}$, where $\mathbf{R}$ is a symmetric matrix with its (k,ℓ)th entry

$$\begin{aligned} \mathbf{R}_{k\ell} = &\mathbf{B}_k'\mathbf{A}_k I_m\left(u_t^{(k)}, z_t^{(\ell+1)}\right)\mathbf{A}^{(\ell+1)\prime}\mathbf{B}_\ell + \mathbf{B}_k'\mathbf{A}^{(k+1)}I_m\left(z_t^{(k+1)}, u_t^{(\ell)}\right)\mathbf{A}_\ell'\mathbf{B}_\ell \\ &+ \mathbf{B}_k'\mathbf{A}^{(k+1)}I_m\left(z_t^{(k+1)}, z_t^{(\ell+1)}\right)\mathbf{A}^{(\ell+1)\prime}\mathbf{B}_\ell, \end{aligned}$$

for $k \leq \ell$, $\ell = 0, 1, \ldots, (s-1)$,

$$\mathbf{R}_{ks} = \mathbf{B}_k'\mathbf{A}^{(k+1)}I_m\left(z_t^{(k+1)}, z_t^{(s)}\right)\mathbf{A}^{(s)\prime}\mathbf{B}_s,$$

for $k < s$, and $\mathbf{R}_{ss} = 0$. Now

$$\mathbf{d}_n^{-1}\mathbf{\Phi}\mathbf{d}_n^{-1} = \mathbf{d}_n^{-1}\mathbf{K}I_m\left(z_t, z_t\right)\mathbf{K}'\mathbf{d}_n^{-1} + \mathbf{d}_n^{-1}\mathbf{R}\mathbf{d}_n^{-1}.$$

Since $\mathbf{d}_n^{-1}\mathbf{K} = \text{diag}\left(n^{-d_{u_0}}\mathbf{B}_0'\mathbf{A}_0, \ldots, n^{-d_{u_s}}\mathbf{B}_s'\mathbf{A}_s\right)$, we have

$$\mathbf{d}_n^{-1}\mathbf{K}I_m\left(z_t, z_t\right)\mathbf{K}'\mathbf{d}_n^{-1} \xrightarrow{D} \mathbf{K}\left(\mathbf{U}\mathbf{U}' + \mathbf{V}\mathbf{V}'\right)\mathbf{K}',$$

by Lemma 1. We complete the proof by showing $\left\|\mathbf{d}_n^{-1}\mathbf{R}\mathbf{d}_n^{-1}\right\| = o_p\left(1\right)$.Note that

$$\mathbf{R}_{k\ell} = O_p\left(n^{d_k+d_{\ell+1}} + n^{d_{k+1}+d_\ell}\right), \text{ for } k \leq \ell,$$

and the (k,ℓ)th entry of $\mathbf{d}_n^{-1}\mathbf{R}\mathbf{d}_n^{-1}$ is

$$n^{-d_k-d_\ell}\mathbf{R}_{k\ell} = o_p(1), \text{ for } k \leq \ell.$$

Since $\mathbf{d}_n^{-1}\mathbf{R}\mathbf{d}_n^{-1}$ is symmetric, we have completed the proof.□

Corollary 6 *For $k = 0, 1, \ldots, s$, the limiting distribution of $n^{-2d_k}\boldsymbol{\Phi}_{kk}$ is positive definite and has distinct eigenvalues with probability 1. Furthermore,*

$$n^{-2d_k}\lambda_j\left(\boldsymbol{\Phi}_{kk}\right) \xrightarrow{D} \Omega_j^{(k)},$$

where $\Omega_j^{(k)}$ is a random variable that has no mass at 0.

Proof. By Lemma 6,

$$n^{-2d_k}\boldsymbol{\Phi}_{kk} \xrightarrow{D} \mathbf{B}_k'\mathbf{A}_k\left(\mathbf{U}^{(k)}\mathbf{U}^{(k)\prime} + \mathbf{V}^{(k)}\mathbf{V}^{(k)\prime}\right)\mathbf{A}_k'\mathbf{B}_k,$$

where $\mathbf{U}^{(k)}{=}\left(U_1^{(k)}, \ldots, U_m^{(k)}\right)$ and $\mathbf{V}^{(k)}{=}\left(V_1^{(k)}, \ldots, V_m^{(k)}\right)$, $U_i^{(k)}, V_j^{(k)}$ are $a_k{\times}1$ vectors, and vec $\left\{U_i^{(k)}, V_j^{(k)}\right\}_{j,k=1}^m$ is a $2ma_k$-variate normal random variable with zero mean, and covariance determined by

$$\begin{aligned}
\mathbb{E}\left(U_i^{(k)}U_j^{(k)\prime}\right) &= \int_{\mathbb{R}} \upsilon_i(x)\,\overline{\upsilon_j(x)}\mathbf{G}_{0,kk}(dx),\\
\mathbb{E}\left(V_i^{(k)}V_j^{(k)\prime}\right) &= \int_{\mathbb{R}} \nu_i(x)\,\overline{\nu_j(x)}\mathbf{G}_{0,kk}(dx),\\
\mathbb{E}\left(U_i^{(k)}V_j^{(k)\prime}\right) &= \int_{\mathbb{R}} \upsilon_i(x)\,\overline{\nu_j(x)}\mathbf{G}_{0,kk}(dx).
\end{aligned}$$

Here $\mathbf{G}_{0,kk}(dx)$ is the kth diagonal block $\mathbf{G}_0(dx)$. We see that the limiting distribution of $n^{-2d_k}\boldsymbol{\Phi}_{kk}$ is positive definite and has distinct eigenvalues with probability 1. Hence all of its eigenvalues converge in distribution to random variables with no mass at 0.□

Corollary 7 *For $i \in N_k$, $k = 0, \ldots, s$, $n^{-2d_k}\lambda_i\left(\boldsymbol{\Phi}_D\right) \xrightarrow{D} \xi_i^{(k)}$,where $\xi_i^{(k)}$ is a random variable that has no mass at zero.*

Lemma 7 $\left\|\mathbf{X}_\ell^*(\mathbf{H})\mathbf{X}_k\right\|_F = O_p\left(n^{-|d_k-d_\ell|}\right)$, *for all $\ell, k \in \{0, 1, \ldots, s\}$ with $\ell \neq k$.*

Proof: Since $\left\|\mathbf{X}_\ell^*(\mathbf{H})\mathbf{X}_k\right\|_F = \left\|\mathbf{X}_\ell^*(\boldsymbol{\Phi}_D)\mathbf{B}'\mathbf{B}\mathbf{X}_k(\boldsymbol{\Phi})\right\|_F = \left\|\mathbf{X}_\ell^*(\boldsymbol{\Phi}_D)\mathbf{X}_k(\boldsymbol{\Phi})\right\|_F$, we prove this lemma by showing that

$$\left\|\mathbf{X}_\ell^*(\boldsymbol{\Phi}_D)\mathbf{X}_k(\boldsymbol{\Phi})\right\|_F = O_p\left(n^{-|d_k-d_\ell|}\right).$$

Let $\boldsymbol{\Lambda} = \operatorname{diag}\{\lambda_j, j = 1, \ldots, q\}$ and $\Lambda^{(k)} = \{\lambda_j,\ j \in N_k\}$. We define $\boldsymbol{\Lambda}(\boldsymbol{\Phi}_D)$ and $\boldsymbol{\Lambda}^{(k)}(\boldsymbol{\Phi}_D)$ similarly for $\boldsymbol{\Phi}_D$. We will use the bound for the error in two subspaces within the nonzero space form Theorem 4.1, Barlow and Slapničar (2000) (which can be shown to apply in our context with probability one), that is,

$$\left\|\mathbf{X}_\ell^*(\boldsymbol{\Phi}_D)\mathbf{X}_k(\boldsymbol{\Phi})\right\|_F \le \frac{\left\|\boldsymbol{\Lambda}^{-1/2}(\boldsymbol{\Phi}_D)\mathbf{X}^*(\boldsymbol{\Phi}_D)\Delta\boldsymbol{\Phi}\mathbf{X}(\boldsymbol{\Phi})\boldsymbol{\Lambda}^{-1/2}\right\|_F}{\operatorname{relgap}\left(\Lambda^{(\ell)}(\boldsymbol{\Phi}_D), \Lambda^{(k)}\right)},$$

where

$$\operatorname{relgap}\left(\Lambda^{(\ell)}(\boldsymbol{\Phi}_D), \Lambda^{(k)}\right) = \min_{i\in N_k, j\in N_\ell}\left|\frac{\lambda_i(\boldsymbol{\Phi}) - \lambda_j(\boldsymbol{\Phi}_D)}{\lambda_i^{1/2}(\boldsymbol{\Phi}_D)\lambda_j^{1/2}(\boldsymbol{\Phi})}\right|.$$

We will prove this lemma by showing that

$$\left\| \mathbf{\Lambda}^{-1/2} \left(\mathbf{\Phi}_D \right) \mathbf{X}^* \left(\mathbf{\Phi}_D \right) \Delta \mathbf{\Phi} \mathbf{X} \left(\mathbf{\Phi} \right) \mathbf{\Lambda}^{-1/2} \right\|_F = O_p\left(1\right) \tag{17}$$

and

$$\frac{1}{\text{relgap}\left(\Lambda^{(\ell)}\left(\mathbf{\Phi}_D\right), \Lambda^{(k)}\right)} = O_p\left(n^{-|d_k - d_\ell|}\right). \tag{18}$$

By Lemmas 2, 3 and Corollary 7, relgap $\left(\Lambda^{(\ell)}\left(\mathbf{\Phi}_D\right), \Lambda^{(k)}\right) = O_p\left(n^{|d_k-d_\ell|}\right)$ and $n^{-|d_k-d_\ell|}$ relgap $\left(\Lambda^{(\ell)}\left(\mathbf{\Phi}_D\right), \Lambda^{(k)}\right) \geq \varsigma_{\ell,k}$, where $\varsigma_{\ell,k}$ is a random variable that has no mass at 0. We have (18). We next prove (17). Note that by Lemmas 1 and 6,

$$\mathbf{d}_n \mathbf{\Phi}^{-1} \mathbf{d}_n \xrightarrow{D} \mathbf{K}'^{-1} \left(\mathbf{U}\mathbf{U}' + \mathbf{V}\mathbf{V}' \right)^{-1} \mathbf{K}^{-1}.$$

Hence

$$\mathbf{d}_n \mathbf{X}\left(\mathbf{\Phi}\right) \mathbf{\Lambda}^{-1/2} = O_p\left(1\right),$$

since $\mathbf{d}_n \mathbf{\Phi}^{-1} \mathbf{d}_n = \mathbf{d}_n \mathbf{X}\left(\mathbf{\Phi}\right) \mathbf{\Lambda}^{-1/2} \mathbf{\Lambda}^{-1/2} \mathbf{X}'\left(\mathbf{\Phi}\right) \mathbf{d}_n = O_p\left(1\right).$ Similarly,

$$\mathbf{\Lambda}^{-1/2}\left(\mathbf{\Phi}_D\right) \mathbf{X}^*\left(\mathbf{\Phi}_D\right) \mathbf{d}_n = O_p\left(1\right).$$

We have

$$\begin{aligned}
&\left\| \mathbf{\Lambda}^{-1/2} \left(\mathbf{\Phi}_D \right) \mathbf{X}^* \left(\mathbf{\Phi}_D \right) \Delta \mathbf{\Phi} \mathbf{X} \left(\mathbf{\Phi} \right) \mathbf{\Lambda}^{-1/2} \right\|_F \\
&= \left\| \mathbf{\Lambda}^{-1/2} \left(\mathbf{\Phi}_D \right) \mathbf{X}^* \left(\mathbf{\Phi}_D \right) \mathbf{d}_n \mathbf{d_n^{-1}} \Delta \mathbf{\Phi} \mathbf{d_n^{-1}} \mathbf{d_n} \mathbf{X} \left(\mathbf{\Phi} \right) \mathbf{\Lambda}^{-1/2} \right\|_F \\
&\leq \left\| \mathbf{\Lambda}^{-1/2} \left(\mathbf{\Phi}_D \right) \mathbf{X}^* \left(\mathbf{\Phi}_D \right) \mathbf{d}_n \right\|_F \left\| \mathbf{d_n^{-1}} \Delta \mathbf{\Phi} \mathbf{d_n^{-1}} \right\|_F \left\| \mathbf{d_n} \mathbf{X} \left(\mathbf{\Phi} \right) \mathbf{\Lambda}^{-1/2} \right\|_F \\
&= O_p\left(1\right),
\end{aligned}$$

by Lemma 6. Hence $\left\| \mathbf{X}_\ell^*\left(\mathbf{\Phi}_D\right) \mathbf{X}_k\left(\mathbf{\Phi}\right) \right\|_F = O_p\left(n^{-|d_k-d_\ell|}\right).\square$

We need the following two lemmas for the proof of Lemma 10.

Lemma 8 *Under Assumption 1, there exists a finite constant C not depending on n such that for all sufficiently large n,*

$$\mathbb{E}\left[\lambda_1^2\left(\mathbf{Q}_n^{-1}\right)\right] < C.$$

Proof: Note that

$$\mathbf{Q}_n = \left(\mathbf{U}_n, \mathbf{V}_n\right)\left(\mathbf{U}_n, \mathbf{V}_n\right)',$$

where $\mathbf{U}_n$ and $\mathbf{V}_n$ are defined in Equation (10). Let

$$\mathcal{T}\left(W_n\right) = \lambda_1^2\left(\mathbf{Q}_n^{-1}\right),$$

where $W_n=$ vec $\left(\mathbf{U}_n, \mathbf{V}_n\right).$ By Assumption 1, $W_n \sim N\left(0, \mathbf{\Xi}_n\right),$ where $\mathbf{\Xi}_n = cov\left(W_n\right)$ and $\mathbf{\Xi}_n \longrightarrow \mathbf{\Xi}$, the covariance matrix of vec $\left(\mathbf{U}, \mathbf{V}\right)$ in Lemma 1. It was shown in Chen and Hurvich (2003b) that $\mathbf{\Xi}$ is positive definite. Thus for all sufficiently large n, $\mathbf{\Xi}_n$ is invertible and $\lambda_1\left(\mathbf{\Xi}_n\right) \rightarrow \lambda_1\left(\mathbf{\Xi}\right) > 0$.

For all sufficiently large n,

$$\mathbb{E}_{\mathbf{\Xi}_n}\left[\mathcal{T}\left(W_n\right)\right] = \left(2\pi\right)^{-mq} \left|\mathbf{\Xi}_n\right|^{-1/2} \int_{\mathbb{R}^{2mq}} \mathcal{T}\left(x\right) e^{-x'\mathbf{\Xi}_n^{-1}x/2} dx.$$

Since $x'\mathbf{\Xi}_n^{-1}x' \geq x'x/\lambda_1(\mathbf{\Xi}_n)$, we have

$$e^{-x'\mathbf{\Xi}_n^{-1}x/2} \leq e^{-x'x/2\lambda_1(\mathbf{\Xi}_n)}.$$

Since $\lambda_1(\mathbf{\Xi}_n) \to \lambda_1(\mathbf{\Xi}) > 0$ and since $|\mathbf{\Xi}_n|^{-1/2} \to |\mathbf{\Xi}|^{-1/2} > 0$, there exist constants $C_1 > 0$ and $C_2 > 0$ such that for all sufficiently large n,

$$\mathbb{E}_{\mathbf{\Xi}_n}[\mathcal{T}(W_n)] \leq C_1 \int_{\mathbb{R}^{2mq}} \mathcal{T}(x)\, e^{-C_2 x'x/2} dx = C,$$

a finite constant which does not depend on n. The above integral is the second moment of the largest eigenvalue of an inverse Wishart matrix and hence is bounded by a finite constant (Siskind,1972), in view of our assumption that $m > q+3$.□

Lemma 9 *Let $\mathbf{d}_n^{(k)} = \mathrm{diag}\left(n^{d_k}, \ldots, n^{d_k}, \ldots, n^{d_{u_s}}, \ldots, n^{d_{u_s}}\right)'$. Then, under Assumption 1, there exists a positive constant C such that for all sufficiently large n,*

$$\mathbb{E}^{1/2}\left[\lambda_1^2\left(n^{2d_k}\mathbf{\Phi}_{kk}^{-1}\right)\right] < C \quad , \quad k = 0, \ldots, s \quad .$$

Proof: We have

$$\begin{aligned}
\mathbf{\Phi}_{kk} &= \mathbf{B}_k'\mathbf{A}^{(k)} I_m\left(z_t^{(k)}, z_t^{(k)}\right)\mathbf{A}^{(k)\prime}\mathbf{B}_k \\
&= \mathbf{B}_k'\mathbf{A}^{(k)}\mathbf{d}_n^{(k)}\left(\mathbf{d}_n^{(k)}\right)^{-1} I_m\left(z_t^{(k)}, z_t^{(k)}\right)\left(\mathbf{d}_n^{(k)}\right)^{-1}\mathbf{d}_n^{(k)}\mathbf{A}^{(k)\prime}\mathbf{B}_k \\
&= \mathbf{J}\left\{\left(\mathbf{d}_n^{(k)}\right)^{-1} I_m\left(z_t^{(k)}, z_t^{(k)}\right)\left(\mathbf{d}_n^{(k)}\right)^{-1}\right\}\mathbf{J}'
\end{aligned}$$

where

$$\mathbf{J} = \mathbf{B}_k'\mathbf{A}^{(k)}\mathbf{d}_n^{(k)}.$$

We will use the inequality of Exercise 19 on page 238 of Magnus and Neudecker (1999). That is,

$$\begin{aligned}
\mathbf{\Phi}_{kk}^{-1} &= \left[\mathbf{J}\left\{\left(\mathbf{d}_n^{(k)}\right)^{-1} I_m\left(z_t^{(k)}, z_t^{(k)}\right)\left(\mathbf{d}_n^{(k)}\right)^{-1}\right\}\mathbf{J}'\right]^{-1} \\
&\leq \left(\mathbf{J}\mathbf{J}'\right)^{-1}\mathbf{J}\left\{\left(\mathbf{d}_n^{(k)}\right)^{-1} I_m\left(z_t^{(k)}, z_t^{(k)}\right)\left(\mathbf{d}_n^{(k)}\right)^{-1}\right\}^{-1}\mathbf{J}'\left(\mathbf{J}\mathbf{J}'\right)^{-1}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\mathrm{trace}\,\mathbf{\Phi}_{kk}^{-1} &\leq \mathrm{trace}\left(\mathbf{J}\mathbf{J}'\right)^{-1}\mathbf{J}\left\{\left(\mathbf{d}_n^{(k)}\right)^{-1} I_m\left(z_t^{(k)}, z_t^{(k)}\right)\left(\mathbf{d}_n^{(k)}\right)^{-1}\right\}^{-1}\mathbf{J}'\left(\mathbf{J}\mathbf{J}'\right)^{-1} \\
&\leq \lambda_1\left\{\left(\mathbf{d}_n^{(k)}\right)^{-1} I_m\left(z_t^{(k)}, z_t^{(k)}\right)\left(\mathbf{d}_n^{(k)}\right)^{-1}\right\}^{-1}\mathrm{trace}\left\{\left(\mathbf{J}\mathbf{J}'\right)^{-1}\mathbf{J}\mathbf{J}'\left(\mathbf{J}\mathbf{J}'\right)^{-1}\right\} \\
&= \lambda_1\left\{\left(\mathbf{d}_n^{(k)}\right)^{-1} I_m\left(z_t^{(k)}, z_t^{(k)}\right)\left(\mathbf{d}_n^{(k)}\right)^{-1}\right\}^{-1}\mathrm{trace}\left(\mathbf{J}\mathbf{J}'\right)^{-1}.
\end{aligned}$$

Since there exists a finite constant C such that $E\left[\lambda_1^2\left\{\left(\mathbf{d}_n^{(k)}\right)^{-1} I_m\left(z_t^{(k)}, z_t^{(k)}\right)\left(\mathbf{d}_n^{(k)}\right)^{-1}\right\}^{-1}\right] < C$ for all sufficiently large n by Lemma 8, the proof will be completed by showing that

$$\mathrm{trace}\left(\mathbf{J}\mathbf{J}'\right)^{-1} = O\left(n^{-2d_k}\right).$$

For $k < s$, we write

$$\begin{aligned}\mathbf{J} &= \mathbf{B}_k' \mathbf{A}^{(k)} \mathbf{d}_n^{(k)} \\ &= \mathbf{B}_k' \left[\begin{array}{cc} \mathbf{A}_k & \mathbf{A}^{(k+1)} \end{array} \right] \left[\begin{array}{cc} n^{d_k} \mathbf{I}_k & 0 \\ 0 & \mathbf{d}_n^{(k+1)} \end{array} \right],\end{aligned}$$

then

$$\mathbf{J}\mathbf{J}' = n^{2d_k} \mathbf{B}_k' \mathbf{A}_k \mathbf{A}_k' \mathbf{B}_k + \mathbf{B}_k' \mathbf{A}^{(k+1)} \mathbf{d}_n^{(k+1)} \mathbf{d}_n^{(k+1)} \mathbf{A}^{(k+1)} \mathbf{B}_k.$$

For $k = s$, the second term on the RHS is 0. Since both matrices on the RHS are symmetric and positive definite,

$$\lambda_{a_k} \mathbf{J}\mathbf{J}' \geq \lambda_{a_k} \left[n^{2d_k} \mathbf{B}_k' \mathbf{A}_k \mathbf{A}_k' \mathbf{B}_k \right],$$

(see, for example, Exercise 1 on page 204 of Magnus and Neudecker 1999.) We have

$$\lambda_1 \left(\mathbf{J}\mathbf{J}' \right)^{-1} \leq n^{-2d_k} \lambda_1 \left[\mathbf{B}_k' \mathbf{A}_k \mathbf{A}_k' \mathbf{B}_k \right]^{-1} = O\left(n^{-2d_k} \right).$$

□

Lemma 10 *Define $E_{k\ell}$ to be an event, $E_{k\ell} = \{\lambda_{a_k}(\mathbf{\Phi}_{kk}) > \lambda_1(\mathbf{\Phi}_{\ell\ell})\}$, $0 \leq k < \ell \leq s$. Then, under Assumption 1,*

$$P\left(E_{k\ell}^c\right) = O\left(n^{-2d_k + 2d_\ell}\right).$$

Proof: For $\ell > k$, $\ell = 1, \ldots, s$, we have, by Chebyshev's inequality and the Cauchy-Schwartz inequality,

$$\begin{aligned}P\left(E_{k\ell}^c\right) &= P\left\{\lambda_{a_k}(\mathbf{\Phi}_{kk}) \leq \lambda_1(\mathbf{\Phi}_{\ell\ell})\right\} \\ &= P\left\{ n^{-2d_k + 2d_\ell} \left(\frac{n^{-2d_\ell} \lambda_1(\mathbf{\Phi}_{\ell\ell})}{n^{-d_k} \lambda_{a_k}(\mathbf{\Phi}_{kk})} \right) \geq 1 \right\} \\ &\leq n^{-2d_k + 2d_\ell} \mathbb{E}\left[\frac{n^{-2d_\ell} \lambda_1(\mathbf{\Phi}_{\ell\ell})}{n^{-2d_k} \lambda_{a_k}(\mathbf{\Phi}_{kk})} \right] \\ &\leq n^{-2d_k + 2d_\ell} \mathbb{E}^{1/2}\left[\lambda_1^2\left(n^{-2d_\ell} \mathbf{\Phi}_{\ell\ell} \right) \right] \mathbb{E}^{1/2}\left[\lambda_1^2\left(n^{-2d_k} \mathbf{\Phi}_{kk} \right)^{-1} \right] \\ &\leq n^{-2d_k + 2d_\ell} \mathbb{E}^{1/2}\left[\operatorname{trace}^2\left(n^{-2d_\ell} \mathbf{\Phi}_{\ell\ell} \right) \right] \mathbb{E}^{1/2}\left[\lambda_1^2\left(n^{-2d_k} \mathbf{\Phi}_{kk} \right)^{-1} \right] \\ &= O\left(n^{-2d_k + 2d_\ell} \right),\end{aligned}$$

since $\mathbb{E}\left[\operatorname{trace}^2\left(n^{-2d_\ell}\mathbf{\Phi}_{\ell\ell}\right)\right] < C$ by Assumption 1 and Lemma 1 and $\mathbb{E}\left[\lambda_1^2\left(n^{-2d_k}\mathbf{\Phi}_{kk}\right)^{-1}\right] < C$ by Lemma 9. □

Lemma 11 *Under Assumption 1,*

$$P\left\{\mathcal{M}\mathbf{X}_k(\mathbf{H}) \cap \oplus_{\ell \leq h_1} \mathcal{B}_j \neq \mathbf{0}\right\} = O\left(n^{-2d_{h_1} + 2d_k}\right) \tag{19}$$

for $h_1 < k$, $k = 1, \ldots, s$ and

$$P\left\{\mathcal{M}\mathbf{X}_k(\mathbf{H}) \cap \oplus_{\ell \geq h_2} \mathcal{B}_j \neq \mathbf{0}\right\} = O\left(n^{-2d_k + 2d_{h_2}}\right) \tag{20}$$

for $h_2 > k$, $k = 0, \ldots, s-1$.

Proof: Since $\mathbf{H} = \mathbf{B}\mathbf{\Phi}_D\mathbf{B}'$, we have $\mathbf{X}_\ell(\mathbf{H}) = \mathbf{B}\mathbf{X}_\ell(\mathbf{\Phi}_D)$. Since $\mathbf{\Phi}_D$ is a block diagonal matrix,

$$\lambda_i(\mathbf{\Phi}_D) \in \{\lambda_j(\mathbf{\Phi}_{kk}) \,|k = 0, \ldots, s,\ j = 1, \ldots, a_k\}$$

and for $\lambda_i(\mathbf{\Phi}_D)$ such that $\lambda_i(\mathbf{\Phi}_D) = \lambda_j(\mathbf{\Phi}_{kk})$,

$$\chi_i(\mathbf{\Phi}_D) = \left(0, \ldots, 0, \chi_j'(\mathbf{\Phi}_{kk}), 0, \ldots 0\right)',$$

i.e., the first j^*_{k-1} entries are all zero. Define $E_{h\ell}$ to be an event, $E_{h\ell} = \{\lambda_{a_h}(\mathbf{\Phi}_{hh}) > \lambda_1(\mathbf{\Phi}_{\ell\ell})\}$, $0 \le h < \ell \le s$. We first prove (19).

$$P\{\mathcal{M}\mathbf{X}_k(\mathbf{H}) \cap \oplus_{\ell \le h_1}\mathcal{B}_\ell \ne \mathbf{0}\} = P\left(\mathbf{X}_k(\mathbf{\Phi}_D) \ne \begin{bmatrix} \mathbf{0} & \mathbf{Y} \end{bmatrix}'\right),$$

where the 0 in $\begin{bmatrix} \mathbf{0} & \mathbf{Y} \end{bmatrix}'$ has dimension $j^*_{h_1} \times a_k$, and $\mathbf{Y}$ has full rank. We have for $h_1 < k$, $k = 1, \ldots, s$

$$P\left(\mathbf{X}_k(\mathbf{\Phi}_D) \ne \begin{bmatrix} \mathbf{0} & \mathbf{Y} \end{bmatrix}'\right) = P\left(\bigcup_{\ell:\ell \le h_1} E^c_{\ell k}\right) \le \sum_{\ell:\ell \le h_1} P\left(E^c_{\ell k}\right) = O\left(\sum_{\ell:\ell \le h_1} n^{-2d_\ell + 2d_k}\right) = O\left(n^{-2d_{h_1} + 2d_k}\right)$$

by Lemma 10. Similarly, for (20),

$$P\{\mathcal{M}\mathbf{X}_k(\mathbf{H}) \cap \oplus_{\ell \ge h_2}\mathcal{B}_\ell \ne \mathbf{0}\} = P\left(\mathbf{X}_k(\mathbf{\Phi}_D) \ne \begin{bmatrix} \mathbf{Z} & \mathbf{0} \end{bmatrix}'\right),$$

where the 0 in $\begin{bmatrix} \mathbf{Z} & \mathbf{0} \end{bmatrix}'$ has dimension $\left(q - j^*_{h_2}\right) \times a_k$, and $\mathbf{Z}$ has full rank. We have for $h_2 > k, k = 0, \ldots, s-1$,

$$P\left(\mathbf{X}_k(\mathbf{\Phi}_D) \ne \begin{bmatrix} \mathbf{Z} & \mathbf{0} \end{bmatrix}'\right) = P\left(\bigcup_{\ell:\ell \ge h_2} E^c_{k\ell}\right) \le \sum_{\ell:\ell \ge h_2} P\left(E^c_{k\ell}\right) = O\left(\sum_{\ell:\ell \ge h_2} n^{-2d_k + 2d_\ell}\right) = O\left(n^{-2d_k + 2d_{h_2}}\right).$$

□

9.3 Proofs For Section 6

In this section, we will use the following decomposition and notation for the proofs. We write

$$b'\mathbf{A}I_{zz}(\omega_j)\mathbf{A}'b - b'\mathbf{A}\mathbf{f}\left(\omega_{\tilde{j}}\right)\mathbf{A}'b = b'\mathbf{A}\mathbf{R}(\omega_j)\mathbf{A}'b + b'\mathbf{A}\mathbf{S}(\omega_j)\mathbf{A}'b, \tag{21}$$

where

$$\mathbf{R}(\omega_j) = I_{zz}(\omega_j) - \mathbf{\Psi}\left(\omega_{\tilde{j}}\right) I_{\varepsilon\varepsilon}(\omega_j)\mathbf{\Psi}^*\left(\omega_{\tilde{j}}\right)$$

and

$$\mathbf{S}(\omega_j) = \mathbf{\Psi}\left(\omega_{\tilde{j}}\right) I_{\varepsilon\varepsilon}(\omega_j)\mathbf{\Psi}^*\left(\omega_{\tilde{j}}\right) - \mathbf{f}\left(\omega_{\tilde{j}}\right).$$

We will also use the following notation:

$$\mathcal{L}_{m_n}(d) = \frac{1}{m_n}\sum_{j=1}^{m_n} \omega_{\tilde{j}}^{2d} b'\mathbf{A}\mathbf{R}(\omega_j)\mathbf{A}'b,$$

$$\mathcal{M}_{m_n}(d) = \frac{1}{m_n}\sum_{j=1}^{m_n} \omega_{\tilde{j}}^{2d} b'\mathbf{A}\mathbf{S}(\omega_j)\mathbf{A}'b,$$

$$\mathcal{F}_{m_n}(d) = \frac{1}{m_n}\sum_{j=1}^{m_n} \omega_{\tilde{j}}^{2d} b'\mathbf{A}\mathbf{f}(\omega_j)\mathbf{A}'b. \tag{22}$$

Proof of Theorem 2. For $1/4 > \delta > 0$, let $N_\delta = \{d : |d - d_k| < \delta\}$. Then for $S(d) = R(d) - R(d_k)$,

$$P\left(\left|\hat{d}_k - d_k\right| \geq \delta\right) = P\left(\hat{d}_k \in N_\delta^c \cap \Theta\right) = P\left(\inf_{N_\delta^c \cap \Theta} R(d) \leq \inf_{N_\delta \cap \Theta} R(d)\right) \leq P\left(\inf_{N_\delta^c \cap \Theta} S(d) \leq 0\right),$$

where Define $\Theta_1 = \{d : \Delta \leq d \leq \Delta_2\}$, where $\Delta = \Delta_1$ when $d_k < 1/2 + \Delta_1$ and $d_k \geq \Delta > d_k - 1/2$ otherwise. Note that $d - d_k > -1/2$ for all $d \in \Theta_1$. When $d_k \geq 1/2 + \Delta_1$, define $\Theta_2 = \{d : \Delta_1 \leq d < \Delta\}$, and otherwise take Θ_2 to be empty. Hence

$$P\left(\left|\hat{d}_k - d_k\right| \geq \delta\right) \leq P\left(\inf_{N_\delta^c \cap \Theta_1} S(d) \leq 0\right) + P\left(\inf_{\Theta_2} S(d) \leq 0\right) = o(1)$$

by Lemmas 14 and 15.□

Proof of Theorem 3. By Theorem 2, $\widehat{d}_k$ satisfies

$$0 = \frac{\partial R\left(\hat{d}_k\right)}{\partial d} = \frac{\partial R(d_k)}{\partial d} + \frac{\partial^2 R\left(\tilde{d}\right)}{\partial d^2}\left(\hat{d}_k - d_k\right), \tag{23}$$

where $\left|\tilde{d} - d_k\right| \leq \left|\hat{d}_k - d_k\right|$. Let

$$\mathbf{Z}_n = 2m_n^{-1/2} \sum_{j=1}^{m_n} \nu_j \left(I_{\varepsilon\varepsilon}(\omega_j) - \mathbf{\Sigma}\right), \qquad \nu_j = \log \tilde{j} - \frac{1}{m_n} \sum_{j=1}^{m_n} \log \tilde{j},$$

and let

$$\mathfrak{Z}_n = \frac{1}{\mathcal{G}} b' \mathbf{A}_k \mathbf{\Psi}_k^{\dagger\prime}(0) \mathbf{Z}_n \mathbf{\Psi}_k^{\dagger}(0) \mathbf{A}_k' b,$$

where

$$\mathcal{G} = b' \mathbf{A}_k \mathbf{f}^{\dagger}(0) \mathbf{A}_k' b = b' \mathbf{A}_k \mathbf{\Psi}_k^{\dagger\prime}(0) \mathbf{\Sigma} \mathbf{\Psi}_k^{\dagger}(0) \mathbf{A}_k' b$$

and $\mathbf{\Psi}_k^{\dagger}(\omega)$ is a $q \times a_k$ sub-matrix of $\mathbf{\Psi}^{\dagger}(\omega) = \left[\mathbf{\Psi}_0^{\dagger}(\omega) \ \ \cdots \ \ \mathbf{\Psi}_s^{\dagger}(\omega)\right]$ in (8). By Lemmas 24 and 25,

$$\frac{\partial^2 R\left(\tilde{d}\right)}{\partial d^2} \xrightarrow{p} 4 \tag{24}$$

and

$$m_n^{1/2} \frac{\partial R(d_k)}{\partial d} = \mathfrak{Z}_n + o_p(1). \tag{25}$$

Following from Lemmas 0 and 8 of Hurvich and Chen (2000), the (u,v)th entry of $\mathbf{Z}_n$,

$$Z_{n,uv} \xrightarrow{D} N\left(0, 4\Phi_p \sigma_{uv}^2\right).$$

With a similar computation for the variance above and equation (46) in the proof of Lemma 17, we obtain

$$E\left(Z_{n,u_1v_1} Z_{n,u_2v_2}\right) \rightarrow 4\Phi_p \sigma_{u_1v_2} \sigma_{u_2v_1}$$

Following from the Cramer-Wold device, we have

$$\operatorname{vec} \mathbf{Z}_n \xrightarrow{D} \mathbf{Z} \sim N\left(0, 4\Phi_p \mathbf{\Sigma} \otimes \mathbf{\Sigma}\right). \tag{26}$$

By Corollary 15 , $b \xrightarrow{D} \mathring{b}$,

$$\mathfrak{z}_n \xrightarrow{D} \frac{\mathring{b}'\mathbf{A}_k \mathbf{\Psi}_k^{\dagger\prime}(0)\,\mathbf{Z}\mathbf{A}_k'\mathbf{\Psi}_k^{\dagger}(0)\,\mathring{b}}{\mathring{b}'\mathbf{A}_k \mathbf{\Psi}_k^{\dagger\prime}(0)\,\mathbf{\Sigma}\mathbf{A}_k'\mathbf{\Psi}_k^{\dagger}(0)\,\mathring{b}} := \mathfrak{z}, \tag{27}$$

where $\mathring{b}$ is independent of $\mathbf{Z}$ by Lemma 26. Let

$$\varphi = (\varphi_1, \ldots, \varphi_q)' = \mathbf{\Psi}_k^{\dagger\prime}(0)\,\mathbf{A}_k'\mathring{b},$$

we have for $\ell = 1, 2, \ldots,$

$$\begin{aligned}
\mathbb{E}\left(\mathfrak{z}^{2\ell}\right) &= \mathbb{E}\left[\mathbb{E}\left(\mathfrak{z}^{2\ell}|\mathring{b}\right)\right] \\
&= \mathbb{E}\left[\frac{1}{\left(\varphi'\mathbf{\Sigma}\varphi\right)^{2\ell}} \sum_{u_1,v_1,\cdots,u_{2\ell},v_{2\ell}=1}^{q} \varphi_{u_1}\varphi_{v_1}\cdots\varphi_{u_{2\ell}}\varphi_{v_{2\ell}}\mathbb{E}\left(Z_{u_1v_1}\ldots Z_{u_{2\ell}v_{2\ell}}|\mathring{b}\right)\right] \\
&= \mathbb{E}\left[\frac{1}{\left(\varphi'\mathbf{\Sigma}\varphi\right)^{2\ell}} \sum_{u_1,v_1,\cdots,u_{2\ell},v_{2\ell}=1}^{q} \varphi_{u_1}\varphi_{v_1}\cdots\varphi_{u_{2\ell}}\varphi_{v_{2\ell}}\mathbb{E}\left(Z_{u_1v_1}\ldots Z_{u_{2\ell}v_{2\ell}}\right)\right] \\
&= \mathbb{E}\left[\frac{(2\ell)!\,(4\Phi_p)^{\ell}}{2^{\ell}\ell!}\frac{1}{\left(\varphi'\mathbf{\Sigma}\varphi\right)^{2\ell}} \sum_{u_1,v_1,\cdots,u_{2\ell},v_{2\ell}=1}^{q} \varphi_{u_1}\varphi_{v_1}\cdots\varphi_{u_{2\ell}}\varphi_{v_{2\ell}}\sigma_{u_1v_1}\cdots\sigma_{u_{2\ell}v_{2\ell}}\right] \\
&= \frac{(2\ell)!\,(4\Phi_p)^{\ell}}{2^{\ell}\ell!}
\end{aligned}$$

and

$$\mathbb{E}\left(\mathfrak{z}^{2\ell-1}\right) = 0,$$

since for a zero mean multivariate normal $(Y_1, \ldots, Y_{2\ell})$ with $cov\,(Y_aY_b) = \rho_{a,b}$, $\mathbb{E}\,(Y_1, \ldots, Y_{2\ell-1}) = 0$ and

$$\mathbb{E}\,(Y_1, \ldots, Y_{2\ell}) = \sum_{\varpi} \rho_{\varpi_1}\cdots\rho_{\varpi_\ell},$$

where $(\varpi_1, \ldots, \varpi_\ell)$ is a partition of $(1, \ldots, 2\ell)$ with all $|\varpi_u| = 2$ and ϖ consists all these partitions (a total of $2^{-\ell}\,(2\ell)!/\ell!$ partitions). Since all the moments of $\mathfrak{z}$ match those of a $N\,(0, 4\Phi_p)$, we have

$$\mathfrak{z} \sim N\,(0, 4\Phi_p)\,.$$

Together with (23), (24), (25), and (27), we have proved the Theorem.□

We will need the following two lemmas.

Lemma 12 *If* $b \in \mathcal{M}\,(\mathbf{X}_k)$ *and* $\|b\| = 1$, *then under Assumption 1,*

$$b'\mathbf{A}_h = O_p\left(n^{-d_h+d_k}\right)$$

for $h < k$, $k = 1, \ldots, s$, *and*

$$b'\mathbf{A}_k = O_p\,(1)\,,$$

for $k = 0, \ldots, s$.

Lemma 13 *If $b \in \mathcal{M}(\mathbf{X}_k)$, $k = 0, \ldots, s,$ and $\|b\| = 1$, then under Assumption 1,*

$$\|b'\mathbf{A}_k\| \geq c(1 - \epsilon_k),$$

where $c > 0$ and $\epsilon_k = O_p(n^{-\alpha_k})$.

Proof of Lemma 12: Since $\mathbf{X}(\mathbf{H})$ is an orthogonal matrix and $\mathcal{M}\mathbf{X}(\mathbf{H}) = \mathbb{R}^q$,

$$b = \sum_{\ell=0}^{s} \mathbf{X}_\ell(\mathbf{H}) c_\ell, \tag{28}$$

where

$$c_\ell = \mathbf{X}'_\ell(\mathbf{H}) b = O_p\left(n^{-|d_k - d_\ell|}\right)$$

by Lemma 7. Furthermore, for $\ell > h$,

$$\begin{aligned}
\mathbb{E}[\|\mathbf{X}'_\ell(\mathbf{H})\mathbf{A}_h\|] &= \mathbb{E}\left[\|\mathbf{X}'_\ell(\mathbf{H})\mathbf{A}_h\| \mathbf{1}_{\{\mathcal{M}\mathbf{X}_\ell(\mathbf{H}) \subset \oplus_{j>h}\mathcal{B}_j\}}\right] + \mathbb{E}\left[\|\mathbf{X}'_\ell(\mathbf{H})\mathbf{A}_h\| \mathbf{1}_{\left\{\mathcal{M}\mathbf{X}_\ell(\mathbf{H}) \cap \oplus_{j\leq h}\mathcal{B}_j \neq \mathbf{0}\right\}}\right] \\
&\leq 0 + \mathbb{E}\left[\|\mathbf{X}'_\ell(\mathbf{H})\mathbf{A}_h\| \mathbf{1}_{\left\{\mathcal{M}\mathbf{X}_\ell(\mathbf{H}) \cap \oplus_{j\leq h}\mathcal{B}_j \neq \mathbf{0}\right\}}\right] \\
&= \mathbb{E}\left[\text{trace}^{1/2}\left(\mathbf{A}'_h \mathbf{X}_\ell(\mathbf{H}) \mathbf{X}'_\ell(\mathbf{H}) \mathbf{A}_h\right) \mathbf{1}_{\left\{\mathcal{M}\mathbf{X}_\ell(\mathbf{H}) \cap \oplus_{j\leq h}\mathcal{B}_j \neq \mathbf{0}\right\}}\right] \\
&\leq \mathbb{E}\left[\text{trace}^{1/2}(\mathbf{A}'_h\mathbf{A}_h)\,\text{trace}^{1/2}\left(\mathbf{X}_\ell(\mathbf{H})\mathbf{X}'_\ell(\mathbf{H})\right) \mathbf{1}_{\left\{\mathcal{M}\mathbf{X}_\ell(\mathbf{H}) \cap \oplus_{j\leq h}\mathcal{B}_j \neq \mathbf{0}\right\}}\right] \\
&= \|\mathbf{A}_h\| \left|P\left\{\mathcal{M}\mathbf{X}_\ell(\mathbf{H}) \cap \oplus_{j\leq h}\mathcal{B}_j \neq \mathbf{0}\right\}\right|^{1/2} \\
&= O\left(n^{-d_h + d_\ell}\right)
\end{aligned} \tag{29}$$

by exercise 12 (iii) of Chapter 11 in Magnus and Neudecker (1999) and Lemma 11. For $\ell \leq h$,

$$\mathbb{E}[\|\mathbf{X}'_\ell(\mathbf{H})\mathbf{A}_h\|] = O(1). \tag{30}$$

We have for $h < k$,

$$\begin{aligned}
b'\mathbf{A}_h &= \sum_{\ell=0}^{s} c'_\ell \mathbf{X}'_\ell(\mathbf{H})\mathbf{A}_h \\
&= \sum_{\ell:\ell\leq h} c'_\ell \mathbf{X}'_\ell(\mathbf{H})\mathbf{A}_h + \sum_{\ell:\ell>h} c'_\ell \mathbf{X}'_\ell(\mathbf{H})\mathbf{A}_h \\
&= O_p\left(\sum_{\ell:\ell\leq h} n^{-d_\ell + d_k} + \sum_{\ell:h<\ell\leq k} n^{-d_h + d_\ell - d_\ell + d_k} + \sum_{\ell:\ell>k} n^{-d_h + d_\ell - d_k + d_\ell}\right) \\
&= O_p\left(n^{-d_h + d_k}\right).
\end{aligned}$$

For $h = k$, the above equation is of $O_p(1)$ since $c_k = O_p(1)$ and $\mathbb{E}[\|\mathbf{X}'_k(\mathbf{H})\mathbf{A}_k\|] = O(1)$.□

Proof of Lemma 13: Note that

$$\begin{aligned}
\|b'\mathbf{A}_k\|^2 &= \left\|c'_k \mathbf{X}'_k(\mathbf{H})\mathbf{A}_k + \sum_{\ell=0,\ell\neq k}^{s} c'_\ell \mathbf{X}'_\ell(\mathbf{H})\mathbf{A}_k\right\| \\
&\geq \left|\|c'_k \mathbf{X}'_k(\mathbf{H})\mathbf{A}_k\| - \left\|\sum_{\ell=0,\ell\neq k}^{s} c'_\ell \mathbf{X}'_\ell(\mathbf{H})\mathbf{A}_k\right\|\right|.
\end{aligned} \tag{31}$$

Using (28), we have

$$1 = \|b\|^2 = \sum_{\ell=0}^{s} \|\mathbf{X}_\ell (\mathbf{H}) c_\ell\|^2 = \sum_{\ell=0}^{s} \|c_\ell\|^2 = \|c_k\|^2 + \sum_{\ell=0,\ell\neq k}^{s} \|c_\ell\|^2,$$

and

$$\sum_{\ell=0,\ell\neq k}^{s} \|c_\ell\|^2 = O_p\left(n^{-2\alpha_k}\right). \tag{32}$$

by Lemma 4. Thus,

$$\|c_k\|^2 = 1 - O_p\left(n^{-2\alpha_k}\right). \tag{33}$$

By (32), (29) and (30),

$$\left\|\sum_{\ell=0,\ell\neq k}^{s} c_\ell' \mathbf{X}_\ell' (\mathbf{H}) \mathbf{A}_k\right\| \leq \left(\sum_{\ell=0,\ell\neq k}^{s} \|c_\ell\|^2 \sum_{\ell=0,\ell\neq k}^{s} \|\mathbf{X}_\ell' (\mathbf{H}) \mathbf{A}_k\|^2\right)^{1/2} = O_p\left(n^{-\alpha_k}\right). \tag{34}$$

Furthermore, if $\mathcal{M}\mathbf{X}_k(\mathbf{H}) = \mathcal{B}_k$, then there exists an $a_k \times a_k$ orthogonal matrix $\mathbf{D}$ such that

$$\mathbf{X}_k(\mathbf{H}) = \mathbf{B}_k \mathbf{D}$$

since both $\mathbf{X}_k(\mathbf{H})$ and $\mathbf{B}_k$ are matrices with orthonormal columns. We have

$$\begin{aligned}
\|c_k\|^2 &= \operatorname{trace}\left\{c_k' \mathbf{D}\mathbf{B}_k' \mathbf{A}_k \left(\mathbf{B}_k'\mathbf{A}_k\right)^{-1} \left(\mathbf{A}_k'\mathbf{B}_k\right)^{-1} \mathbf{A}_k' \mathbf{B}_k \mathbf{D}' c_k\right\} \\
&\leq \left\|\left(\mathbf{B}_k'\mathbf{A}_k\right)^{-1}\right\|^2 \left\|c_k'\mathbf{D}\mathbf{B}_k'\mathbf{A}_k\right\|^2 \\
&= \left\|\left(\mathbf{B}_k'\mathbf{A}_k\right)^{-1}\right\|^2 \left\|c_k'\mathbf{X}_k'(\mathbf{H})\mathbf{A}_k\right\|^2.
\end{aligned}$$

It follows that

$$\left\|c_k'\mathbf{X}_k'(\mathbf{H})\mathbf{A}_k\right\|^2 \geq \left\|\left(\mathbf{B}_k'\mathbf{A}_k\right)^{-1}\right\|^{-2} \|c_k\|^2 = C\left(1-\delta_k\right)$$

where $\delta_k = O_p\left(n^{-2\alpha_k}\right)$ by (33). By (31), (34) and the above equation, $\|b'\mathbf{A}_k\| \geq C\left(1-\delta_k-\tilde{\epsilon}_k\right)$, where $\tilde{\epsilon}_k = O_p\left(n^{-\alpha_k}\right)$. We have completed the proof.□

Lemma 14 *Under the assumptions of Theorem 2,* $P\left(\inf_{N_\delta^c\cap\Theta_1} S(d) \leq 0\right) = o(1)$.

Proof. Let

$$U(d) = 2(d-d_k) - \log\{2(d-d_k)+1\}$$

and

$$T(d) = \log\frac{\widehat{G}(d_k)}{\mathcal{G}} - \log\frac{\widehat{G}(d)}{G(d)} - \log\left\{\frac{2(d-d_k)+1}{m_n}\sum_{j=1}^{m_n}\left(\frac{\tilde{j}}{m_n}\right)^{2(d-d_k)}\right\}$$
$$+ 2(d-d_k)\left\{\frac{1}{m_n}\sum_{j=1}^{m_n}\log\tilde{j} - (\log m_n - 1)\right\},$$

where

$$G(d) = \mathcal{G}\frac{1}{m_n}\sum_{j=1}^{m_n}\omega_{\tilde{j}}^{2(d-d_k)}$$

and

$$\mathcal{G} = b'\mathbf{A}_k\mathbf{f}^{\dagger}(0)\,\mathbf{A}_k'b.$$

Then

$$S(d) = U(d) - T(d).$$

We have

$$P\left(\inf_{N_\delta^c\cap\Theta_1} S(d)\le 0\right)\le P\left(\inf_{N_\delta^c\cap\Theta_1} U(d)\le \sup_{\Theta_1}|T(d)|\right).$$

Following the same arguments as in page 1635 of Robinson (1995b), it is sufficient to show that

$$\sup_{\Theta_1}\left|\frac{\widehat{G}(d)-G(d)}{G(d)}\right| = o_p(1). \tag{35}$$

Note that, by Corollary 11,

$$G(d) = C\mathcal{G}\omega_{m_n}^{2(d-d_k)} \ge C(1-\epsilon_k)\,\omega_{m_n}^{2(d-d_k)}. \tag{36}$$

where $\epsilon_k = O_p(n^{-\alpha_k})$. By Lemmas 18, 19 and 22, for $d\in\Theta_1$,

$$\left|\widehat{G}(d)-G(d)\right| = \mathcal{L}_{m_n}(d)+\mathcal{M}_{m_n}(d)+\mathcal{F}_{m_n}(d)-\mathcal{G}\frac{1}{m_n}\sum_{j=1}^{m_n}\omega_{\tilde{j}}^{2(d-d_k)} = o_p\left(\omega_{m_n}^{2d-2d_k}m_n^{-\epsilon}\right), \tag{37}$$

where $\mathcal{L}_{m_n}$, $\mathcal{M}_{m_n}$ and $\mathcal{F}_{m_n}$ are defined in (22). We have completed the proof.□

Lemma 15 *Under the assumptions of Theorem 2,* $P\left(\inf_{\Theta_2} S(d)\le 0\right) = o_p(1)$.

Proof. Following from the proof on page 1638-39 of Robinson (1995b), we write

$$S(d) = \log\left\{\widehat{D}(d)/\widehat{D}(d_k)\right\},$$

where

$$\widehat{D}(d) = \frac{1}{m_n}\sum_{j=1}^{m_n}\left(\frac{\tilde{j}}{e^{\nu}}\right)^{2(d-d_k)}\tilde{j}^{-2d_k}I_{vv,j} \qquad \text{and} \qquad \nu = \frac{1}{m_n}\sum_{j=1}^{m_n}\log\tilde{j}.$$

Let

$$\alpha_j = \begin{cases}\left(\frac{\tilde{j}}{e^{\nu}}\right)^{2(\Delta-d_k)}, & 1\le j\le e^{\nu}\\ \left(\frac{\tilde{j}}{e^{\nu}}\right)^{2(\Delta_1-d_k)} & e^{\nu}\le j\le m_n\end{cases}.$$

Note that $e^{\nu}\sim m_n/e$, thus

$$\alpha_j \sim \begin{cases}\left(\frac{ej}{m_n}\right)^{2(\Delta-d_k)} & 1\le j\le e^{\nu}\\ \left(\frac{ej}{m_n}\right)^{2(\Delta_1-d_k)} & e^{\nu}<j<m_n\end{cases}. \tag{38}$$

By choosing $\Delta < d_k - \frac{1}{2} + \frac{1}{4e}$, so that $m_n^{-1} \sum_{j=1}^{m_n} (\alpha_j - 1) \geq 1$ for all sufficiently large m_n, we have

$$P\left(\inf_{\Theta_2} S(d) \leq 0\right) \leq P\left(\frac{1}{m_n}\sum_{j=1}^{m_n}(\alpha_j - 1)\tilde{j}^{-2d_k} I_{vv}(\omega_j) \leq 0\right) = P\left(\frac{1}{m_n}\sum_{j=1}^{m_n}(\alpha_j - 1)\frac{I_{vv}(\omega_j)}{\mathcal{G}\omega_{\tilde{j}}^{-2d_k}} \leq 0\right)$$
$$\leq P\left(\left|\frac{1}{m_n}\sum_{j=1}^{m_n}(\alpha_j - 1)\left(\frac{I_{vv}(\omega_j)}{\mathcal{G}\omega_{\tilde{j}}^{-2d_k}} - 1\right)\right| \geq 1\right).$$

Now by (21),

$$\frac{1}{m_n}\sum_{j=1}^{m_n}(\alpha_j - 1)\left(\frac{I_{vv}(\omega_j)}{\mathcal{G}\omega_{\tilde{j}}^{-2d_k}} - 1\right) = \frac{1}{m_n}\sum_{j=1}^{m_n}(\alpha_j - 1)\left(\frac{I_{vv}(\omega_j)}{\mathcal{G}\omega_{\tilde{j}}^{-2d_k}} - \frac{I_{vv}(\omega_j)}{b'\mathbf{Af}\left(\omega_{\tilde{j}}\right)\mathbf{A}'b}\right)$$
$$+ \frac{1}{m_n}\sum_{j=1}^{m_n}(\alpha_j - 1)\frac{b'\mathbf{AR}(\omega_j)\mathbf{A}'b}{b'\mathbf{Af}\left(\omega_{\tilde{j}}\right)\mathbf{A}'b}$$
$$+ \frac{1}{m_n}\sum_{j=1}^{m_n}(\alpha_j - 1)\frac{b'\mathbf{AS}(\omega_j)\mathbf{A}'b}{b'\mathbf{Af}\left(\omega_{\tilde{j}}\right)\mathbf{A}'b}. \quad (39)$$

We will first show that the second and the third terms of (39) are $o_p(1)$. By (48) in the proof of Lemma 18, $b'\mathbf{AR}(\omega_j)\mathbf{A}'b = O_p\left(\omega_j^{-2d_k} j^{-\rho/2}\right)$. Thus by Corollary 11, the second term is of

$$O_p\left(\frac{1}{m_n}\sum_{j=1}^{m_n}(\alpha_j + 1) j^{-\rho/2}\right) = O_p\left(\frac{1}{m_n}\left(\sum_{j=1}^{m_n}\alpha_j^2\right)^{1/2} + m_n^{-\rho/2}\right) = o_p(1) \quad (40)$$

since $\sum_{j=1}^{m_n} \alpha_j^2 = O\left(m_n^{4(d_k - \Delta)} + m \log m\right)$ by equation 3.24 of Robinson (1995b). The third term of (39) is bounded by

$$\left|\frac{1}{m_n}\sum_{j=1}^{[e^v]}(\alpha_j - 1)\frac{b'\mathbf{AS}(\omega_j)\mathbf{A}'b}{b'\mathbf{Af}\left(\omega_{\tilde{j}}\right)\mathbf{A}'b}\right| + \left|\frac{1}{m_n}\sum_{j=[e^v]+1}^{m_n}(\alpha_j - 1)\frac{b'\mathbf{AS}(\omega_j)\mathbf{A}'b}{b'\mathbf{Af}\left(\omega_{\tilde{j}}\right)\mathbf{A}'b}\right|.$$

Following from Corollary 11 and (38), the first term of the above equation is of

$$O_p\left(\omega_{m_n}^{2(d_k - \Delta)}\frac{1}{m_n}\sum_{j=1}^{[e^v]}\omega_{\tilde{j}}^{2\Delta} b'\mathbf{AS}(\omega_j)\mathbf{A}'b\right) = O_p\left(\omega_{m_n}^{2(d_k - \Delta)}\mathcal{M}_{m_n}(\Delta)\right) = o_p\left(\omega_{m_n}^{2(d_k - \Delta)}\omega_{m_n}^{2(\Delta - d_k)}\right) = o_p(1).$$

by Lemma 19 since $0 \geq \Delta - d_k > -1/2$, and since $e^\nu \sim m_n/e$, the second term is

$$O_p\left(\frac{1}{m_n}\sum_{j=[e^v]+1}^{m_n}\left(\frac{j}{m_n}\right)^{2(\Delta_1 - d_k)}\frac{b'\mathbf{AS}(\omega_j)\mathbf{A}'b}{\omega_j^{-2d_k}}\right) = o_p(1)$$

by Lemma 20. The lemma will follow if the first term of (39) is also $o_p(1)$. By Lemma 21 and Corollary 12,

$$\frac{I_{vv}(\omega_j)}{\mathcal{G}\omega_{\tilde{j}}^{-2d_k}} = O_p(1)$$

and

$$1-\frac{\mathcal{G}\omega_{\tilde{j}}^{-2d_k}}{b'\mathbf{A}\mathbf{f}\left(\omega_{\tilde{j}}\right)\mathbf{A}'b}=O_p\left(\frac{\omega_{\tilde{j}}^{-2d_k}\left(j^{d_k-d_{k-1}}+\omega_{\tilde{j}}^{d_k-d_{k+1}}+\omega_{\tilde{j}}^{\rho}\right)}{\omega_{\tilde{j}}^{-2d_k}}\right)=O_p\left(j^{d_k-d_{k-1}}\right).$$

Thus, the first term of (39) is

$$\begin{aligned}\frac{1}{m_n}\sum_{j=1}^{m_n}(\alpha_j-1)\left(1-\frac{\mathcal{G}\omega_{\tilde{j}}^{-2d_k}}{b'\mathbf{A}\mathbf{f}\left(\omega_{\tilde{j}}\right)\mathbf{A}'b}\right)\frac{I_{vv}(\omega_j)}{\mathcal{G}\omega_{\tilde{j}}^{-2d_k}}&=O_p\left(\frac{1}{m_n}\sum_{j=1}^{m_n}(\alpha_j+1)\,j^{d_k-d_{k-1}}\right)\\&=O_p\left(\frac{1}{m_n}\left(\sum_{j=1}^{m_n}\alpha_j^2\right)^{1/2}+m_n^{d_k-d_{k-1}}\right)\\&=o_p(1)\end{aligned}$$

by the same argument for (40). We have completed the proof. □

Lemma 16 *Let $R_{ab}(\omega_j)$ be the (a,b)th entry of $\mathbf{R}(\omega_j)$,*

$$\mathbb{E}\left|R_{ab}(\omega_j)\right|\leq C\left|1-e^{-i\omega_{\tilde{j}}}\right|^{-(d_{aa}+d_{bb})}j^{-\rho/2},\quad a,b=1,\ldots,q\quad\text{and}\quad 1\leq j\leq[n/2]$$

under Assumption 2.

Proof. Let $J_{z_a}(\omega_j)$ be the jth element of $J_z(\omega_j)$, the DFT of z_t. By (4),

$$J_{z_a}(\omega_j)=\sum_{b=1}^{q}J_{z_{a_b}}(\omega_j) \tag{41}$$

where

$$J_{z_{a_b}}(\omega_j)=\frac{1}{\sqrt{2\pi\sum\left|h_t^{p-1}\right|^2}}\sum_{t=1}^{n}h_t^{p-1}\left(\sum_{k=\infty}^{\infty}\psi_{k,ab}\varepsilon_{t-k,b}\right)e^{i\omega_j t}.$$

Hence

$$\begin{aligned}R_{ab}(\omega_j)&=J_{z_a}(\omega_j)\overline{J}_{z_b}(\omega_j)-\sum_{u=1}^{q}\Psi_{au}\left(\omega_{\tilde{j}}\right)J_{\varepsilon_u}(\omega_j)\sum_{v=1}^{q}\overline{\Psi}_{bv}\left(\omega_{\tilde{j}}\right)\overline{J}_{\varepsilon_v}(\omega_j)\\&=\sum_{u=1}^{q}\sum_{v=1}^{q}\left(J_{z_{a_u}}(\omega_j)\overline{J}_{z_{b_v}}(\omega_j)-\Psi_{au}\left(\omega_{\tilde{j}}\right)J_{\varepsilon_u}(\omega_j)\overline{\Psi}_{bv}\left(\omega_{\tilde{j}}\right)\overline{J}_{\varepsilon_v}\right)\\&=\sum_{u=1}^{q}\sum_{v=1}^{q}\left(\Psi_{au}\left(\omega_{\tilde{j}}\right)\overline{\Psi}_{bv}\left(\omega_{\tilde{j}}\right)\left(A_{au,j}\overline{A}_{bv,j}-B_{u,j}\overline{B}_{v,j}\right)\right),\end{aligned} \tag{42}$$

where

$$A_{au,j}=\frac{J_{z_{a_u}}(\omega_j)}{\Psi_{au}\left(\omega_{\tilde{j}}\right)}\quad\text{and}\quad B_{u,j}=J_{\varepsilon_u}(\omega_j). \tag{43}$$

From Lemmas 9 and 10 of Hurvich et. al (2002),

$$\begin{aligned}\mathbb{E}\left|A_{au,j}-B_{u,j}\right|^{2\ell} &\leq C\left(\int_{-\pi}^{\pi}\left|\frac{\Psi_{au}(\omega)}{\Psi_{au}\left(\omega_{\tilde{j}}\right)}-1\right|^{2}\left|D_{p,n}\left(\omega_{j}-\omega\right)\right|^{2}d\omega\right)^{\ell} \\ &\leq Cj^{-\ell\rho}, \end{aligned} \tag{44}$$

Now

$$A_{au,j}\overline{A}_{bv,j}-B_{u,j}\overline{B}_{v,j}=\left(A_{au,j}-B_{u,j}\right)\left(\overline{A}_{bv,j}-\overline{B}_{v,j}\right)+B_{u,j}\left(\overline{A}_{bv,j}-\overline{B}_{v,j}\right)+\overline{B}_{v,j}\left(A_{au,j}-B_{u,j}\right).$$

By Cauchy Schwarz inequality,

$$\begin{aligned}&\mathbb{E}\left|A_{au,j}\overline{A}_{bv,j}-B_{u,j}\overline{B}_{v,j}\right|^{2} \\ &\leq 3\mathbb{E}\left|A_{au,j}-B_{u,j}\right|^{2}\left|\overline{A}_{bv,j}-\overline{B}_{v,j}\right|^{2}+\mathbb{E}\left|B_{u,j}\right|^{2}\left|\overline{A}_{bv,j}-\overline{B}_{v,j}\right|^{2}+\mathbb{E}\left|\overline{B}_{v,j}\right|^{2}\left|A_{au,j}-B_{u,j}\right|^{2} \\ &\leq 3\left(\mathbb{E}\left|A_{au,j}-B_{u,j}\right|^{4}\mathbb{E}\left|A_{bv,j}-B_{v,j}\right|^{4}\right)^{1/2}+\left(\mathbb{E}\left|B_{u,j}\right|^{4}\mathbb{E}\left|A_{bv,j}-B_{v,j}\right|^{4}\right)^{1/2}+\left(\mathbb{E}\left|B_{v,j}\right|^{4}\mathbb{E}\left|A_{au,j}-B_{u,j}\right|^{4}\right)^{1/2} \\ &\leq C\left[\left(j^{-2\rho}j^{-2\rho}\right)^{1/2}+\left(j^{-2\rho}\right)^{1/2}\right]=Cj^{-\rho}. \end{aligned} \tag{45}$$

We have from (42),

$$\begin{aligned}\mathbb{E}\left|R_{ab}\right| &\leq \sum_{u=1}^{q}\sum_{v=1}^{q}\left|\Psi_{au}\left(\omega_{\tilde{j}}\right)\overline{\Psi}_{bv}\left(\omega_{\tilde{j}}\right)\right|\left(\mathbb{E}\left|A_{au,j}\overline{A}_{bv,j}-B_{u,j}\overline{B}_{v,j}\right|^{2}\right)^{1/2} \\ &\leq C\sum_{u=1}^{q}\sum_{v=1}^{q}\left|1-e^{-i\omega_{\tilde{j}}}\right|^{-(d_{au}+d_{bv})}\tau_{au}\left(\omega_{j}\right)\tau_{av}\left(\omega_{j}\right)j^{-\rho/2} \\ &\leq C\sup_{\substack{\omega\in(0,\omega_{m_n}] \\ a,b=1,\ldots,q}}\tau_{ab}(\omega)\cdot\left|1-e^{-i\omega_{\tilde{j}}}\right|^{-(d_{aa}+d_{bb})}\sum_{u=1}^{q}\sum_{v=1}^{q}j^{-\rho/2} \\ &= C\left|1-e^{-i\omega_{\tilde{j}}}\right|^{-(d_{aa}+d_{bb})}j^{-\rho/2}\end{aligned}$$

where the constant C does not depend on n. □

Lemma 17 *Let $S_{ab}(\omega)$ be the (a,b)th entry of $\mathbf{S}(\omega_j)$. Then, for $1\leq j,k\leq[n/2]$*

$$\mathbb{E}\left|S_{ab}\left(\omega_{j}\right)S_{ab}\left(\omega_{k}\right)\right|\leq\begin{cases}C\left|\left(1-e^{-i\omega_{\tilde{j}}}\right)\left(1-e^{-i\omega_{\tilde{k}}}\right)\right|^{-(d_{aa}+d_{bb})}, & |j-k|<p \\ C/n, & \textit{otherwise}\end{cases}.$$

Proof: Note that

$$\mathbb{E}I_{\varepsilon\varepsilon}\left(\omega_{j}\right)=\mathbf{\Sigma},$$

and

$$S_{ab}\left(\omega_{j}\right)=\sum_{u=1}^{q}\sum_{v=1}^{q}\Psi_{au}\left(\omega_{\tilde{j}}\right)\overline{\Psi}_{bv}\left(\omega_{\tilde{j}}\right)\left(I_{\varepsilon\varepsilon,uv}\left(\omega_{j}\right)-\sigma_{uv}\right).$$

Now

$$\mathbb{E}\left|S_{ab}(\omega_j)S_{ab}(\omega_k)\right| = \sum_{u_1,u_2=1}^{q}\sum_{v_1,v_2=1}^{q}\Psi_{au_1}(\omega_{\tilde{k}})\overline{\Psi}_{au_2}(\widetilde{\omega}_k)\overline{\Psi}_{bv_1}\left(\omega_{\tilde{j}}\right)\Psi_{bv_2}(\omega_{\tilde{k}})$$
$$\times\,\mathbb{E}\left[\left(I_{\varepsilon\varepsilon,u_1v_1}(\omega_j)-\sigma_{u_1v_1}\right)\left(I_{\varepsilon\varepsilon,u_2v_2}(\omega_k)-\sigma_{u_2v_2}\right)\right].$$

Note that $\mathbb{E}\left(J_{\varepsilon_u}(\omega_j)J_{\varepsilon_v}(\omega_k)\right)=0$, $1\le j,k\le n/2$ and $\mathbb{E}\left(J_{\varepsilon_u}(\omega_j)\overline{J}_{\varepsilon_v}(\omega_k)\right)=0$ if $|j-k|\ge p$ and

$$\mathbb{E}\left(J_{\varepsilon_u}(\omega_j)\overline{J}_{\varepsilon_v}(\omega_k)\right)=\frac{\sigma_{uv}}{c_p}(-1)^{j-k}\begin{pmatrix}2p-2\\ p-1+j-k\end{pmatrix}\mathbf{1}_{\{|j-k|<p\}},\tag{46}$$

where

$$c_p=\begin{pmatrix}2p-2\\ p-1\end{pmatrix},$$

see Hurvich et. al (2002). Hence

$$\begin{aligned}&\mathbb{E}\left[\left(I_{\varepsilon\varepsilon,u_1v_1}(\omega_j)-\sigma_{u_1v_1}\right)\left(I_{\varepsilon\varepsilon,u_2v_2}(\omega_k)-\sigma_{u_2v_2}\right)\right]\\ &=\mathbb{E}\left[I_{\varepsilon\varepsilon,u_1v_1}(\omega_j)I_{\varepsilon\varepsilon,u_2v_2}(\omega_k)\right]-\sigma_{u_1v_1}\sigma_{u_2v_2}\\ &=\operatorname{cum}\left(J_{\varepsilon_{u_1}}(\omega_j),J_{\varepsilon_{u_2}}(\omega_k),\overline{J}_{\varepsilon_{v_1}}(\omega_j),\overline{J}_{\varepsilon_{v_2}}(\omega_k)\right)+\mathbb{E}\left(J_{\varepsilon_{u_1}}(\omega_j)\overline{J}_{\varepsilon_{v_2}}(\omega_k)\right)\mathbb{E}\left(J_{\varepsilon_{u_2}}(\omega_j)\overline{J}_{\varepsilon_{v_1}}(\omega_k)\right).\end{aligned}\tag{47}$$

The second term is nonzero only if $|j-k|<p$. We now show that the cumulant is of $O\left(n^{-1}\right)$. Let

$$\kappa_4=\operatorname{cum}\left(\varepsilon_{t,u_1},\varepsilon_{t,u_2},\varepsilon_{t,v_1},\varepsilon_{t,v_2}\right),\quad 1\le u_1,u_2,v_1,v_2\le q.$$

Then

$$\begin{aligned}&\operatorname{cum}\left(J_{\varepsilon_{u_1}}(\omega_j),J_{\varepsilon_{u_2}}(\omega_k),\overline{J}_{\varepsilon_{v_1}}(\omega_j),\overline{J}_{\varepsilon_{v_2}}(\omega_k)\right)\\ &=\kappa_4\left(\frac{1}{c_pn}\right)^2\int\int\int D_{n,p}(\omega_j-x_1)D_{n,p}(\omega_k-x_2)\overline{D}_{n,p}(\omega_j+x_3)\overline{D}_{n,p}(\omega_k-x_1-x_2-x_3)\,dx_1dx_2dx_3\\ &=\kappa_4\left(\frac{1}{c_pn}\right)^2\sum_{t_1,t_2,t_3,t_4=1}^{n}\sum_{\ell_1,\ell_2,\ell_3\ell_4=0}^{p-1}\binom{p-1}{\ell_1}\binom{p-1}{\ell_2}\binom{p-1}{\ell_3}\binom{p-1}{\ell_4}(-1)^{\ell_1+\ell_2+\ell_3+\ell_4}\\ &\quad\times e^{i\left(\omega_{j+\ell_1}t_1+\omega_{k+\ell_2}t_2-\omega_{j+\ell_3}t_3-\omega_{k+\ell_4}t_4\right)}\int\int\int e^{ix_1(t_1-t_4)}e^{ix_2(t_2-t_4)}e^{ix_3(t_3-t_4)}dx_1dx_2dx_3\\ &=\kappa_4\left(\frac{1}{c_pn}\right)^2\sum_{\ell_1,\ell_2,\ell_3\ell_4=0}^{p-1}\binom{p-1}{\ell_1}\binom{p-1}{\ell_2}\binom{p-1}{\ell_3}\binom{p-1}{\ell_4}(-1)^{\ell_1+\ell_2+\ell_3+\ell_4}\sum_{t=1}^{n}e^{i\omega_{\ell_1+\ell_2-\ell_3-\ell_4}t}\\ &\le\frac{\kappa_4}{nc_p}\left|\sum_{\ell=0}^{p-1}\binom{p-1}{\ell}\right|^4.\end{aligned}$$

Combining the above bound and equation (46) and (47), we have

$$\mathbb{E}\left[\left(I_{\varepsilon\varepsilon,u_1v_1}(\omega_j)-\sigma_{u_1v_1}\right)\left(I_{\varepsilon\varepsilon,u_2v_2}(\omega_k)-\sigma_{u_2v_2}\right)\right]=c\left(j,k,p,u_1,v_1,u_2,v_2\right)\mathbf{1}_{\{|j-k|<p\}}+O\left(n^{-1}\right)$$

and

$$\begin{aligned}
\mathbb{E}\left|S_{ab}\left(\omega_j\right)S_{ab}\left(\omega_k\right)\right| &\leq C\sum_{u_1,u_2=1}^{q}\sum_{v_1,v_2=1}^{q}\left(\left|\Psi_{au_1}\left(\omega_{\tilde{j}}\right)\right|\left|\overline{\Psi}_{au_2}\left(\omega_{\tilde{k}}\right)\right|\left|\overline{\Psi}_{bv_1}\left(\omega_{\tilde{j}}\right)\right|\left|\Psi_{bv_2}\left(\omega_{\tilde{k}}\right)\right|\right)\mathbf{1}_{\{|j-k|<p\}}+O\left(n^{-1}\right)\\
&\leq C\sup_{\substack{\omega\in(0,\omega_{m_n}]\\ a,b=1,\ldots,q}}\tau_{ab}\left(\omega\right)\cdot\left|1-e^{-i\omega_{\tilde{j}}}\right|^{-(d_{aa}+d_{bb})}\left|1-e^{-i\widetilde{\omega}_k}\right|^{-(d_{aa}+d_{bb})}\mathbf{1}_{\{|j-k|<p\}}+O\left(n^{-1}\right)\\
&= C\left|1-e^{-i\omega_{\tilde{j}}}\right|^{-(d_{aa}+d_{bb})}\left|1-e^{-i\omega_{\tilde{k}}}\right|^{-(d_{aa}+d_{bb})}\mathbf{1}_{\{|j-k|<p\}}+O\left(n^{-1}\right)
\end{aligned}$$

□

Lemma 18 *Let $\mathcal{L}_{m_n}(d)$ be defined in (22). Then for $d-d_k>-\frac{1}{2}$, there exists an $\epsilon>0$ such that*

$$\mathcal{L}_{m_n}\left(d\right)=o_p\left(\omega_{m_n}^{2d-2d_k}m_n^{-\epsilon}\right)$$

under Assumptions 1 and 2.

Proof. Let $\mathbf{R}_{h\ell}\left(\omega_j\right)$ be the (h,ℓ)th block matrix of $\mathbf{R}\left(\omega_j\right)$. By Lemmas 12 and 16,

$$b'\mathbf{A}_h\mathbf{R}_{h\ell}\left(\omega_j\right)\mathbf{A}_\ell'b=\begin{cases}O_p\left(n^{2d_k-d_h-d_\ell}\omega_j^{-d_h-d_\ell}j^{-\rho/2}\right), & h,\ell<k\\ O_p\left(\omega_j^{-d_h-d_\ell}j^{-\rho/2}\right) & h,\ell\geq k\\ O_p\left(n^{d_k-d_h}\omega_j^{-d_h-d_\ell}j^{-\rho/2}\right) & h<k,\ell\geq k\end{cases}\tag{48}$$

for $0\leq h,\ell\leq s$ and $1\leq j\leq m_n$. We have

$$\begin{aligned}
\mathcal{L}_{m_n}\left(d\right)&=\frac{1}{m_n}\sum_{j=1}^{m_n}\sum_{h=0}^{s}\sum_{\ell=0}^{s}\omega_{\tilde{j}}^{2d}b'\mathbf{A}_h\mathbf{R}_{h\ell}\left(\omega_j\right)\mathbf{A}_\ell'b\\
&=O_p\left(\frac{1}{m_n}\sum_{j=1}^{m_n}\omega_{\tilde{j}}^{2d}\left(n^{2d_k-2d_{k-1}}\omega_j^{-2d_{k-1}}j^{-\rho/2}+\omega_j^{-2d_k}j^{-\rho/2}+n^{d_k-d_{k-1}}\omega_j^{-d_{k-1}-d_k}j^{-\rho/2}\right)\right)\\
&=O_p\left(\frac{1}{m_n}\sum_{j=1}^{m_n}\omega_{\tilde{j}}^{2d-2d_k}j^{-\rho/2}\left(j^{2d_k-2d_{k-1}}+1+j^{d_k-d_{k-1}}\right)\right)\\
&=\begin{cases}O_p\left(\omega_{m_n}^{2d-2d_k}m_n^{2d_k-2d-1}\log m_n\right), & 2d-2d_k-\rho/2\leq 1\\ O_p\left(\omega_{m_n}^{2d-2d_k}m_n^{-\rho/2}\right), & 2d-2d_k-\rho/2>1\end{cases}.
\end{aligned}$$

□

Corollary 8 *For $d-d_k>-\frac{1}{4}$,*

$$\mathcal{L}_{m_n}\left(d\right)=o_p\left(\omega_{m_n}^{2d-2d_k}m_n^{-1/2-\epsilon}\right)$$

under the assumptions of Theorem 2.

Lemma 19 *Let $\mathcal{M}_{m_n}(d)$ be defined in (22). Then for $d-d_k>-\frac{1}{2}$, there exists an $\epsilon>0$ such that*

$$\mathcal{M}_{m_n}(d)=o_p\left(\omega_{m_n}^{2d-2d_k}m_n^{-\epsilon}\right)$$

under Assumptions 1 and 2.

Proof. Let $\mathbf{S}_{h\ell}(d)$ be the (h,ℓ)th block matrix of $\mathbf{S}(d)$. Following from Lemma 17,

$$\begin{aligned} E\left\|\frac{1}{m_n}\sum_{j=1}^{m_n}\omega_{\tilde{j}}^{2d}\mathbf{S}_{h\ell}(\omega_j)\right\|^2 &= O\left(\frac{1}{m_n^2}\sum_{j=1}^{m_n}\sum_{k=j}^{p+j}\omega_{\tilde{j}}^{2d-d_h-d_\ell}\omega_{\tilde{k}}^{2d-d_h-d_\ell}\right)\\ &= O\left(\frac{1}{m_n^2}\sum_{j=1}^{m_n}\sum_{k=j}^{p+j}\omega_{\tilde{j}}^{4d-2d_h-2d_\ell}\right)\\ &= \begin{cases} O\left(n^{2d_h+2d_\ell-4d}m_n^{-2}\log m_n\right), & 4d-2d_h-2d_\ell\le -1\\ O\left(\omega_{m_n}^{4d-2d_h-2d_\ell}m_n^{-1}\right), & 4d-2d_h-2d_\ell>-1\end{cases}.\end{aligned}$$

Hence we have

$$\left\|\frac{1}{m_n}\sum_{j=1}^{m_n}\omega_{\tilde{j}}^{2d}\mathbf{S}_{h\ell}(\omega_j)\right\| = \begin{cases} O_p\left(n^{d_h+d_\ell-2d}m_n^{-1}\log^{1/2}m_n\right), & 2d-d_h-d_\ell\le -1/2\\ O_p\left(\omega_{m_n}^{2d-d_h-d_\ell}m_n^{-1/2}\right), & 2d-d_h-d_\ell>-1/2\end{cases}. \tag{49}$$

Let

$$\mathcal{M}_{m_n}^{(h,\ell)}(d)=b'\mathbf{A}_h\left(\frac{1}{m_n}\sum_{j=1}^{m_n}\omega_{\tilde{j}}^{2d}\mathbf{S}_{h\ell}(\omega_j)\right)\mathbf{A}_\ell' b.$$

By Lemma 13 (subCointe2) and (49), we have, for $h,\ell<k$,

$$\begin{aligned}\mathcal{M}_{m_n}^{(h,\ell)}(d) &= \begin{cases} O_p\left(n^{-d_h-d_\ell+2d_k}n^{d_h+d_\ell-2d}m_n^{-1}\log^{1/2}m_n\right), & 2d-d_h-d_\ell\le -1/2\\ O_p\left(\omega_{m_n}^{2d-d_h-d_\ell}n^{-d_h-d_\ell+2d_k}m_n^{-1/2}\right), & 2d-d_h-d_\ell>-1/2\end{cases}\\ &= \begin{cases} O_p\left(\omega_{m_n}^{2d-2d_k}m_n^{-1-2d+2d_k}\log^{1/2}m_n\right), & 2d-d_h-d_\ell\le -1/2\\ o_p\left(\omega_{m_n}^{2d-2d_k}m_n^{-1/2+\varepsilon}\right), & 2d-d_h-d_\ell>-1/2\end{cases},\end{aligned} \tag{50}$$

where $\varepsilon>0$. By the same lemma and (49), we have, for $h,\ell\ge k$,

$$\mathcal{M}_{m_n}^{(h,\ell)}(d)=O_p\left(\omega_{m_n}^{2d-d_h-d_\ell}m_n^{-1/2}\right)=O_p\left(\omega_{m_n}^{2d-2d_k}\omega_{m_n}^{2d_k-d_h-d_\ell}m_n^{-1/2}\right), \tag{51}$$

and for $h<k$, $\ell\ge k$,

$$\begin{aligned}\mathcal{M}_{m_n}^{(h,\ell)}(d) &= \begin{cases} O_p\left(n^{-d_h+d_k}n^{d_h+d_\ell-2d}m_n^{-1}\log^{1/2}m_n\right), & 2d-d_h-d_\ell\le -1/2\\ O_p\left(\omega_{m_n}^{2d-d_h-d_\ell}n^{-d_h+d_k}m_n^{-1/2}\right), & 2d-d_h-d_\ell>-1/2\end{cases}\\ &= \begin{cases} o_p\left(\omega_{m_n}^{2d-2d_k}m_n^{-1-2d+2d_k}\log^{1/2}m_n\right), & 2d-d_h-d_\ell\le -1/2\\ o_p\left(\omega_{m_n}^{2d-2d_k}\omega_{m_n}^{d_k-d_\ell}m_n^{-1/2+d_k-d_h}\right), & 2d-d_h-d_\ell>-1/2\end{cases}.\end{aligned} \tag{52}$$

Hence,

$$\mathcal{M}_{m_n}(d) = \sum_{h=0}^{s}\sum_{\ell=0}^{s}\mathcal{M}_{m_n}^{(h,\ell)}(d) = o_p\left(\omega_{m_n}^{2d-2d_k} m_n^{-\epsilon}\right),$$

since $2d_k - d_h - d_\ell > 0$ in (51) and $-1-2d+2d_k < 0$ in (50) and (52). □

Corollary 9 *For* $d - d_k > -\frac{1}{4}$,

$$\mathcal{M}_{m_n}^{(h,\ell)}(d) = \begin{cases} O_p\left(\omega_{m_n}^{2d-2d_k} m_n^{-1/2}\right), & h=\ell=k \\ o_p\left(\omega_{m_n}^{2d-2d_k} m_n^{-1/2-\epsilon}\right), & \text{otherwise} \end{cases},$$

under the assumptions of Theorem 2.

Proof. The corollary follows by the fact that $-1-2d+2d_k < -1/2$ for $d-d_k > -\frac{1}{4}$ in (50) and (52).□

Lemma 20 *Let* δ *be a positive constant,* $\delta < 1$, *then*

$$\frac{1}{m_n}\sum_{j=\delta m_n}^{m_n}\left(\frac{j}{m_n}\right)^{2(\Delta_1-d_k)}\frac{b'\mathbf{A}\mathbf{S}(\omega_j)\mathbf{A}'b}{\omega_j^{-2d_k}} = o_p(1)$$

under the assumptions of Theorem 2.

Proof. Following the similar computation for (49)

$$\begin{aligned}\mathbb{E}\left\|\frac{1}{m_n}\sum_{j=\delta m_n}^{m_n}\left(\frac{j}{m_n}\right)^{2(\Delta_1-d_k)}\frac{\mathbf{S}_{h\ell}(\omega_j)}{\omega_j^{-2d_k}}\right\|^2 &= O\left(n^{-4d_k+2d_h+2d_\ell} m_n^{2(2d_k-2\Delta_1-1)}\sum_{j=\delta m_n}^{m_n} j^{4\Delta_1-2d_h-2d_\ell}\right) \\ &= O\left(\omega_{m_n}^{4d_k-2d_h-2d_\ell} m_n^{-1}\right).\end{aligned}$$

Hence,

$$\left\|\frac{1}{m_n}\sum_{j=\delta m_n}^{m_n}\left(\frac{j}{m_n}\right)^{2(\Delta_1-d_k)}\frac{\mathbf{S}_{h\ell}(\omega_j)}{\omega_j^{-2d_k}}\right\| = O_p\left(\omega_{m_n}^{2d_k-d_h-d_\ell} m_n^{-1/2}\right).$$

By Lemma 12, we have

$$\begin{aligned}b'\mathbf{A}_h\left(\frac{1}{m_n}\sum_{j=\delta m_n}^{m_n}\left(\frac{j}{m_n}\right)^{2(\Delta_1-d_k)}\frac{\mathbf{S}_{h\ell}(\omega_j)}{\omega_j^{-2d_k}}\right)\mathbf{A}'_\ell b &= \begin{cases} O_p\left(m_n^{2d_k-d_h-d_\ell-1/2}\right), & h,\ell < k \\ O_p\left(\omega_{m_n}^{2d_k-d_h-d_\ell} m_n^{-1/2}\right) & h,\ell \geq k \\ O_p\left(\omega_{m_n}^{d_k-d_\ell} m_n^{-1/2+d_k-d_h}\right) & h<k, \ell \geq k \end{cases} \\ &= o_p(1).\end{aligned}$$

The Lemma follows from the fact that

$$\left|\frac{1}{m_n}\sum_{j=\delta m_n}^{m_n}\left(\frac{j}{m_n}\right)^{2(\Delta_1-d_k)}\frac{b'\mathbf{A}\mathbf{S}(\omega_j)\mathbf{A}'b}{\omega_j^{-2d_k}}\right| \leq \sum_{h,\ell=0}^{s}\left|\frac{1}{m_n}\sum_{j=\delta m_n}^{m_n}\left(\frac{j}{m_n}\right)^{2(\Delta_1-d_k)}\frac{b'\mathbf{A}_h\mathbf{S}_{h\ell}(\omega_j)\mathbf{A}'_\ell b}{\omega_j^{-2d_k}}\right|.$$

□

Lemma 21 *Under Assumptions 1 and 2,*

$$b'\mathbf{A}\mathbf{f}(\omega_j)\mathbf{A}'b - b'\mathbf{A}_k\mathbf{f}_{kk}(\omega_j)\mathbf{A}_k'b = O_p\left(\omega_j^{-2d_k}\left(j^{d_k-d_{k-1}} + \omega_j^{d_k-d_{k+1}}\right)\right),$$

for $1 \le j \le m_n$, *and*

$$b'\mathbf{A}_k\mathbf{f}_{kk}(\omega_j)\mathbf{A}_k'b - \mathcal{G}\omega_j^{-2d_k} = O_p\left(\omega_j^{-2d_k+\rho}\right).$$

Furthermore, there exist two constants, C_1 *and* $\mathcal{C}_2$ *such that*

$$C_1 > \mathcal{G} \ge \mathcal{C}_2(1-\epsilon_k)$$

where $\epsilon_k = O_p(n^{-\alpha_k})$.

Proof. Since by Lemma 12,

$$b'\mathbf{A}_h\mathbf{f}_{h\ell}(\omega_j)\mathbf{A}_\ell'b = \begin{cases} O_p\left(\omega_j^{-2d_k}j^{2d_k-d_h-d_\ell}\right), & h<k, \ell \le k \\ O_p\left(\omega_j^{-d_h-d_\ell}\right), & h,\ell>k \\ O_p\left(\omega_j^{-d_k-d_\ell}j^{d_k-d_h}\right) & h\le k, \ell>k \end{cases}, \tag{53}$$

we have

$$\begin{aligned} b'\mathbf{A}\mathbf{f}(\omega_j)\mathbf{A}'b &= b'\mathbf{A}_k\mathbf{f}_{kk}(\omega_j)\mathbf{A}_k'b + \sum_{h=0}^{s}\sum_{\substack{\ell=0\\ \ell\neq k}}^{s} b'\mathbf{A}_h\mathbf{f}_{h\ell}(\omega_j)\mathbf{A}_\ell'b \\ &= b'\mathbf{A}_k\mathbf{f}_{kk}(\omega_j)\mathbf{A}_k'b + O_p\left(\omega_j^{-2d_k}j^{d_k-d_{k-1}} + \omega_j^{-2d_{k+1}} + \omega_j^{-d_k-d_{k+1}}\right) \\ &= b'\mathbf{A}_k\mathbf{f}_{kk}(\omega_j)\mathbf{A}_k'b + O_p\left(\omega_j^{-2d_k}\left(j^{d_k-d_{k-1}} + \omega_j^{d_k-d_{k+1}}\right)\right). \end{aligned}$$

Since (7) and Assumption 2 imply that

$$\mathbf{f}_{kk}(\omega) = \mathbf{f}_{kk}^{\dagger}(0)\,\omega^{-2d_k} + O\left(\omega^{-2d_k+\rho}\right) \quad \text{as} \quad \omega \to 0,$$

we have by Lemma 12,

$$\begin{aligned} b'\mathbf{A}\mathbf{f}_{kk}(\omega_j)\mathbf{A}_k b' &= b'\mathbf{A}_k\mathbf{f}_{kk}^{\dagger}(0)\mathbf{A}_k'b\omega_j^{-2d_k} + O_p\left(\|b'\mathbf{A}_k\|^2\omega_j^{-2d_k+\rho}\right) \\ &= \mathcal{G}\omega_j^{-2d_k} + O_p\left(\omega_j^{-2d_k+\rho}\right). \end{aligned}$$

Furthermore,

$$b'\mathbf{A}_k\mathbf{f}_{kk}^{\dagger}(0)\mathbf{A}_k'b\omega_j^{-2d_k} \ge \omega_j^{-2d_k}\lambda_{\min}\left(\mathbf{f}_{kk}^{\dagger}(0)\right)\|b'\mathbf{A}_k\|^2 \ge C\omega_j^{-2d_k}(1-\epsilon_k)$$

by Lemma 13. The upper bound for $\mathcal{G}$ is due to the fact that

$$\mathcal{G} \le \lambda_{\max}\left(\mathbf{f}_{kk}^{\dagger}(0)\right)\|\mathbf{A}_k\|\,\|b\| = \lambda_{\max}\left(\mathbf{f}_{kk}^{\dagger}(0)\right)\|\mathbf{A}_k\|.$$

□

Lemma 22 *Let $\mathcal{F}_{m_n}(d)$ be defined in (22). Then for $d - d_k > -\frac{1}{2}$, there exists an $\epsilon > 0$ such that*

$$\mathcal{F}_{m_n}(d) - \mathcal{G}\omega_{m_n}^{2d-2d_k} = o_p\left(\omega_{m_n}^{2d-2d_k} m_n^{-\epsilon}\right)$$

under the assumptions of Theorem 2.

Proof. Let the $a_h \times a_\ell$ matrix $\mathbf{f}_{h\ell}(\omega)$ be the (h,ℓ) block matrix of $\mathbf{f}(\omega)$, then

$$\begin{aligned}\mathcal{F}_{m_n}(d) &= \frac{1}{m_n}\sum_{j=1}^{m_n}\sum_{h,\ell=0}^{s}\omega_j^{2d} b'\mathbf{A}_h\mathbf{f}_{h\ell}(\omega_j)\mathbf{A}_\ell' b \\ &= \frac{1}{m_n}\sum_{j=1}^{m_n}\omega_j^{2d} b'\mathbf{A}_k\mathbf{f}_{kk}(\omega_j) b'\mathbf{A}_k + \frac{1}{m_n}\sum_{j=1}^{m_n}\sum_{h=0}^{s}\sum_{\substack{\ell=0\\ \ell\neq k}}^{s}\omega_j^{2d} b'\mathbf{A}_h\mathbf{f}_{h\ell}(\omega_j)\mathbf{A}_\ell' b. \end{aligned} \tag{54}$$

By Lemma 21,

$$\begin{aligned}\frac{1}{m_n}\sum_{j=1}^{m_n}\omega_j^{2d} b'\mathbf{A}_k\mathbf{f}_{kk}(\omega_j) b'\mathbf{A}_k &= \frac{\mathcal{G}}{m_n}\sum_{j=1}^{m_n}\omega_j^{2d-d_k} + O_p\left(\frac{1}{m_n}\sum_{j=1}^{m_n}\omega_j^{2d-2d_k+\rho}\right) \\ &= \mathcal{G}\omega_{m_n}^{2d-2d_k} + O_p\left(\omega_{m_n}^{2d-2d_k+\rho}\right)\end{aligned}$$

the second term of (54) is

$$\begin{aligned}&O_p\left(\frac{1}{m_n}\sum_{j=1}^{m_n}\omega_j^{2d-2d_k}\left(j^{d_k-d_{k-1}} + \omega_j^{d_k-d_{k+1}}\right)\right) \\ &= \begin{cases} O_p\left(\omega_{m_n}^{2d-2d_k}\left(m_n^{-1-2d+2d_k} + \omega_{m_n}^{d_k-d_{k+1}}\right)\right) & 2d-d_k-d_{k-1} \le -1 \\ O_p\left(\omega_{m_n}^{2d-2d_k}\left(m_n^{d_k-d_{k-1}} + \omega_{m_n}^{d_k-d_{k+1}}\right)\right) & 2d-d_k-d_{k-1} > -1 \end{cases}.\end{aligned}$$

□

Corollary 10 *Under the assumptions of Theorem 2,*

$$\mathcal{F}_{m_n}(d_k) - \mathcal{G} = O_p\left(m_n^{d_k-d_{k-1}} + \omega_{m_n}^{d_k-d_{k+1}}\right),$$

the $O_p\left(m_n^{d_k-d_{k-1}}\right)$ term is vacuous if $k=0$, and the $O_p\left(\omega_{m_n}^{d_k-d_{k+1}}\right)$ term is vacuous if $k=s$.

Corollary 11 *Under Assumptions 1 and 2, $b'\mathbf{Af}(\omega_j)\mathbf{A}'b \geq \mathcal{C}\omega_j^{-2d_k}(1-\epsilon_k)$*

Proof. By (53),

$$b'\mathbf{Af}\left(\omega_{\tilde{j}}\right)\mathbf{A}'b = b'\left(\sum_{h,\ell=0}^{k}\mathbf{A}_h\mathbf{f}_{h\ell}\left(\omega_{\tilde{j}}\right)\mathbf{A}_\ell'\right)b + O_p\left(\omega_j^{-d_k-d_{k+1}}\right).$$

By Assumption and Lemmas 12 and 13,

$$b'\left(\sum_{h,\ell=0}^{k}\mathbf{A}_h\mathbf{f}_{h\ell}\left(\omega_{\tilde{j}}\right)\mathbf{A}'_\ell\right)b = b'\left(\sum_{h,\ell=0}^{k}\omega_j^{-d_h-d_\ell}\mathbf{A}_h\mathbf{f}^{\dagger}_{h\ell}\left(0\right)\mathbf{A}'_\ell\right)b + O_p\left(\omega_j^{-\rho}\sum_{h,\ell=0}^{k}b'\mathbf{A}_h\mathbf{A}'_\ell b\right)$$

$$\geq \omega_j^{-2d_k}\lambda_{\min}\left\{\mathbf{f}^{\dagger}\left(0\right)\right\}\sum_{h,\ell=0}^{k}b'\mathbf{A}_h\mathbf{A}'_\ell b\mathbf{A}_h + O_p\left(\omega_j^{-\rho}\right)$$

$$\geq C\omega_j^{-2d_k}\left(1-\epsilon_k\right) + O_p\left(\omega_j^{-\rho}\right).$$

The corollary follows.□

Corollary 12 *Under Assumptions 1 and 2,* $I_{vv}\left(\omega_j\right) = O_p\left(\omega_j^{-2d_k}\right)$.

Proof. By (21), (48) and (53),

$$I_{vv}\left(\omega_j\right) = b'\mathbf{AR}\left(\omega_j\right)\mathbf{A}'b + b'\mathbf{AS}\left(\omega_j\right)\mathbf{A}'b + b'\mathbf{Af}\left(\omega_j\right)\mathbf{A}'b$$
$$= b'\mathbf{AS}\left(\omega_j\right)\mathbf{A}'b + O_p\left(\omega_j^{-2d_k}\right).$$

By Lemmas 12 and 17,

$$b'\mathbf{AS}\left(\omega_j\right)\mathbf{A}'b = O_p\left(\sum_{h,\ell=0}^{k-1}\omega_j^{-d_h-d_\ell}n^{d_{2k}-d_h-d_\ell} + \sum_{h,\ell=k}^{s}\omega_j^{-d_h-d_\ell} + \sum_{h<k,\ell\geq k}\omega_j^{-d_h-d_\ell}n^{d_k-d_h}\right)$$
$$= O_p\left(\omega_j^{-2d_k}\left(j^{2d_k-2d_{k-1}} + 1 + j^{d_k-d_{k-1}}\right)\right).$$

□

Lemma 23 *Define*

$$\widehat{F}_a\left(d\right) = \frac{1}{m_n}\sum_{j=c}^{m_n}\left(\log\tilde{j}\right)^a\frac{I_{vv}\left(\omega_j\right)}{\omega_{\tilde{j}}^{-2d}},$$

then for any $\tilde{d}$ *such that* $|\tilde{d}-d_k| \leq |\hat{d}_k - d_k|$,

$$\widehat{F}_a\left(\tilde{d}\right) = \widehat{F}_a\left(d_k\right) + o_p\left(1\right),$$

for $a = 0, 1, 2$, *under the assumptions of Theorem 2.*

Proof. Let

$$\widehat{E}_a\left(d\right) = \frac{1}{m_n}\sum_{j=c}^{m_n}\left(\log\tilde{j}\right)^a\tilde{j}^{2d}I_{vv}\left(\omega_j\right).$$

It is sufficient to show that

$$\widehat{E}_a\left(\tilde{d}\right) = \widehat{E}_a\left(d_k\right) + o_p\left(n^{2d_k}\right),$$

for $a = 0, 1, 2$. Let $M = \{d : \log^3 m_n \times |d - d_k| \le \epsilon\}$, where $\epsilon > 0$, is fixed to be such that $2\epsilon < \log^2 m_n$ with a proper n. Following the same lines of the proof on page 1642 in Robinson (1995b), for $\eta > 0$,

$$
\begin{aligned}
& P\left(\left|\widehat{E}_a\left(\tilde{d}\right) - \widehat{E}_a\left(d_k\right)\right| > \eta\left(\frac{2\pi}{n}\right)^{-2d_k}\right) \\
& \le P\left(\widehat{G}\left(d_k\right) > \frac{\eta}{2e\varepsilon}\left(\log m_n\right)^{2-a}\right) + P\left(\log^3 m_n\left|\tilde{d} - d_k\right| > \varepsilon\right). \qquad (55)
\end{aligned}
$$

The first probability is bounded by

$$
P\left(\left|\widehat{G}\left(d_k\right) - \mathcal{G}\right| > \frac{\eta}{4e\varepsilon}\left(\log m_n\right)^{2-a}\right) + P\left(\mathcal{G} > \frac{\eta}{4e\varepsilon}\left(\log m_n\right)^{2-a}\right).
$$

Both probabilities tend to 0 for ε sufficiently small since $\left|\widehat{G}\left(d_k\right) - \mathcal{G}\right| = o_p(1)$ and $\mathcal{G} < C$ by Lemma 21.

To show the second probability of (55) tending to 0, we only have to verify that

$$
\sup_{\Theta_1 \cap N_\delta}\left|\frac{\widehat{G}(d) - G(d)}{G(d)}\right| = o_p\left(\log^{-6} m_n\right).
$$

Following from (36) and (37) in the proof of Lemma 14,

$$
\sup_{\Theta_1 \cap N_\delta}\left|\frac{\widehat{G}(d) - G(d)}{G(d)}\right| \le \sup_{\Theta_1}\left|\frac{\widehat{G}(d) - G(d)}{G(d)}\right| = o_p\left(m_n^{-\epsilon}\right).
$$

□

Lemma 24 *Let $\tilde{d}$ be such that $\left|\tilde{d} - d_k\right| \le \left|\hat{d}_k - d_k\right|$, then under the assumptions of Theorem 2,*

$$
\frac{\partial^2 R\left(d_k\right)}{\partial d^2} \xrightarrow{p} 4.
$$

Proof. By Lemma 23, we have

$$
\begin{aligned}
\frac{\partial^2 R\left(\tilde{d}\right)}{\partial d^2} &= \frac{4\left\{\widehat{G}_2\left(\tilde{d}\right)\widehat{G}\left(\tilde{d}\right) - \widehat{G}_1^2\left(\tilde{d}\right)\right\}}{\widehat{G}^2(d)} \\
&= \frac{4\left\{\left(\widehat{F}_2\left(d_k\right) + o_p(1)\right)\left(\widehat{F}_0\left(d_k\right) + o_p(1)\right) - \left(\widehat{F}_1\left(d_k\right) + o_p(1)\right)^2\right\}}{\left(\widehat{F}_0\left(d_k\right) + o_p(1)\right)^2} \\
&= \frac{4\left\{\widehat{F}_2\left(d_k\right)\widehat{F}_0\left(d_k\right) - \widehat{F}_1^2\left(d_k\right)\right\}}{\widehat{F}_0^2\left(d_k\right)} + o_p(1) \quad \text{as} \quad n \to \infty.
\end{aligned}
$$

By Lemmas 18, 19 and 22,

$$
\begin{aligned}
\left|\widehat{F}_a\left(d_k\right) - \mathcal{G}\frac{1}{m_n}\sum_{j=1}^{m_n}\log^a\tilde{j}\right| &= \left|\frac{1}{m_n}\sum_{j=1}^{m_n}\log^a\tilde{j}\left(\frac{I_{vv}\left(\omega_j\right)}{\omega_{\tilde{j}}^{-2d_k}} - \mathcal{G}\right)\right| \\
&\le \log^a m_n\left|\mathcal{L}_{m_n}\left(d_k\right) + \mathcal{M}_{m_n}\left(d_k\right) + \mathcal{F}_{m_n}\left(d_k\right) - \mathcal{G}\right| \\
&= O_p\left(m_n^{-\epsilon}\log^a m_n\right).
\end{aligned}
$$

By the same reasoning as (4.10) in Robinson (1995b),

$$\frac{\partial^2 R\left(\tilde{d}\right)}{\partial d^2} = 4\left\{\frac{1}{m_n}\sum_{j=1}^{m_n}\log^2\tilde{j} - \left(\frac{1}{m_n}\sum_{j=1}^{m_n}\log\tilde{j}\right)^2\right\}(1+o_p(1)) + o_p(1) \xrightarrow{p} 4.$$

□

Lemma 25 *Under the assumptions of Theorem 3,*

$$m_n^{1/2}\frac{\partial R(d_k)}{\partial d} - \frac{1}{\mathcal{G}}b'\mathbf{A}_k\mathbf{\Psi}_k^{\dagger *}(0)\,\mathbf{Z}_n\mathbf{\Psi}_k^{\dagger}(0)\,\mathbf{A}_k'b = o_p(1),$$

where

$$\mathbf{Z}_n = 2m_n^{-1/2}\sum_{j=1}^{m_n}\nu_j\left(I_{\varepsilon\varepsilon}(\omega_j) - \mathbf{\Sigma}\right) \qquad \text{and} \qquad \nu_j = \log\tilde{j} - \frac{1}{m_n}\sum_{j=1}^{m_n}\log\tilde{j}.$$

Proof. Note that

$$\frac{\partial R(d)}{\partial d} = \frac{2}{m_n}\sum_{j=1}^{m_n}\frac{v_j I_{vv}(\omega_j)}{\omega_{\tilde{j}}^{-2d}\widehat{G}(d)}.$$

Since $\widehat{G}(d_k) - \mathcal{G} = o_p(1)$ by (37) and $\sum_{j=1}^{m_n}\nu_j = 0$,

$$m_n^{1/2}\frac{\partial R(d_k)}{\partial d} = 2m_n^{-1/2}\sum_{j=1}^{m_n}\nu_j\left(\frac{I_{vv}(\omega_j)}{\mathcal{G}\omega_{\tilde{j}}^{-2d_k}} - 1\right)(1+o_p(1)).$$

and

$$\begin{aligned}
&m_n^{1/2}\frac{\partial R(d_k)}{\partial d} - \frac{1}{\mathcal{G}}b'\mathbf{A}_k\mathbf{\Psi}_k^{\dagger *}(0)\,\mathbf{Z}_n\mathbf{\Psi}_k^{\dagger}(0)\,\mathbf{A}_k'b \\
&= 2m_n^{-1/2}\left\{\sum_{j=1}^{m_n}\nu_j\left(\frac{I_{vv}(\omega_j)}{\mathcal{G}\omega_{\tilde{j}}^{-2d_k}} - 1\right) - \sum_{j=1}^{m_n}\frac{\nu_j}{\mathcal{G}\omega_{\tilde{j}}^{-2d_k}}b'\mathbf{A}_k\mathbf{S}_{kk}(\omega_j)\,\mathbf{A}_k'b\right\} \\
&+ 2m_n^{-1/2}\left\{\sum_{j=1}^{m_n}\frac{\nu_j}{\mathcal{G}}\left(\omega_{\tilde{j}}^{2d_k}b'\mathbf{A}_k\mathbf{S}_{kk}(\omega_j)\,\mathbf{A}_k'b\right) - b'\mathbf{A}_k\mathbf{\Psi}_k^{\dagger *}(0)\,\mathbf{Z}_n\mathbf{\Psi}_k^{\dagger}(0)\,\mathbf{A}_k'b\right\} + o_p(1),
\end{aligned} \tag{56}$$

where $\mathbf{S}_{h\ell}(\omega_j)$ is the (h,ℓ)th block of $\mathbf{S}(\omega_j)$ defined in (21). Let

$$\mathcal{M}_{m_n}^{(h,\ell)}(d) = \frac{1}{m_n}\sum_{j=1}^{m_n}\omega_{\tilde{j}}^{2d}b'\mathbf{A}_h\mathbf{S}_{h\ell}(\omega_j)\,\mathbf{A}_\ell'b,$$

the first term of (56) is

$$\left| 2m_n^{-1/2} \sum_{j=1}^{m_n} \frac{\nu_j}{\mathcal{G}\omega_{\tilde{j}}^{-2d_k}} \left\{ I_{vv}(\omega_j) - \mathcal{G}\omega_{\tilde{j}}^{-2d_k} + \sum_{\substack{u,v=0 \\ v\neq k}}^{q} b'\mathbf{A}_h \mathbf{S}_{h\ell}(\omega_j) \mathbf{A}_\ell' b \right\} \right|$$

$$\leq \frac{2m_n^{1/2} \log m_n}{\mathcal{G}} \left| \mathcal{L}_{m_n}(d_k) + \mathcal{F}_{m_n}(d_k) - \mathcal{G} + \sum_{\substack{u,v=0 \\ v\neq k}}^{q} \mathcal{M}_{m_n}^{(h,\ell)}(d_k) \right|$$

$$= o_p(1),$$

by (21), (22), Corollaries 8, 9, 10 and Assumption 3b. Since

$$\begin{aligned} \mathbf{S}_{kk}(\omega_j) &= \left|1 - e^{i\omega_j}\right|^{-2d_k} \mathbf{\Psi}_k^{\dagger *}(\omega_j) \left[I(\omega_j) - \mathbf{\Sigma}\right] \mathbf{\Psi}_k^{\dagger}(\omega_j) \\ &= \omega_{\tilde{j}}^{-2d_k} \mathbf{\Psi}_k^{\dagger *}(0) \left[I(\omega_j) - \mathbf{\Sigma}\right] \mathbf{\Psi}_k^{\dagger}(0) + O_p\left(\omega_{\tilde{j}}^{-2d_k+\rho}\right), \end{aligned}$$

the second term in (56) is

$$O_p\left(m_n^{-1/2} \sum_{j=1}^{m_n} \nu_j \omega_{\tilde{j}}^{\rho}\right) = O_p\left(\frac{m_n^{\rho+1/2} \log m}{n^{\rho}}\right) = o_p(1)$$

by Lemma 12 and Assumption 3b. We have shown that both terms on the RHS of (56) are $o_p(1)$ hence completed the proof.□

Lemma 26 *Let* $\mathbf{U}_n$ *and* $\mathbf{V}_n$ *be defined as in Equation (10),* $\mathbf{U}_n$, $\mathbf{V}_n$ *and* $\mathbf{\Xi}$ *as in Lemma 1, and* $\mathbf{Z}_n$ *as in Lemma 25. Under Assumptions 1 and 2,*

$$\text{vec}(\mathbf{Z}_n, \mathbf{U}_n, \mathbf{V}_n) \xrightarrow{D} \text{vec}(\mathbf{Z}, \mathbf{U}, \mathbf{V})$$

where

$$\text{vec}(\mathbf{Z}, \mathbf{U}, \mathbf{V}) \sim N(0, \mathbf{\Delta})$$

and

$$\mathbf{\Delta} = \begin{bmatrix} 2\pi\Phi_p \mathbf{\Sigma} \otimes \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Xi} \end{bmatrix},$$

hence vec($\mathbf{Z}$) *and* vec$(\mathbf{U}, \mathbf{V})$ *are independent.*

Proof. Note that

$$\mathbb{E}\left[\text{vec}(\mathbf{Z}_n) \text{vec}(\mathbf{U}_n, \mathbf{V}_n)'\right] = \mathbf{0},$$

for every n since by Assumption 1, each element of the LHS is the product of three zero mean normal random variables. By Lemma 28 and the corollary of Theorem 25.12 of Billingsley (1995),

$$\mathbb{E}\left[\text{vec}(\mathbf{Z}) \text{vec}(\mathbf{U}, \mathbf{V})'\right] = \mathbf{0}.$$

The lemma follows from the Cramer-Wold device. □

Lemma 28 gives the uniform integrability of $\mathrm{vec}\left(\mathbf{Z}_n\right)\mathrm{vec}\left(\mathbf{U}_n,\mathbf{V}_n\right)'$ so that the previous lemma will follow. The proof of Lemma 28 requires a bound for the covariance of the normalized DFTs at ω_k and ω_j, $k<j$. Although Lemma 12 of Hurvich et. al (2002) provides such a bound for all d and p, their bound tends to infinity for $d\in(-p+1/2,-p+1)$ when k is fixed and $p\geq 2$. We obtain a new bound for $d<0$ and $p\geq 2$ in the next lemma which gives a sufficient bound for our purposes when combined with Lemma 12 of Hurvich et. al (2002).

Lemma 27 *Let $\{\xi_t\}$ be a univariate process with spectral density*

$$f(\omega)=\frac{\sigma^2}{2\pi}\left|1-e^{-i\omega}\right|^{-2d}\left|\alpha^*(\omega)\right|^2$$

where $d\in(-p+1,0)$, $p\geq 2$ and $\alpha^*(\omega)\in\mathcal{L}^*(\mu,\rho)$ for some $\mu>1$ and $\rho\in(1,2]$, and $\mathcal{L}^*$ is defined in Assumption 2. Let $J(\omega)$ be the tapered DFT of $\{\xi_t\}_{t=1}^n$. Then for $d<0$ and $1\leq k<j\leq[(n-1)/2]$,

$$\left|\mathbb{E}\left(\frac{J(\omega_k)}{f^{1/2}(\omega_{\tilde{k}})}\frac{J(\omega_j)}{f^{1/2}(\omega_{\tilde{j}})}\right)\right|+\left|\mathbb{E}\left(\frac{J(\omega_k)}{f^{1/2}(\omega_{\tilde{k}})}\frac{\overline{J}(\omega_j)}{f^{1/2}(\omega_{\tilde{j}})}\right)\right|\leq Cj^dk^d(j+k)^{-2d-p}.$$

Proof. We will only derive this bound for the first term on the LHS, since the derivation for the other term is similar. Let $x=(\omega_j+\omega_k)/2$, $y=(\omega_j-\omega_k)/2$ so that $\omega_j=x+y$ and $\omega_k=x-y$. Then

$$\begin{aligned}\mathbb{E}\left(J(\omega_k)J(\omega_j)\right)&=\int_{-\pi}^{\pi}f(\omega)D_{n,p}(\omega_j-\omega)D_{n,p}(\omega_k-\omega)\,d\omega\\&=\int_0^{\pi}\left[f(\omega+y)+f(\omega-y)\right]D_{n,p}(x-\omega)D_{n,p}(x+\omega)\,d\omega\\&:=\int_0^{\pi}g(y;\omega)\Delta_n(x;\omega)\,dx.\end{aligned}$$

Let $W_1=\{\omega:0<\omega<x/2\}$, $W_2=\{\omega:x/2<\omega<\min(3x/2,\pi)\}$ and $W_3=\{\omega:\min(3x/2,\pi)<\omega<\pi\}$,

$$g(y;\omega)=f(\omega-y)+f(\omega+y)$$

and

$$\Delta_n(x;\omega)=D_{n,p}(x-\omega)D_{n,p}(x+\omega).$$

Then

$$\mathbb{E}\left(\frac{J(\omega_k)}{f^{1/2}(\omega_{\tilde{k}})}\frac{J(\omega_j)}{f^{1/2}(\omega_{\tilde{j}})}\right)=f^{-1/2}(\omega_{\tilde{j}})f^{-1/2}(\omega_{\tilde{k}})\int_0^{\pi}g(y;\omega)\Delta_n(x;\omega)\,dx=\sum_{a=1}^{3}q_a(x,y),$$

where

$$q_a(x,y)=f^{-1/2}(\omega_{\tilde{j}})f^{-1/2}(\omega_{\tilde{k}})\int_{W_a}g(y;\omega)\Delta_n(x;\omega)\,d\omega.$$

By Assumption 2,

$$|g(y;\omega)|\leq C\left(|\omega-y|^{-2d}+|\omega+y|^{-2d}\right).$$

Also by Lemma 0 of Hurvich and Chen (2000),

$$|\Delta_n(x;\omega)|\leq Cn\left(1+n|x-\omega|\right)^{-p}\left(1+n|x+\omega|\right)^{-p},\tag{57}$$

for $-\alpha\pi < x-\omega,\ x+\omega < \alpha\pi,\ \ 0<\alpha<2.$

Since $\omega_{(j+k)/4} < x-\omega < \omega_{(j+k)/2}$ and $\omega_{(j+k)/2} < x+\omega < \omega_{3(j+k)/2}$ for $\omega \in W_1$,

$$\Delta_n(x;\omega) \le Cn(j+k)^{-2p},$$

by (57) and since $\omega_{(j-k)/4} - \omega_{k/2} < y-\omega < y = \omega_{(j-k)/2}$ and $\omega_{(j-k)/2} < y+\omega < \omega_{j/2} + \omega_{(j+k)/4}$ for $\omega \in W_1$,

$$\begin{aligned} |q_1(x,y)| &\le C\omega_j^d\omega_k^d n(j+k)^{-2p}\int_0^{x/2}\left(|\omega-y|^{-2d} + |\omega+y|^{-2d}\right)d\omega \\ &\le C\omega_j^d\omega_k^d n(j+k)^{-2p}\,\omega_{(j+k)/4}\left(\left|\omega_{(j-k)/2}\right|^{-2d} + \omega_{k/2}^{-2d} + \left|\omega_{j/2}+\omega_{(j+k)/4}\right|^{-2d}\right) \\ &\le C\omega_j^d\omega_k^d n(j+k)^{-2p}\,\omega_{(j+k)}^{1-2d} \le Cj^dk^d(j+k)^{-2p+1-2d}. \end{aligned} \tag{58}$$

For $\omega \in W_2$,

$$\Delta_n(x;\omega) \le Cn(j+k)^{-p}(1+n|x-\omega|)^{-p}$$

because $\omega_{3(j+k)/4} < x+\omega < \min\left\{\omega_{5(j+k)/4}, \pi+\omega_{(j+k)/2}\right\}$. Furthermore, since $\int_{-\infty}^{\infty}(1+n|x|)^{-p}dx \le cn^{-1}$ and $|g(y;\omega)| \le C\{\omega_{5j+k}^{-2d} + \omega_{j+5k}^{-2d}\} \le C\omega_{j+k}^{-2d}$, we have

$$|q_2(j,k)| \le C\omega_j^d\omega_k^d\omega_{j+k}^{-2d} n(j+k)^{-p}\int_{-\infty}^{\infty}(1+n|x-\omega|)^{-p}d\omega \le Cj^dk^d(j+k)^{-2d-p}. \tag{59}$$

For $\omega \in W_3 = \{\omega : \min(3x/2,\pi) < \omega < \pi\}$ and assuming that $3x/2 < \pi$, we have $\omega_{5(j+k)/4} < x+\omega < \pi + \omega_{(j+k)/2}$. We further break q_3 into two integrals,

$$q_3^{(1)}(j,k) = \int_{W_3}[g(y;\omega) - g(y;x)]\,\Delta_n(x;\omega)\,\mathbf{1}_{\{x+\omega\le\pi\}}d\omega$$

and

$$q_3^{(2)}(j,k) = \int_{W_3}[g(y;\omega) - g(y;x)]\,\Delta_n(x;\omega)\,\mathbf{1}_{\{x+\omega>\pi\}}d\omega.$$

Since $|x-\omega| \ge \frac{1}{3}\omega$ for $\omega \in W_3$,

$$\left|\Delta_n(x;\omega)\,\mathbf{1}_{\{x+\omega\le\pi\}}\right| \le Cn(1+n\omega)^{-p}(1+n(x+\omega))^{-p} \le Cn^{-2p+1}\omega^{-2p}. \tag{60}$$

Because $D_{n,p}(\cdot)$ is 2π-periodic and $|D_{n,p}(\cdot)|$ is symmetric around $-\omega_{p-1}$,

$$|D_{n,p}(x+\omega)| = |D_{n,p}(2\pi - (x+\omega) - \omega_{p-1})|.$$

Hence

$$\begin{aligned} \left|\Delta_n(x;\omega)\,\mathbf{1}_{\{x+\omega>\pi\}}\right| &\le Cn(1+n|x-\omega|)^{-p}(1+n|2\pi-(x+\omega)|)^{-p} \\ &\le Cn(1+n|x-\omega|)^{-p}(1+n|x-\omega|)^{-p} \\ &\le Cn^{-2p+1}\omega^{-2p}, \end{aligned} \tag{61}$$

since $|2\pi - (x+\omega)| \ge |x-\omega|$ for $x,\omega \in [0,\pi]$. Since $|\omega+y| < 2\omega$ and $|y-\omega| = \omega - y < \omega$ for $\omega \in W_3$,

$$g(y;\omega) \le C\omega^{-2d}.$$

Thus for $b \in \{1,2\}$, we have by (60), (61) and the above equation,

$$\left|q_3^{(b)}(j,k)\right| \le Cn^{-2p+1}\omega_j^d\omega_k^d\int_{\omega_{3(j+k)/4}}^{\infty}\omega^{-2d-2p}d\omega \le Cn^{-2p+1}\omega_j^d\omega_k^d\omega_{j+k}^{-2d-2p+1} = Cj^dk^d(j+k)^{-2d-2p+1}. \tag{62}$$

Combining (58), (59) and (62),

$$\mathbb{E}\left|\frac{J(\omega_k)}{f^{1/2}(\omega_{\tilde{k}})}\frac{J(\omega_j)}{f^{1/2}(\omega_{\tilde{j}})}\right| \le Cj^dk^d(j+k)^{-2d-p}\left\{1+(j+k)^{-p+1}\right\} \le Cj^dk^d(j+k)^{-2d-p}.$$

□

Corollary 13 *Let $\{\xi_t\}$ be the process defined in Lemma 27 with $d \in (-p+1/2, 1/2)$. For $k < K < \infty$, and $1 \le k \le (j-p+1) \le [(n-1)/2]$,*

$$\mathbb{E}\left|\frac{J(\omega_k)}{f^{1/2}(\omega_{\tilde{k}})}\frac{J(\omega_j)}{f^{1/2}(\omega_{\tilde{j}})}\right| \le \begin{cases} C\log(1+j)\,j^{|d|-1}, & d \ge 0. \\ Cj^{|d|-p}, & d < 0. \end{cases}.$$

Proof. By Lemma 12 of Hurvich, et. al (2002), the LHS is bounded by

$$\begin{cases} Ck^{-|d|}j^{|d|-1}\log(1+j), & p=1 \\ C(j-k)^{-p}h^{-p}j\left(\frac{k}{j}\right)^d + (j-h)^{-p+1}\left\{h^{-1}\left(\frac{j}{h}\right)^d + j^{-1}\left(\frac{h}{j}\right)^d\right\}, & p \ge 2 \end{cases},$$

which reduces to

$$C\log(1+j)\,j^{|d|-1}$$

for $d \ge 0$. If we use this together with Lemma 27 , the corollary follows.□

Lemma 28 *Let $J_{z_a}(\omega_h)$ be the ath element of $J_z(\omega_h)$ and $Z_{n,uv}$ be the (u,v)th entry of $\mathbf{Z}_n$. Then under the assumptions of Theorem 3,*

$$\mathbb{E}\left(n^{-2d_{aa}}I_{zz,aa}(\omega_h)Z_{n,uv}^2\right) < \infty,$$

for $1 \le h < m+p$ and $1 \le a,u,v \le q$.

Proof. We will use the notation $J_{z_{a_b}}(\omega_j)$, $A_{ab,j}$, and $B_{u,j}$ defined in (41) and (43). Following from (41),

$$\begin{aligned}
\mathbb{E}\left(n^{-2d_{aa}}I_{zz,aa}(\omega_h)Z_{n,uv}^2\right) &= \frac{1}{n^{2d_{aa}}m_n}\sum_{b_1,b_2=1}^{q}\sum_{j_1,j_2=1}^{m_n}\nu_{j_1}\nu_{j_2}\mathbb{E}\left[J_{z_{a_{b_1}}}(\omega_h)\overline{J}_{z_{a_{b_2}}}(\omega_h)I_{\varepsilon\varepsilon,uv}(\omega_{j_1})I_{\varepsilon\varepsilon,uv}(\omega_{j_2})\right] \\
&= \frac{1}{n^{2d_{aa}}m_n}\sum_{b_1,b_2=1}^{q}\Psi_{ab_1}(\omega_{\tilde{h}})\overline{\Psi}_{ab_2}(\omega_{\tilde{h}}) \\
&\qquad \times \sum_{j_1,j_2=1}^{m_n}\nu_{j_1}\nu_{j_2}\mathbb{E}\left[A_{ab_1,h}\overline{A}_{ab_2,h}B_{u,j_1}\overline{B}_{v,j_1}B_{u,j_2}\overline{B}_{v,j_2}\right] \\
&\le \frac{C}{m_n}\sum_{b_1,b_2=1}^{q}\left|\sum_{j_1,j_2=1}^{m_n}\nu_{j_1}\nu_{j_2}\mathbb{E}\left[A_{ab_1,h}\overline{A}_{ab_2,h}B_{u,j_1}\overline{B}_{v,j_1}B_{u,j_2}\overline{B}_{v,j_2}\right]\right|,
\end{aligned}$$

since $n^{-2d_{aa}}\left|\Psi_{ab_1}\left(\omega_{\tilde{h}}\right)\overline{\Psi}_{ab_2}\left(\omega_{\tilde{h}}\right)\right|\le C$ by (5). It is sufficient to show that

$$\frac{1}{m_n}\sum_{j_1,j_2=1}^{m_n}\nu_{j_1}\nu_{j_2}\mathbb{E}\left[A_{ab_1,h}\overline{A}_{ab_2,h}B_{u,j_1}\overline{B}_{v,j_1}B_{u,j_2}\overline{B}_{v,j_2}\right]<\infty \tag{63}$$

for $1\le b_1,b_2\le q$. Since $A_{ab,j}$ and $B_{u,j}$ are zero mean normal random variables. The expectation in (63) is a sum of 15 products of three expectations which are all the possible pair-partitions of the 6 random variables. Thus the LHS of (63) can be decomposed accordingly. Many terms of this decomposition are zero since $\mathbb{E}\left[B_{u,j_1}B_{u,j_2}\right]=0$ and $\sum_j v_j=0$. Here we show the finiteness for only two nonzero terms, since the rest of the terms can be handled analogously. By Corollary 13, for $1\le a,b,u\le q$, $h<j$, $h\le m+p$ and $j=1,\dots,m_n$,

$$\left|\mathbb{E}\left(A_{ab,h}A_{au,j}\right)\right|+\left|\mathbb{E}\left(A_{ab,h}\overline{A}_{av,j}\right)\right|=\begin{cases}C\log\left(1+j\right)j^{d-1}, & d\ge 0.\\ Cj^{|d|-p}, & d<0.\end{cases},$$

Note that by (45), $\mathbb{E}\left|A_{ab_1,h}\overline{A}_{ab_2,h}\right|\le C_h$. Applying the above bound together with (44), we have

$$\begin{aligned}\mathbb{E}\left(A_{ab,h}B_{u,j}\right)&=\mathbb{E}\left(A_{ab,h}A_{au,j}\right)+\mathbb{E}\left[A_{ab,h}\left(B_{u,j}-A_{au,j}\right)\right]\\&\le C\left(j^{|d|+\epsilon-1}\mathbf{1}\left\{d\ge 0\right\}+Cj^{|d|-p}\right)+\left(\mathbb{E}\left|A_{ab,h}\right|^2\mathbb{E}\left|B_{u,j}-A_{au,j}\right|^2\right)^{1/2}\\&<Cj^{|d|+\epsilon-1}\mathbf{1}\left\{d\ge 0\right\}+Cj^{|d|-p}+j^{-\rho/2}=O\left(j^{-1/2-\varepsilon}\right)\end{aligned} \tag{64}$$

where $\epsilon>0$. Let $c=m+p$,

$$\begin{aligned}&\frac{1}{m_n}\sum_{j_1,j_2=1}^{m_n}\nu_{j_1}\nu_{j_2}\mathbb{E}\left[A_{ab_1,h}B_{u,j_1}\right]\mathbb{E}\left[\overline{A}_{ab_2,h}\overline{B}_{v,j_2}\right]\mathbb{E}\left[\overline{B}_{v,j_1}B_{u,j_2}\right]\\&\le C\frac{\log^2 m_n}{m_n}\sum_{j_1=c}^{m_n}\sum_{j_2=c}^{m_n}\left|\mathbb{E}\left[A_{ab_1,h}B_{u,j_1}\right]\mathbb{E}\left[\overline{A}_{ab_2,h}\overline{B}_{v,j_2}\right]\mathbb{E}\left[\overline{B}_{v,j_1}B_{u,j_2}\right]\right|\\&+C\sum_{j_1=1}^{c-1}\sum_{j_2=1}^{m_n}\left|\mathbb{E}\left[A_{ab_1,h}B_{u,j_1}\right]\mathbb{E}\left[\overline{A}_{ab_2,h}\overline{B}_{v,j_2}\right]\mathbb{E}\left[\overline{B}_{v,j_1}B_{u,j_2}\right]\right|\\&=O\left(\frac{\log m_n^2}{m_n}\sum_{j_1,j_2=c}^{m_n}j_1^{-1/2-\varepsilon}j_2^{-1/2-\varepsilon}\mathbf{1}_{\{|j_1-j_2|<p\}}+\frac{\log m_n}{m_n}\right)=O\left(\frac{\log m_n^2}{m_n}\right),\end{aligned}$$

since the summation over $j_1,j_2\in\{1,\dots,c-1\}$ is $O\left(m_n^{-1}\right)$. Also

$$\begin{aligned}&\frac{1}{m_n}\sum_{j_1,j_2=1}^{m_n}\nu_{j_1}\nu_{j_2}\mathbb{E}\left[A_{ab_1,h}\overline{A}_{ab_2,h}\right]\mathbb{E}\left[B_{u,j_1}\overline{B}_{v,j_2}\right]\mathbb{E}\left[\overline{B}_{v,j_1}B_{u,j_2}\right]\\&=C_h\operatorname{var}\left(Z_{uv}\right)<\infty.\end{aligned}$$

The proof is completed.□

The next corollary is a corollary of Lemma 4.

Corollary 14 $\left\|\mathbf{X}_k\left(\mathbf{H}\right)-\mathbf{X}_k\right\|_F=O_p\left(n^{-\alpha_k}\right).$

Proof. Note that since both $\mathbf{H}$ and I_m are symmetric, a multiplication of -1 to $\chi_j(\mathbf{H})$ or χ_j will still retain the properties. We can assume that $\chi_j'\chi_j(\mathbf{H}) \geq 0$. We have

$$\begin{aligned}\|\mathbf{X}_k(\mathbf{H}) - \mathbf{X}_k\|_F^2 &\leq a_k \max_{j \in N_k} \|\chi_j(\mathbf{H}) - \chi_j\|^2 \leq C \max_{j \in N_k} \sin^2 \theta(\chi_j, \chi_j(\mathbf{H})) \\ &\leq C \|\sin\Theta\{\mathcal{M}(\mathbf{X}_k(\mathbf{H})), \mathcal{M}(\mathbf{X}_k)\}\|_F^2\end{aligned}$$

by the definition of the $\sin\Theta$ bound. □

Lemma 29 *Under Assumption 1, the eigenvectors of* $\mathbf{H}$,

$$\chi_j(\mathbf{H}) \xrightarrow{D} \mathring{\xi}_j(\operatorname{vec}(\mathbf{U}, \mathbf{V})), \quad j = 1, \ldots, q,$$

where $\mathring{\xi}_j$ *are continuous functions of* $\operatorname{vec}(\mathbf{U}, \mathbf{V})$ *and* $\mathbf{U}, \mathbf{V}$ *are defined in Lemma 1.*

Proof. Since $\chi_j(\mathbf{H}) = \mathbf{B}'\chi_j(\mathbf{\Phi}_D)$, it is sufficient to show that

$$\chi_j(\mathbf{\Phi}_D) \xrightarrow{D} \mathring{\varsigma}_j(\operatorname{vec}(\mathbf{U}, \mathbf{V})), \quad j = 1, \ldots, q,$$

where $\mathring{\varsigma}_j$ are continuous functions of $\operatorname{vec}(\mathbf{U}, \mathbf{V})$. Let $\widetilde{\mathbf{\Phi}}_D = \mathbf{d}_n^{-1}\mathbf{\Phi}_D\mathbf{d}_n^{-1}$

$$\widetilde{\mathbf{\Phi}}_D = \mathbf{X}'\left(\widetilde{\mathbf{\Phi}}_D\right)\mathbf{\Lambda}\left(\widetilde{\mathbf{\Phi}}_D\right)\mathbf{X}\left(\widetilde{\mathbf{\Phi}}_D\right).$$

First note that under Assumption 1, the eigenvalues of $\widetilde{\mathbf{\Phi}}_D$ are distinct with probability 1 by Okamoto (1973). Since both $\widetilde{\mathbf{\Phi}}_D$ and $\mathbf{\Phi}_D$ are block diagonal matrices,

$$\mathbf{\Phi}_D = \mathbf{d}_n^{-1}\mathbf{X}'\left(\widetilde{\mathbf{\Phi}}_D\right)\mathbf{\Lambda}\left(\widetilde{\mathbf{\Phi}}_D\right)\mathbf{X}\left(\widetilde{\mathbf{\Phi}}_D\right)\mathbf{d}_n^{-1} = \mathbf{X}'\left(\widetilde{\mathbf{\Phi}}_D\right)\mathbf{d}_n^{-1}\mathbf{\Lambda}\left(\widetilde{\mathbf{\Phi}}_D\right)\mathbf{d}_n^{-1}\mathbf{X}\left(\widetilde{\mathbf{\Phi}}_D\right).$$

This implies that

$$\mathbf{X}'(\mathbf{\Phi}_D) = \mathbf{X}'\left(\widetilde{\mathbf{\Phi}}_D\right) \quad \text{and} \quad \mathbf{\Lambda}(\mathbf{\Phi}_D) = \mathbf{d}_n^{-1}\mathbf{\Lambda}\left(\widetilde{\mathbf{\Phi}}_D\right)\mathbf{d}_n^{-1}.$$

Let $\mathbf{K}$ be defined as in Lemma 6 and we rewrite $\mathbf{U}_n$ and $\mathbf{V}_n$ in (10) as

$$\mathbf{U}_n = \left(\mathbf{U}_n^{(0)}, \ldots, \mathbf{U}_n^{(s)}\right)' \text{ and } \mathbf{V}_n = \left(\mathbf{V}_n^{(0)}, \ldots, \mathbf{V}_n^{(s)}\right)'$$

where $\mathbf{U}_n^{(k)}$ and $\mathbf{V}_n^{(k)}$ are $a_k \times m$ matrices. Since $\mathbf{K}$ is a block diagonal matrix,

$$\begin{aligned}\widetilde{\mathbf{\Phi}}_D &= \mathbf{d}_n^{-1}\mathbf{K}'\mathbf{d}_n \operatorname{diag}\left(\begin{array}{ccc}\mathbf{U}_n^{(0)}\mathbf{U}_n^{(0)\prime} & \cdots & \mathbf{U}_n^{(s)}\mathbf{U}_n^{(s)\prime}\end{array}\right)\mathbf{d}_n\mathbf{K}\mathbf{d}_n^{-1} \\ &= \mathbf{K}' \operatorname{diag}\left(\begin{array}{ccc}\mathbf{U}_n^{(0)}\mathbf{U}_n^{(0)\prime} & \cdots & \mathbf{U}_n^{(s)}\mathbf{U}_n^{(s)\prime}\end{array}\right)\mathbf{K}.\end{aligned}$$

It follows that

$$\chi_j(\mathbf{\Phi}_D) = \chi_j\left(\widetilde{\mathbf{\Phi}}_D\right) := \mathring{\varsigma}_j(\operatorname{vec}(\mathbf{U}_n, \mathbf{V}_n)) \xrightarrow{D} \mathring{\varsigma}_j(\operatorname{vec}(\mathbf{U}, \mathbf{V})),$$

and

$$\chi_j(\mathbf{H}) = \mathbf{B}'\chi_j(\mathbf{\Phi}_D) \xrightarrow{D} \mathbf{B}'\mathring{\varsigma}_j(\operatorname{vec}(\mathbf{U}, \mathbf{V})).$$

□

Remark 3 *Let* $\mathring{\mathbf{X}}(\mathbf{\Phi}_D) = [\mathring{\varsigma}_1(\mathrm{vec}(\mathbf{U},\mathbf{V})), \ldots, \mathring{\varsigma}_s(\mathrm{vec}(\mathbf{U},\mathbf{V}))]$ *and* $\mathring{\mathbf{X}}(\mathbf{H}) = \left[\mathring{\xi}_1(\mathrm{vec}(\mathbf{U},\mathbf{V})), \ldots, \mathring{\xi}_s(\mathrm{vec}(\mathbf{U},\mathbf{V}))\right] = \left[\begin{array}{ccc} \mathring{\mathbf{X}}_0(\mathbf{H}) & \cdots & \mathring{\mathbf{X}}_s(\mathbf{H}) \end{array}\right]$. *Since* $\mathring{\mathbf{X}}(\mathbf{\Phi}_D)$ *is a diagonal block matrix,* $\mathring{\mathbf{X}}_k(\mathbf{H}) = \mathbf{B}'\mathring{\mathbf{X}}_k(\mathbf{\Phi}_D) \in \mathcal{B}_k$.

Corollary 15 *Let* b *be a* $q \times 1$ *vector in* $\mathcal{M}(\mathbf{X}_k)$ *with length one, then under Assumption 1,*

$$b \xrightarrow{D} \mathring{b} := \mathring{\xi}(\mathrm{vec}(\mathbf{U},\mathbf{V})),$$

where $\mathring{\xi}$ *is a continuous function with respect to* $\mathrm{vec}(\mathbf{U},\mathbf{V})$ *and* $\mathbf{U},\mathbf{V}$ *are defined in Lemma 1.*

Proof. The corollary follows from Lemma 29 and Corollary 14 and $b = c'\mathbf{X}_k$ for some $a_k \times 1$ vector c.□

References

[1] Anderson, T.W. and Das Gupta, S. (1963). "Some inequalities on characteristic roots of matrices," *Biometrika* **50**, 522-524.

[2] Barlow, J.L. and Slapničar, I. (2000). "Optimal perturbation bounds for the Hermitian eigenvalue problem," *Linear Algebra and its Applications* **309**, 19-43.

[3] Billingsley, P. (1986). *Probability and Measure*, 2'nd Ed., New York: Wiley.

[4] Chen, W.W. and Hurvich, C.M. (2003a). "Estimating fractional cointegration in the presence of polynomial trends," *Journal of Econometrics* **117**, 95-121.

[5] Chen, W.W. and Hurvich, C.M. (2003b). "Semiparametric estimation of multivariate fractional cointegration," *Journal of the American Statistical Association* **463**, 629-642.

[6] Davies, R.B. and Harte, D.S. (1987). "Tests for Hurst effect," *Biometrika* **74**, 95 -102.

[7] Davis, C. and Kahan, W.M. (1970). "The rotation of eigenvectors by a perturbation. III," *SIAM J. Numer. Anal.* **7**, 1-46.

[8] Engle, R.F., and Granger, C.W.J. (1987). "Co-integration and error correction: representation, estimation and testing," *Econometrica* **55**, 251-276.

[9] Fujikoshi, Y. (1985). "Two methods for estimation of dimensionality in canonical correlation analysis and the multivariate linear model." In: Matusita, K. (Ed.), Statistical Theory and Data Analysis. Elsevier, Amsterdam, pp. 233-240.

[10] Fujikoshi, Y. and Veitch, L.G. (1979). "Estimation of dimensionality in canonical correlation analysis," *Biometrika* **66**, 345-351.

[11] Geweke, J, Porter-Hudak, S.(1983). "The estimation and application of long memory time series models." *J. Time Ser. Anal.* **4**, 221-238.

[12] Gunderson, B.K. and Muirhead, R.J. (1997). "On estimating the dimensionality in canonical correlation analysis," *Journal of Multivariate Analysis* **62**, 121-136.

[13] Hausman, J. (1978). "Misspecification tests in econometrics," *Econometrica* **46**, 1251-1271.

[14] Hurvich, C.M. and Chen, W.W. (2000). "An efficient taper for potentially overdifferenced long-memory time series," *J. Time Ser. Anal.* **21**, 155-180.

[15] Hurvich, C.M., Moulines E and Soulier, P (2002). "The FEXP estimator for potentially non-stationary linear time series," *Stochastic Processes and their Applications* **97**, 307 340.

[16] Künsch, H. R. (1987). "Statistical aspects of self-similar processes." Proceedings of the World Congress of the Bernoulli Society, Tashkent, Vol. 1, 67–74.

[17] Lobato, I.G. (1999). "A semiparametric two-step estimation on a multivariate long memory model," *Journal of Econometrics* **90**, 129-153.

[18] Magnus, I.R. and Neudecker, H. (1999). *Matrix differential calculus with applications in statistics and econometrics*, Chichester: Wiley.

[19] Marinucci, D. and Robinson, P.M. (2001). "Semiparametric fractional cointegration analysis," *J. Econometrics* **105,** 225-247.

[20] Okamoto, M. (1973). "Distinctness of the eigenvalues of a quadratic form in a multivariate sample," *Annals of Statistics* **1**, 763-765.

[21] Rao, C.R. (1973). *Linear Statistical Inference and its Applications*, 2'nd Ed. New York: Wiley.

[22] Robinson, P.M. (1994). "Semiparametric analysis of long-memory time series." *Annals of Statistics* **22**, 515-539.

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

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

[25] Robinson, P.M. and Marinucci, D. (2003). "Semiparametric Frequency-Domain Analysis of Fractional Cointegration," *Time Series with Long Memory*, Ed. P.M. Robinson, Oxford University Press, 334-373.

[26] Robinson, P.M. and Marinucci, D. (2001). "Narrow-band analysis of nonstationary processes," *Ann. Stat.* **29**, 947-86.

[27] Robinson, P.M. and Yajima, Y. (2002). "Determination of cointegrating rank in fractional systems," *J. Econometrics* **106**, 217-241.

[28] Siskind, V. (1972). "Second moments of inverse Wishart-matrix elements," *Biometrika* **59**, 690-691.

[29] Stewart, G.W. and Sun, J. (1990). *Matrix Perturbation Theory.* San Diego: Academic Press.

[30] Terrin, N. and Hurvich, C.M. (1994). "An asymptotic Wiener-Ito representation for the low frequency ordinates of the periodogram of a long memory time series." *Stochastic Processes and their Applications* **54,** 297-307.

[31] Terrin, N. and Taqqu, M.S. (1991). "Convergence in distribution of sums of bivariate Appell polynomials with long-range dependence. " *Probability Theory and Related Fields* **90,** 57-81.

[32] Velasco, C. (2003). "Gaussian semi-parametric estimation of fractional cointegration." *Journal of Time Series Analysis* **24**, 345-378.

[33] Wedin, P.A. (1983). "On angles between subspaces of a finite dimensional inner product space." *Lecture Notes in Mathematics* **973,** 263-285. New York: Springer-Verlag.