Econometrica Supplementary Material

SUPPLEMENT TO "SPURIOUS FACTOR ANALYSIS"

(*Econometrica*, Vol. 89, No. 2, March 2021, 591–614)

ALEXEI ONATSKI
Faculty of Economics, University of Cambridge

CHEN WANG
Department of Statistics and Actuarial Science, University of Hong Kong

This note contains supplementary material for Onatski and Wang (2021) (OW in what follows). It is lined up with sections in the main text to make it easy to locate the required proofs.

S1. INTRODUCTION

THERE IS NO SUPPLEMENTARY MATERIAL for this section of OW.

S2. BASIC SETUP AND MAIN RESULTS

In this section, we provide proofs of Lemmas OW5 and OW6 referred to in the proof of Theorem OW1.

S2.1. *Proof of Lemma OW5*

Note that

$$UMU' = \begin{pmatrix} 0 & 0 \\ 0 & U_{T-1} \end{pmatrix} M \begin{pmatrix} 0 & 0 \\ 0 & U'_{T-1} \end{pmatrix},$$

where U_{T-1} is the $T-1$-dimensional upper triangular matrix of ones. Denoting the $T-1$-dimensional vector of ones as l_{T-1}, we obtain

$$UMU' = \begin{pmatrix} 0 & 0 \\ 0 & U_{T-1}(I_{T-1} - l_{T-1}l'_{T-1}/T)U'_{T-1} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & Q \end{pmatrix}.$$

We have

$$Q^{-1} = (U'_{T-1})^{-1}(I_{T-1} + l_{T-1}l'_{T-1})(U_{T-1})^{-1}.$$

On the other hand, $(U_{T-1})^{-1}$ is a two-diagonal matrix with 1 on the main diagonal and -1 on the super-diagonal. Therefore, Q^{-1} is a three-diagonal matrix with 2 on the main diagonal, and -1 on the sub- and super-diagonals. As is well known (e.g., Sargan and Bhargava (1983)), the eigenvalues of such a three-diagonal matrix, indexed in the increasing order, are $\mu_k = 2 - 2\cos(\varpi_k/2)$, $k = 1, \dots, T-1$, where $\varpi_k = 2\pi k/T$. The corresponding (normalized) eigenvectors are $\bar{v}_k = (\bar{v}_{k1}, \dots, \bar{v}_{k,T-1})'$ with $\bar{v}_{kj} = \sqrt{2/T}\sin(j\varpi_k/2)$. This implies that the singular values of MU' (in decreasing order) are

$$\sigma_k = \sqrt{\mu_k^{-1}} = \big(2\sin(\varpi_k/4)\big)^{-1}$$

Alexei Onatski: ao319@cam.ac.uk
Chen Wang: stacw@hku.hk

 https://doi.org/10.3982/ECTA16703

for $k = 1, \ldots, T-1$ and $\sigma_T = 0$, and the components of the corresponding normalized right singular vectors are

$$v_{ks} = \sqrt{2/T}\sin\bigl((s-1)\varpi_k/2\bigr), \quad s = 1, \ldots, T$$

for $k = 1, \ldots, T-1$; and $v_{Ts} = 1$ for $s = 1$ and $v_{Ts} = 0$ for $s > 1$. Notice that v_{ks}, $s = 1, \ldots, T$, are proportional to the values at $(s-1)/T$ of the kth principal eigenfunction of the covariance operator of the Brownian bridge process (e.g., Shorack and Wellner (1986, pp. 213–214)).

To find the kth left singular vectors w_k with $k < T$, we multiply MU' by $\sigma_k^{-1}v_k$. We have $w_k = 2\sin(\varpi_k/4)MU'v_k$. On the other hand, the jth element of $U'v_k$ equals

$$\sqrt{2/T}\,\mathrm{Im}\sum_{s=0}^{j-1} e^{\mathrm{i}s\varpi_k/2} = \sqrt{2/T}\,\mathrm{Im}\,\frac{e^{\mathrm{i}j\varpi_k/2}-1}{e^{\mathrm{i}\varpi_k/2}-1}.$$

Therefore, $\sqrt{T/2}$ times the jth element of $MU'v_k$ equals

$$\mathrm{Im}\,\frac{e^{\mathrm{i}j\varpi_k/2}-1}{e^{\mathrm{i}\varpi_k/2}-1} - \frac{1}{T}\,\mathrm{Im}\sum_{j=1}^{T}\frac{e^{\mathrm{i}j\varpi_k/2}-1}{e^{\mathrm{i}\varpi_k/2}-1} = \mathrm{Im}\,\frac{e^{\mathrm{i}j\varpi_k/2}}{e^{\mathrm{i}\varpi_k/2}-1} = -\frac{\cos\bigl((2j-1)\varpi_k/4\bigr)}{2\sin(\varpi_k/4)}.$$

Hence,

$$w_{ks} = -\sqrt{2/T}\cos\bigl((s-1/2)\varpi_k/2\bigr), \quad s = 1, \ldots, T,$$

for $k < T$. Clearly, the left singular vector of MU' corresponding to zero singular value equals $w_T = \sqrt{1/T}\,l_T$.

REMARK: From (OW7), we see that w_{ks} with $s = 1, \ldots, T$ and $k < T$ are proportional to the values at $(s-1/2)/T$ of the kth principal eigenfunction of the covariance operator of the demeaned Wiener process.

S2.2. *Proof of Lemma OW6*

Our proof of Lemma OW6 is based on the following result.

LEMMA S1: *Suppose Assumption* A1 *of OW holds. Let a, b, c, d and A, B be any deterministic T-dimensional vectors and $N_\varepsilon \times N_\varepsilon$ matrices, respectively. Then*

$$\mathbb{E}\bigl(a'\varepsilon' A\varepsilon b\bigr) = a'b\,\mathrm{tr}\,A, \quad \textit{and} \tag{S1}$$

$$\begin{aligned}&\bigl|\mathrm{Cov}\bigl(a'\varepsilon' A\varepsilon b, c'\varepsilon' B\varepsilon d\bigr) - \bigl(a'c\bigr)\bigl(b'd\bigr)\mathrm{tr}\bigl(A'B\bigr) - \bigl(a'd\bigr)\bigl(b'c\bigr)\mathrm{tr}(AB)\bigr| \\ &\quad \le 2\varkappa_4\sum_{i=1}^{N_\varepsilon}\sum_{t=1}^{T}|A_{ii}B_{ii}a_t b_t c_t d_t|,\end{aligned} \tag{S2}$$

where a_t, b_t, c_t, and d_t are the tth components of vectors a, b, c, and d.

First, let us derive Lemma OW6 from Lemma S1. After such a derivation, we will prove Lemma S1. Since W is positive semi-definite, we have

$$\bigl(\|W\|/\,\mathrm{tr}\,W\bigr)^2 \le \mathrm{tr}\bigl(W^2\bigr)/(\mathrm{tr}\,W)^2 \le \|W\|/\,\mathrm{tr}\,W.$$

Further, $\|W\| = \|\Omega\|$ and $\operatorname{tr} W = \operatorname{tr} \Omega$. Therefore, Assumption A3 of OW is equivalent to the requirement

$$\operatorname{tr}(W^2) = o(1)(\operatorname{tr} W)^2. \tag{S3}$$

The first and second equalities of Lemma OW6 follow from Lemma S1, Chebyshev's inequality, and (S3). The last equality follows from the fact that $\sum_{r=1}^{T-1} \sigma_r^2 = O(T^2)$.

Now, let us turn to the proof of Lemma S1. We have

$$\begin{aligned}\mathbb{E}(a'\varepsilon' A\varepsilon b) &= \sum_{t,s=1}^{T} \sum_{i,j=1}^{N_\varepsilon} \mathbb{E}(a_t \varepsilon_{it} A_{ij} \varepsilon_{js} b_s) \\ &= \sum_{t=1}^{T} \sum_{i=1}^{N_\varepsilon} a_t A_{ii} b_t = a'b \operatorname{tr} A.\end{aligned}$$

Further, denoting the ith row of ε as $\varepsilon_{i\cdot}$, we have

$$\begin{aligned}&\mathbb{E}(a'\varepsilon' A\varepsilon bc'\varepsilon' B\varepsilon d) \\ &\quad= \sum_{i,j=1}^{N_\varepsilon} \sum_{p,l=1}^{N_\varepsilon} \mathbb{E}(\varepsilon_{i\cdot} a A_{ij} \varepsilon_{j\cdot} b \varepsilon_{p\cdot} c B_{pl} \varepsilon_{l\cdot} d) \\ &\quad= \sum_{i=1}^{N_\varepsilon} \sum_{j\neq i}^{N_\varepsilon} \mathbb{E}\big((\varepsilon_{i\cdot} a)(\varepsilon_{i\cdot} b)(\varepsilon_{j\cdot} c)(\varepsilon_{j\cdot} d) A_{ii} B_{jj}\big) \\ &\qquad+ \sum_{i=1}^{N_\varepsilon} \sum_{j\neq i}^{N_\varepsilon} \mathbb{E}\big((\varepsilon_{i\cdot} a)(\varepsilon_{i\cdot} c)(\varepsilon_{j\cdot} b)(\varepsilon_{j\cdot} d) A_{ij} B_{ij}\big) \\ &\qquad+ \sum_{i=1}^{N_\varepsilon} \sum_{j\neq i}^{N_\varepsilon} \mathbb{E}\big((\varepsilon_{i\cdot} a)(\varepsilon_{i\cdot} d)(\varepsilon_{j\cdot} b)(\varepsilon_{j\cdot} c) A_{ij} B_{ji}\big) \\ &\qquad+ \sum_{i=1}^{N_\varepsilon} \mathbb{E}\big((\varepsilon_{i\cdot} a)(\varepsilon_{i\cdot} b)(\varepsilon_{i\cdot} c)(\varepsilon_{i\cdot} d) A_{ii} B_{ii}\big).\end{aligned}$$

We have, first,

$$\begin{aligned}&\sum_{i=1}^{N_\varepsilon} \sum_{j\neq i}^{N_\varepsilon} \mathbb{E}\big((\varepsilon_{i\cdot} a)(\varepsilon_{i\cdot} b)(\varepsilon_{j\cdot} c)(\varepsilon_{j\cdot} d) A_{ii} B_{jj}\big) \\ &\quad= \sum_{i=1}^{N_\varepsilon} \sum_{j\neq i}^{N_\varepsilon} (a'b)(c'd) A_{ii} B_{jj} \\ &\quad= (a'b)(c'd)\left[(\operatorname{tr} A)(\operatorname{tr} B) - \sum_{i=1}^{N_\varepsilon} A_{ii} B_{ii}\right],\end{aligned}$$

second,

$$\sum_{i=1}^{N_\varepsilon}\sum_{j\neq i}^{N_\varepsilon}\mathbb{E}\big((\varepsilon_{i\cdot}a)(\varepsilon_{i\cdot}c)(\varepsilon_{j\cdot}b)(\varepsilon_{j\cdot}d)A_{ij}B_{ij}\big)$$
$$=\sum_{i=1}^{N_\varepsilon}\sum_{j\neq i}^{N_\varepsilon}\big(a'c\big)\big(b'd\big)A_{ij}B_{ij}$$
$$=\big(a'c\big)\big(b'd\big)\left[\operatorname{tr}\big(A'B\big)-\sum_{i=1}^{N_\varepsilon}A_{ii}B_{ii}\right],$$

third,

$$\sum_{i=1}^{N_\varepsilon}\sum_{j\neq i}^{N_\varepsilon}\mathbb{E}\big((\varepsilon_{i\cdot}a)(\varepsilon_{i\cdot}d)(\varepsilon_{j\cdot}b)(\varepsilon_{j\cdot}c)A_{ij}B_{ji}\big)$$
$$=\sum_{i=1}^{N_\varepsilon}\sum_{j\neq i}^{N_\varepsilon}\big(a'd\big)\big(b'c\big)A_{ij}B_{ji}$$
$$=\big(a'd\big)\big(b'c\big)\left[\operatorname{tr}(AB)-\sum_{i=1}^{N_\varepsilon}A_{ii}B_{ii}\right],$$

and finally,

$$\sum_{i=1}^{N_\varepsilon}\mathbb{E}\big((\varepsilon_{i\cdot}a)(\varepsilon_{i\cdot}b)(\varepsilon_{i\cdot}c)(\varepsilon_{i\cdot}d)A_{ii}B_{ii}\big)$$
$$=\sum_{i=1}^{N_\varepsilon}\mathbb{E}\left(\sum_{t=1}^{T}\varepsilon_{it}a_t\sum_{t=1}^{T}\varepsilon_{it}b_t\sum_{t=1}^{T}\varepsilon_{it}c_t\sum_{t=1}^{T}\varepsilon_{it}d_tA_{ii}B_{ii}\right)$$
$$=\sum_{i=1}^{N_\varepsilon}A_{ii}B_{ii}\left(\sum_{t,s:t\neq s}^{T}a_tb_tc_sd_s+\sum_{t,s:t\neq s}^{T}a_tb_sc_td_s+\sum_{t,s:t\neq s}^{T}a_tb_sc_sd_t\right.$$
$$\left.+\sum_{t=1}^{T}\mathbb{E}\varepsilon_{it}^4a_tb_tc_td_t\right)$$
$$=\sum_{i=1}^{N_\varepsilon}A_{ii}B_{ii}\left(\big(a'b\big)\big(c'd\big)+\big(a'c\big)\big(b'd\big)+\big(a'd\big)\big(b'c\big)\right.$$
$$\left.+\sum_{t=1}^{T}\big(\mathbb{E}\varepsilon_{it}^4-3\big)a_tb_tc_td_t\right).$$

Summing up,

$$\begin{aligned}&\mathbb{E}\big(a'\varepsilon' A\varepsilon b c'\varepsilon' B\varepsilon d\big)\\&\quad= \big(a'b\big)\big(c'd\big)(\operatorname{tr} A)(\operatorname{tr} B) + \big(a'c\big)\big(b'd\big)\operatorname{tr}\big(A'B\big) + \big(a'd\big)\big(b'c\big)\operatorname{tr}(AB)\\&\qquad+\sum_{i=1}^{N_\varepsilon} A_{ii}B_{ii}\sum_{t=1}^{T}\big(\mathbb{E}\varepsilon_{it}^4-3\big)a_t b_t c_t d_t.\end{aligned}$$

Recall that $\mathbb{E}(a'\varepsilon' A\varepsilon b) = a'b\operatorname{tr} A$ and $\mathbb{E}(c'\varepsilon' B\varepsilon d) = c'd\operatorname{tr} B$. These equalities and the last display yield

$$\begin{aligned}\operatorname{Cov}\big(a'\varepsilon' A\varepsilon b, c'\varepsilon' B\varepsilon d\big) &= \big(a'c\big)\big(b'd\big)\operatorname{tr}\big(A'B\big) + \big(a'd\big)\big(b'c\big)\operatorname{tr}(AB)\\&\quad+\sum_{i=1}^{N_\varepsilon} A_{ii}B_{ii}\sum_{t=1}^{T}\big(\mathbb{E}\varepsilon_{it}^4-3\big)a_t b_t c_t d_t.\end{aligned}$$

The inequality (S2) follows because $|\mathbb{E}\varepsilon_{it}^4-3|$ is bounded by $2\varkappa_4$ uniformly over i and t. Indeed, by Assumption A1, $\mathbb{E}\varepsilon_{it}^4 \le \varkappa_4$, and $\mathbb{E}\varepsilon_{it}^4-3\le\varkappa_4$. On the other hand, $\mathbb{E}\varepsilon_{it}^4 \ge (\mathbb{E}\varepsilon_{it}^2)^2 = 1$, and thus, $\varkappa_4\ge 1$ and $\mathbb{E}\varepsilon_{it}^4-3\ge -2\ge -2\varkappa_4$.

S3. EXTENSIONS

S3.1. *Local Level Model*

S3.1.1. *Proof of Theorem OW2*

Consider the decomposition

$$YM = \omega_T XM + ZM, \tag{S4}$$

where $X = [X_1, \dots, X_T]$ and $Z = [Z_1, \dots, Z_T]$. Note that the principal eigenvalues and eigenvectors of $MX'XM/N$ satisfy Theorem OW1(i)–(ii) as long as condition (OW5) of that theorem holds. Statements (i) and (ii) of Theorem OW2 would follow from this fact and the standard perturbation theory (e.g., Kato (1980, Chapter 2)) if we are able to show that $\|MZ'ZM\| = \omega_T^2 T^2 \operatorname{tr}\Omega o_{\mathrm{P}}(1)$.

Assumption A4 of OW yields $\sum_{k=0}^\infty \|\Pi_k\| = O(N^\beta)$ and $k\|\Pi_k\| \le \sum_{k=0}^\infty (1+k)\|\Pi_k\| = O(N^\beta)$. Therefore, $\sum_{k=0}^\infty k\|\Pi_k\|^2 = O(N^{2\beta})$ and, as explained in Remark OW8, we can apply Lemma OW7 to obtain

$$\|ZM\| \le \|Z\| = O_{\mathrm{P}}\big(T^{1/2}N^\beta + N_\eta^{1/2}N^\beta\big).$$

Hence, Theorem OW2(i)–(ii) holds as long as (OW5) and (OW9) hold. But these are the assumptions of Theorem OW2(i)–(ii).

For (iii) to hold, it is sufficient that (OW6) is satisfied and $|\operatorname{tr}\check{\Sigma} - \omega_T^2\operatorname{tr}(MX'XM)/N|$ is asymptotically dominated by $\omega_T^2\operatorname{tr}(MX'XM)/N$. Using arguments very similar to those employed in the proof of Theorem OW1(iii) after equation (OW33), we see that such an asymptotic domination takes place if $\omega_T^2 T^2\operatorname{tr}\Omega/N$ asymptotically dominates $N^{2\beta}(T+N_\eta)\min\{1, T/N\}$, which is implied by the assumptions of Theorem OW2(iii).

S3.1.2. *The Case of the I(1) Weight Proportional to* $1/T$

In this subsection, we would like to revisit the example given in the main text immediately after the formulation of Theorem OW2. We would like to show how, in that example, the theorem would be violated if ω_T converges to zero faster than allowed by condition (OW9).

Consider X_t and Z_t that follow a pure multivariate random walk and white noise processes, respectively. For simplicity, we assume that X_t and Z_t are independent and Gaussian, and that $T/N = c \in (0, \infty)$. In this setting, Assumptions A1–A4 are satisfied with $\alpha = \beta = 0$ and $\operatorname{tr} \Omega = N_\varepsilon = N_\eta = N$, so that condition (OW5) of Theorem OW1 is trivially satisfied while condition (OW9) of Theorem OW2 is violated if and only if ω_T converges to zero as fast or faster than $1/T$. We will assume that $\omega_T = w/T$ for some positive fixed w.

Consider a singular value decomposition $\omega_T XM/\sqrt{N} = USV$. Here U and V are orthonormal matrices and S is a diagonal matrix of the singular values of $\omega_T XM/\sqrt{N}$. By Theorem OW1(i)–(ii), the kth row of V becomes asymptotically collinear with a cosine wave (represented by vector d_k), and the kth diagonal element of S converges to $w/(k\pi)$ as $T \to \infty$. We would like to know whether and how the principal eigenvectors of $\check{\Sigma} = MY'YM/N$ differ from the cosine waves.

From (S4), we have

$$U'YMV'/\sqrt{N} = S + U'ZMV'/\sqrt{N}.$$

Note that the last diagonal element of S is zero (because M has deficient rank), and the last row of V belongs to the null space of M. Denote matrix $U'YMV'$ with the last (zero) column removed as $\tilde{Y}$. Similarly, denote matrices S and $U'ZMV'$ with last (zero) columns removed as $\tilde{S}$ and $\tilde{\varepsilon}$, respectively. With this notation,we have $\tilde{Y}/\sqrt{N} = \tilde{S} + \tilde{\varepsilon}/\sqrt{N}$.

By definition, the entries of the kth principal eigenvector of $\tilde{Y}'\tilde{Y}/N$ equal the scalar products of the kth principal eigenvector of $\check{\Sigma}$ with the rows of V (which become asymptotically collinear with the cosine waves). Further, since we have assumed that X_t and Z_t are independent Gaussian, $\tilde{\varepsilon}$ has i.i.d. (standard) Gaussian entries.

Now, let $\bar{S}$ be an $N \times (T-1)$ matrix with all elements zero, except the first K diagonal elements. For $k \leq K$, let $\bar{S}_{kk} = w/(k\pi)$. Obviously,

$$\tilde{Y}/\sqrt{N} = \bar{S} + \tilde{\varepsilon}/\sqrt{N} + (\tilde{S} - \bar{S}).$$

For *arbitrarily* small $\delta > 0$, we can choose K so large that $\|\tilde{S} - \bar{S}\| < \delta$ with probability at least $1 - \delta$, for all sufficiently large N, T. Therefore, the asymptotic behavior of the kth principal eigenvectors (and eigenvalues) of $\tilde{Y}'\tilde{Y}/N$ and of $(\sqrt{N}\bar{S} + \tilde{\varepsilon})'(\sqrt{N}\bar{S} + \tilde{\varepsilon})/N$ is the same. In particular, the kth components of these two principal eigenvectors converge to the same limit.

By Theorem 1 of Onatski (2018), if $w^2 > (k\pi)^2\sqrt{c}$, the kth component of the kth principal eigenvector of $(\sqrt{N}\bar{S} + \tilde{\varepsilon})'(\sqrt{N}\bar{S} + \tilde{\varepsilon})/N$ converges to

$$l_{kk} := \sqrt{\frac{w^4 - c(k\pi)^4}{w^2\big(w^2 + (k\pi)^2\big)}}.$$

If $w^2 \leq (k\pi)^2\sqrt{c}$, then the kth component converges to zero. Furthermore, by Theorem 5 of Onatski (2018), if $w^2 > (k\pi)^2\sqrt{c}$, the kth principal eigenvalue of $(\sqrt{N}\bar{S} + \tilde{\varepsilon})'(\sqrt{N}\bar{S} +$

$\tilde{\varepsilon})/N$ converges to

$$\mu_k := \big(w^2/(k\pi)^2 + c\big)\big(1 + (k\pi)^2/w^2\big).$$

If $w^2 \leq (k\pi)^2\sqrt{c}$, then the kth eigenvalue converges to $(1+\sqrt{c})^2$.

In our setting, these results show that Theorem OW2 does not hold when $\omega_T = w/T$. Specifically, the scalar product of the kth principal eigenvector of $\check{\Sigma}$ with the "kth cosine wave" does not converge to 1, and the kth principal eigenvalue of $\check{\Sigma}$ does not converge to $w^2/(k\pi)^2$. Instead, if $w^2 > (k\pi)^2\sqrt{c}$, the scalar product converges to $l_{kk} < 1$ and the eigenvalue converges to $\mu_k > w^2/(k\pi)^2$. If $w^2 \leq (k\pi)^2\sqrt{c}$, the kth principal eigenvector of $\check{\Sigma}$ is asymptotically orthogonal to the "kth cosine wave," and the kth eigenvalue asymptotically depends only on c, but not on w or k.

Interestingly, even though Theorem OW2 becomes violated, the principal eigenvalues of $\check{\Sigma}$ still decay very fast, for relatively large w. Hence, the scree plot for matrix $\check{\Sigma}$ still can be wrongfully interpreted as showing the existence of factors in the data. This phenomenon gradually disappears as w becomes smaller and smaller. Similarly, for large w, the kth principal eigenvector of $\check{\Sigma}$ is "almost collinear" with the "kth cosine wave," but the quality of the alignment deteriorates as w decreases.

S3.2. *Demeaned and Standardized Data*

S3.2.1. *Proof of Theorem OW3*

First, we prove the theorem for $k = 1$, and then establish it for general k using mathematical induction. For the demeaned and standardized case, $\hat{F}_1$ is defined as a normalized eigenvector of

$$\hat{\Sigma} = MX'D^{-1}XM/N,$$

corresponding to its largest eigenvalue $\hat{\lambda}_1$. Here $D = \operatorname{diag}\{XMX'/T\}$ and M is the projector on the space orthogonal to the T-dimensional vector of ones.

In contrast to the proof of Theorem OW1, we will not approximate $\hat{\Sigma}$ by $\tilde{\Sigma}$, where the latter matrix is derived from the Beveridge–Nelson decomposition

$$XM = \Psi(1)\varepsilon UM + \Psi^*(L)\varepsilon M.$$

In fact, we will not be using the BN decomposition at all. There are two reasons for this. First, Lemma OW7 cannot be applied to the standardized version of $\Psi^*(L)\varepsilon M$, that is, $D^{-1/2}\Psi^*(L)\varepsilon M$. Second, even if we manage to reduce the analysis of $D^{-1/2}XM$ to that of $D^{-1/2}\Psi(1)\varepsilon UM$, our method of handling $\Psi(1)\varepsilon UM$ would not extend to $D^{-1/2}\Psi(1)\varepsilon UM$, because D and ε are not independent. To summarize, we are not going to use the BN decomposition, and will work directly with the demeaned and standardized data $D^{-1/2}XM = D^{-1/2}eUM$, where $e = [e_1, \dots, e_T]$ with $e_t = \Psi(L)\varepsilon_t$.

Recall that by Lemma OW5, $UM = \sum_{q=1}^{T}\sigma_q v_q w_q'$. Consider a representation of $\hat{F}_1$ in the basis $w_1, \dots, w_T$:

$$\hat{F}_1 = \sum_{q=1}^{T-1}\alpha_q w_q. \tag{S5}$$

Vector $\hat{F}_1$ is orthogonal to w_T, hence summation runs up to $q = T-1$. Representation (S5) yields

$$\hat{\lambda}_1 = \sum_{k,q=1}^{T-1} \alpha_k \alpha_q w_k' \hat{\Sigma} w_q = \sum_{k,q=1}^{T-1} \alpha_k \alpha_q \sigma_k \sigma_q v_k' e' D^{-1} e v_q / N. \tag{S6}$$

It is convenient to represent $\hat{\lambda}_1$ in the form $\hat{\lambda}_1 = A'A$, where

$$A = \frac{1}{\sqrt{N}} \sum_{k=1}^{K} \alpha_k \sigma_k D^{-1/2} e v_k + \frac{1}{\sqrt{N}} \sum_{k=K+1}^{T-1} \alpha_k \sigma_k D^{-1/2} e v_k = A_1 + A_2,$$

with K being a fixed positive integer. Let $e_{j\cdot}$ denote the jth row of e. Then, we have the following explicit expressions for $\|A_1\|^2$ and $\|A_2\|^2$:

$$\|A_1\|^2 = T \sum_{k,q=1}^{K} \alpha_k \alpha_q \frac{1}{N} \sum_{j=1}^{N} \frac{\sigma_k \sigma_q (e_{j\cdot} v_k)(e_{j\cdot} v_q)}{\sum_{t=1}^{T-1} \sigma_t^2 (e_{j\cdot} v_t)^2}, \tag{S7}$$

$$\|A_2\|^2 = \frac{T}{N} \sum_{j=1}^{N} \frac{\left(\sum_{k=K+1}^{T-1} \alpha_k \sigma_k e_{j\cdot} v_k \right)^2}{\sum_{t=1}^{T-1} \sigma_t^2 (e_{j\cdot} v_t)^2}. \tag{S8}$$

Let

$$M_{j,kq} = \sigma_k \sigma_q (e_{j\cdot} v_k)(e_{j\cdot} v_q) \Big/ \sum_{t=1}^{T-1} \sigma_t^2 (e_{j\cdot} v_t)^2.$$

Then

$$\|A_1\|^2 = T \sum_{k,q=1}^{K} \alpha_k \alpha_q \frac{1}{N} \sum_{j=1}^{N} M_{j,kq},$$

and by Assumption A2b, $M_{j,kq}$ are independent for different $j = 1, \dots, N$. Moreover, since

$$\left| \sigma_k \sigma_q v_k' e_{j\cdot}' e_{j\cdot} v_q \right| \le \frac{\sigma_k^2 v_k' e_{j\cdot}' e_{j\cdot} v_k + \sigma_q^2 v_q' e_{j\cdot}' e_{j\cdot} v_q}{2},$$

we have $|M_{j,kq}| \le 1/2$. Therefore, the variance of $\frac{1}{N} \sum_{j=1}^{N} M_{j,kq}$ is no larger than $1/(4N)$, and thus, the asymptotic behavior of $\|A_1\|^2$ is, to a large extent, determined by that of $\frac{1}{N} \sum_{j=1}^{N} \mathbb{E} M_{j,kq}$.

Consider the finite Fourier transform of $e_{j\cdot}$ (e.g., Brillinger (2001, Chapter 3.1)):

$$d(\varpi) = \sum_{t=1}^{T} e_{jt} \exp\{-\mathrm{i}(t-1)\varpi\}, \quad \varpi \in [0, 2\pi].$$

Let us denote $d(\varpi_r/2)$ as d_r and $d(-\varpi_r/2)$ as d_{-r}, where $\varpi_r = 2\pi r/T$. By definition (see Lemma OW5), the tth entry of v_r for $r = 1, \dots, T-1$ equals

$$v_{rt} = \sqrt{2/T}\big(\exp\{\mathrm{i}(t-1)\varpi_r/2\} - \exp\{-\mathrm{i}(t-1)\varpi_r/2\}\big)/(2\mathrm{i}).$$

Therefore,

$$e_{j\cdot}v_r = \sqrt{2/T}(d_{-r} - d_r)/(2\mathrm{i}). \tag{S9}$$

Theorem 13 in Chapter 4 of Hannan (1970) (one of the assumptions of this theorem requires that the spectral density of $e_{j\cdot}$ at zero is positive, which is ensured by Assumption A2b), identity (S9), and the definition of σ_i imply that, for any fixed j, k, and q, as $T \to \infty$, $M_{j,kq} \xrightarrow{d} \mathcal{M}_{kq}$, where

$$\mathcal{M}_{kq} = (kq)^{-1}\eta_k\eta_q \Big/ \sum_{t=1}^{\infty}(t)^{-2}\eta_t^2,$$

and $\{\eta_k\}_{k=1}^{\infty}$ is a sequence of i.i.d. $N(0,1)$ random variables. Since $M_{j,kq}$ is bounded, the convergence in distribution implies the convergence of the moments of $M_{j,kq}$. In particular, as $T \to \infty$,

$$\mathbb{E}M_{j,kq} = \mathbb{E}\mathcal{M}_{kq} + o_j(1). \tag{S10}$$

To proceed further, we need to establish the uniformity of $o_j(1)$ in $j = 1, \dots, N$.

In preparation for the proof of the uniformity, we establish some bounds on the spectral density of the series e_{jt}, $t \in \mathbb{Z}$ at frequency ϖ,

$$f_j(\varpi) = \frac{1}{2\pi}\left|\sum_{k=0}^{\infty}(\Psi_k)_{jj}\exp\{\mathrm{i}k\varpi\}\right|^2. \tag{S11}$$

By Assumption A2b, for all j,

$$\max_{\varpi}\big|f_j(\varpi)\big| \le B^2/(2\pi). \tag{S12}$$

Furthermore, differentiating both sides of (S11) with respect to ϖ, we obtain

$$f_j(\varpi)' = \frac{1}{2\pi}\sum_{k,r=0}^{\infty}\mathrm{i}(k-r)(\Psi_k)_{jj}(\Psi_r)_{jj}\exp\{\mathrm{i}k\varpi - \mathrm{i}r\varpi\}.$$

Since $|k-r| \le (k+1)(r+1)$, we conclude, using Assumption A2b, that for all j,

$$\max_{\varpi}\big|f_j'(\varpi)\big| \le B^2/(2\pi). \tag{S13}$$

Finally, Assumption A2b also implies that, for all j,

$$f_j(0) \ge b^2/(2\pi). \tag{S14}$$

We will need the following two lemmas. Their proofs can be found in Sections S3.2.2 and S3.2.3.

LEMMA S2: *Under the assumptions of Theorem OW3, there exists an absolute constant C such that, for any* $j = 1, \ldots, N$ *and any* $q, r, p, l = 1, \ldots, T-1$, *we have*

(i) $|\mathbb{E}(v_q' e_{j\cdot}' e_{j\cdot} v_r) - 2\pi f_j(\varpi_q/2)\delta_{qr}| \leq CB^2/T$, *where* δ_{qr} *is the Kronecker delta and* $\varpi_q = 2\pi q/T$;

(ii) $|\operatorname{Cov}(v_q' e_{j\cdot}' e_{j\cdot} v_r, v_p' e_{j\cdot}' e_{j\cdot} v_l)| \leq C(\delta_{qp}\delta_{rl} + \delta_{ql}\delta_{rp} + (1+\varkappa_4)/T)B^4$, *where* $\varkappa_4$ *is as defined in Assumption* A1.

LEMMA S3: *Let* $\mathbf{X}$ *be an R-dimensional vector with the kth coordinate* $e_{j\cdot} v_k$ *and let* $\mathbf{Y}$ *be an R-dimensional vector with i.i.d. normal coordinates with mean zero and variance* $2\pi f_j(0)$. *Further, let* $g : \mathbb{R}^R \to \mathbb{R}$ *be a thrice continuously differentiable function with all derivatives up to and including the third order are bounded by absolute value by a constant* M_g. *Then, under Assumptions* A1 *and* A2b, *we have, for all sufficiently large* T,

$$\left|\mathbb{E}g(\mathbf{X}) - \mathbb{E}g(\mathbf{Y})\right| \leq M_g C/\sqrt{T},$$

where C depends only on R, B, and $\varkappa_4$, *with* $\varkappa_4$ *and B as defined in Assumptions* A1 *and* A2b.

Now we are ready to prove the uniformity of $o_j(1)$ in (S10). By definition,

$$M_{j,kq} = \frac{\sigma_k \sigma_q \mathbf{X}_k \mathbf{X}_q}{\sum_{t=1}^{R} \sigma_t^2 \mathbf{X}_t^2 + \mathbf{Z}},$$

where $\mathbf{Z} = \sum_{t=R+1}^{T-1} \sigma_t^2 v_t' e_{j\cdot}' e_{j\cdot} v_t$. For $\max\{k, q\} \leq R$, denote $\sigma_k \sigma_q \mathbf{X}_k \mathbf{X}_q / \sum_{t=1}^{R} \sigma_t^2 \mathbf{X}_t^2$ as $\bar{M}_{j,kq}$.

Consider the event $\mathcal{E} = \{\mathbf{X}_1^2 \leq \delta\}$ and let $\mathbf{1}_{\mathcal{E}}$ and $\mathbf{1}_{\mathcal{E}^c}$ be the indicators of this event and of its complement, respectively. Since $|M_{j,kq}| \leq 1/2$ and $|\bar{M}_{j,kq}| \leq 1/2$, we have

$$\mathbb{E}\big[|M_{j,kq} - \bar{M}_{j,kq}| \times \mathbf{1}_{\mathcal{E}}\big] \leq p_\delta = \Pr(\mathcal{E}).$$

By setting function g in Lemma S3 so that it approximates $\mathbf{1}_{\mathcal{E}}$, we see that p_δ can be made arbitrarily small for all sufficiently large T uniformly in j by choosing δ sufficiently small. On the other hand,

$$\mathbb{E}\big[|M_{j,kq} - \bar{M}_{j,kq}| \times \mathbf{1}_{\mathcal{E}^c}\big] = \mathbb{E}\left[\frac{|M_{j,kq}|\mathbf{Z}\mathbf{1}_{\mathcal{E}^c}}{\sum_{t=1}^{R} \sigma_t^2 \mathbf{X}_t^2}\right] \leq \frac{\mathbb{E}\mathbf{Z}}{2\delta\sigma_1^2}.$$

By Lemma S2(i) and by (S12),

$$\mathbb{E}\mathbf{Z} \leq \sum_{t=R+1}^{T-1} \sigma_t^2\big(B + CB^2/T\big) \leq \tilde{C}\sigma_1^2/R$$

for some absolute constant $\tilde{C}$, where the latter inequality follows from the definition of σ_t^2. Therefore,

$$\mathbb{E}\big[|M_{j,kq} - \bar{M}_{j,kq}| \times \mathbf{1}_{\mathcal{E}^c}\big] \leq \tilde{C}/(2\delta R),$$

and

$$\mathbb{E}|M_{j,kq} - \bar{M}_{j,kq}| = \mathbb{E}\big[|M_{j,kq} - \bar{M}_{j,kq}| \times \mathbf{1}_{\mathcal{E}}\big] + \mathbb{E}\big[|M_{j,kq} - \bar{M}_{j,kq}| \times \mathbf{1}_{\mathcal{E}^c}\big]$$
$$\leq p_\delta + \tilde{C}/(2\delta R),$$

which can be made arbitrarily small for all sufficiently large T uniformly in j by choosing sufficiently small δ and sufficiently large R.

Further, let $\bar{\mathcal{M}}_{kq} = \sigma_k \sigma_q \eta_k \eta_q / \sum_{t=1}^R \sigma_t^2 \eta_t^2$. By choosing R sufficiently large, we can make $\mathbb{E}|\bar{\mathcal{M}}_{kq} - \mathcal{M}_{kq}|$ arbitrarily small for all sufficiently large T.

Now consider $\mathbb{E}\bar{M}_{j,kq} - \mathbb{E}\bar{\mathcal{M}}_{kq}$. To bound this expression uniformly in j, we would like to use Lemma S3 again. Unfortunately, $\bar{M}_{j,kq}$ does not have bounded derivatives as a function of $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_R)'$. The derivatives are unbounded in a neighborhood of $\mathbf{X} = 0$.

To overcome this difficulty, let us introduce $\tau : [0, \infty) \to \mathbb{R}$, a thrice continuously differential function such that

$$\tau(z) = z \quad \text{for } z > \delta,$$
$$\tau(z) > \delta/2 \quad \text{for } z \geq 0,$$

and the first three derivatives of $\tau(z)$ bounded for $z \in [0, \delta]$. Further, let

$$\tilde{M}_{j,kq} = \frac{\sigma_k \sigma_q \mathbf{X}_k \mathbf{X}_q}{\sigma_1^2 \tau\big(\mathbf{X}_1^2\big) + \sum_{t=2}^{R} \sigma_t^2 \mathbf{X}_t^2} = g(\mathbf{X}).$$

Similarly, let $\tilde{\mathcal{M}}_{kq} = g(\eta)$, where $\eta = (\eta_1, \ldots, \eta_R)'$. Note that the derivatives $\partial_i^r g(x)$ for $r = 1, 2, 3$ are bounded, with a bound that depends only δ, but not on R.

We have

$$\bar{M}_{j,kq} = \bar{M}_{j,kq} \mathbf{1}_{\mathcal{E}} + \bar{M}_{j,kq} \mathbf{1}_{\mathcal{E}^c} = \bar{M}_{j,kq} \mathbf{1}_{\mathcal{E}} + \tilde{M}_{j,kq} \mathbf{1}_{\mathcal{E}^c}$$
$$= (\bar{M}_{j,kq} - \tilde{M}_{j,kq}) \mathbf{1}_{\mathcal{E}} + \tilde{M}_{j,kq}.$$

Therefore,

$$|\mathbb{E}\bar{M}_{j,kq} - \mathbb{E}\tilde{M}_{j,kq}| = \big|\mathbb{E}\big[(\bar{M}_{j,kq} - \tilde{M}_{j,kq}) \mathbf{1}_{\mathcal{E}}\big]\big| \leq p_\delta,$$

so that $|\mathbb{E}\bar{M}_{j,kq} - \mathbb{E}\tilde{M}_{j,kq}|$ can be made arbitrarily small, uniformly over j, by choosing sufficiently small δ. By similar arguments, we can show that $|\mathbb{E}\tilde{\mathcal{M}}_{kq} - \mathbb{E}\bar{\mathcal{M}}_{kq}|$ can be made arbitrarily small by choosing sufficiently small δ. Finally, $|\mathbb{E}\tilde{M}_{j,kq} - \mathbb{E}\tilde{\mathcal{M}}_{kq}|$ can be made arbitrarily small uniformly over j for all sufficiently large T by Lemma S3. Summing up the above arguments, we conclude that $o_j(1)$ in (S10) is uniform in j.

By (S10) and Chebyshev's inequality,

$$\frac{1}{N}\sum_{j=1}^{N} M_{j,kq} = \mathbb{E}\mathcal{M}_{kq} + o(1) + O_{\mathrm{P}}\big(N^{-1/2}\big).$$

Furthermore, by a conditioning argument, it is easy to show that $\mathbb{E}\mathcal{M}_{kq} = 0$ for $k \neq q$. Now recall

$$\|A_1\|^2 = T\sum_{k,q=1}^{K} \alpha_k\alpha_q \frac{1}{N}\sum_{j=1}^{N} M_{j,kq}.$$

Therefore, we have

$$\frac{1}{T}\|A_1\|^2 = \sum_{k=1}^{K} \alpha_k^2 \mathbb{E}\mathcal{M}_{kk} + o(1) + O_{\mathrm{P}}\big(N^{-1/2}\big).$$

For A_2, the Cauchy–Schwarz inequality and the identity $\sum \alpha_i^2 = 1$ yield

$$\begin{aligned} \|A_2\|^2/T &= \frac{1}{N}\sum_{j=1}^{N} \frac{\left(\sum_{k=K+1}^{T-1} \alpha_k\sigma_k e_{j\cdot}v_k\right)^2}{\sum_{t=1}^{T-1}\sigma_t^2 v_t' e_{j\cdot}' e_{j\cdot}v_t} \\ &\leq \left(1 - \sum_{i=1}^{K}\alpha_i^2\right)\frac{1}{N}\sum_{j=1}^{N} \frac{\sum_{k=K+1}^{T-1} \sigma_k^2 (e_{j\cdot}v_k)^2}{\sum_{t=1}^{T-1}\sigma_t^2 (e_{j\cdot}v_t)^2}. \end{aligned}$$

Note that

$$\begin{aligned} \mathbb{E}\frac{\sum_{k=K+1}^{T-1} \sigma_k^2 (e_{j\cdot}v_k)^2}{\sum_{t=1}^{T-1}\sigma_t^2 (e_{j\cdot}v_t)^2} &= 1 - \mathbb{E}\frac{\sum_{k=1}^{K} \sigma_k^2 (e_{j\cdot}v_k)^2}{\sum_{t=1}^{T-1}\sigma_t^2 (e_{j\cdot}v_t)^2} = 1 - \sum_{k=1}^{K}\mathbb{E}\mathcal{M}_{kk} + o(1) \\ &= \delta_K + o(1), \end{aligned}$$

where δ_K can be made arbitrarily small by choosing sufficiently large K, and $o(1)$ is uniform in j, but may depend on K. Therefore,

$$\|A_2\|^2/T \leq \left(1 - \sum_{i=1}^{K}\alpha_i^2\right)\big(\delta_K + o(1) + O_{\mathrm{P}}\big(N^{-1/2}\big)\big).$$

For $\hat{\lambda}_1$, we have

$$\hat{\lambda}_1 = \|A\|^2 \leq \|A_1\|^2 + 2\|A_1\|\|A_2\| + \|A_2\|^2.$$

This inequality and the above bounds on $\|A_1\|^2/T$ and $\|A_2\|^2/T$ yield

$$\hat{\lambda}_1/T \le \sum_{k=1}^{K} \alpha_k^2 \mathbb{E}\mathcal{M}_{kk} + \Delta_K + o_{\mathrm{P}}(1),$$

where Δ_K can be made arbitrarily small by choosing sufficiently large K, and the convergence of $o_{\mathrm{P}}(1)$ to 0 as $T \to \infty$ may depend on the choice of K. This implies that

$$\hat{\lambda}_1/T \le \sum_{k=1}^{T-1} \alpha_k^2 \mathbb{E}\mathcal{M}_{kk} + o_{\mathrm{P}}(1).$$

It is easy to see that $\mathbb{E}\mathcal{M}_{aa} > \mathbb{E}\mathcal{M}_{bb}$ for $a < b$. Indeed, consider functions $\mathcal{H}_a$, $\mathcal{H}_b$: $[0,\infty)^2 \to \mathbb{R}$:

$$\mathcal{H}_a(x,y) = \frac{x\eta_a^2}{\sum_{t\ge 1, t\ne a,b} (t)^{-2}\eta_t^2 + x\eta_a^2 + y\eta_b^2},$$

$$\mathcal{H}_b(x,y) = \frac{x\eta_b^2}{\sum_{t\ge 1, t\ne a,b} (t)^{-2}\eta_t^2 + x\eta_b^2 + y\eta_a^2}.$$

Functions $\mathcal{H}_a$ and $\mathcal{H}_b$ are increasing in x and decreasing in y and this monotonicity is strict (unless $\eta_a^2 = 0$ or $\eta_b^2 = 0$, which is a zero probability event). Therefore, with probability 1, for $a < b$, we have

$$\mathbb{E}\mathcal{M}_{aa} = \mathbb{E}\mathcal{H}_a(a^{-2}, b^{-2}) > \mathbb{E}\mathcal{H}_a\left(\frac{a^{-2}+b^{-2}}{2}, \frac{a^{-2}+b^{-2}}{2}\right) \quad \text{and}$$

$$\mathbb{E}\mathcal{M}_{bb} = \mathbb{E}\mathcal{H}_b(b^{-2}, a^{-2}) < \mathbb{E}\mathcal{H}_b\left(\frac{a^{-2}+b^{-2}}{2}, \frac{a^{-2}+b^{-2}}{2}\right).$$

On the other hand,

$$\mathbb{E}\mathcal{H}_a\left(\frac{a^{-2}+b^{-2}}{2}, \frac{a^{-2}+b^{-2}}{2}\right) = \mathbb{E}\mathcal{H}_b\left(\frac{a^{-2}+b^{-2}}{2}, \frac{a^{-2}+b^{-2}}{2}\right).$$

Therefore, $\mathbb{E}\mathcal{M}_{aa} > \mathbb{E}\mathcal{M}_{bb}$. This implies that

$$\sum_{k=1}^{T-1} \alpha_k^2 \mathbb{E}\mathcal{M}_{kk} \le \alpha_1^2 \mathbb{E}\mathcal{M}_{11} + (1-\alpha_1^2)\mathbb{E}\mathcal{M}_{22},$$

and thus,

$$\hat{\lambda}_1/T \le \alpha_1^2 \mathbb{E}\mathcal{M}_{11} + (1-\alpha_1^2)\mathbb{E}\mathcal{M}_{22} + o_{\mathrm{P}}(1). \tag{S15}$$

On the other hand, $\hat{\lambda}_1$ must be no smaller than $v_1'\hat{\Sigma}v_1$, which yields

$$\hat{\lambda}_1/T \ge \mathbb{E}\mathcal{M}_{11} + o_{\mathrm{P}}(1). \tag{S16}$$

Thus,

$$\left(1-\alpha_1^2\right)\mathbb{E}\mathcal{M}_{11} \le \left(1-\alpha_1^2\right)\mathbb{E}\mathcal{M}_{22} + o_{\mathrm{P}}(1),$$

which only holds if

$$\alpha_1^2 \xrightarrow{\mathrm{P}} 1. \tag{S17}$$

This yields statement (i) of the theorem.

To establish statement (ii), note that (S15) and (S16) imply

$$|\hat{\lambda}_1/T - \mathbb{E}\mathcal{M}_{11}| \le \left|1-\alpha_1^2\right|(\mathbb{E}\mathcal{M}_{11} + \mathbb{E}\mathcal{M}_{22}) + o_{\mathrm{P}}(1).$$

Combining this with (S17), we conclude that

$$\hat{\lambda}_1/T = \mathbb{E}\mathcal{M}_{11} + o_{\mathrm{P}}(1).$$

This yields (ii) because $\mathbb{E}\mathcal{M}_{11} = \nu_1$ (the latter being defined in the statement of Theorem OW3).

Further,

$$\operatorname{tr}\hat{\Sigma} = \operatorname{tr}\left(MX'D^{-1}XM/N\right) = \operatorname{tr}\left(D^{-1}XMX'/N\right).$$

But, by definition, $D = \operatorname{diag}\{XMX'/T\}$. Therefore, $\operatorname{tr}\hat{\Sigma} = T$ and

$$\hat{\lambda}_1/T = \hat{\lambda}_1/\operatorname{tr}\hat{\Sigma},$$

which yields statement (iii) of the theorem.

For $k = m > 1$, the theorem follows by mathematical induction. Indeed, suppose it holds for $k < m$. Consider a representation $\hat{F}_m = \sum_{q=1}^{T-1}\alpha_q w_q$. Since $\hat{F}_m'\hat{F}_j = 0$ for all $j < m$, and since $|\hat{F}_j' w_j| = 1 + o_{\mathrm{P}}(1)$ by the induction hypothesis, we must have $\alpha_j = o_{\mathrm{P}}(1)$ for all $j < m$. In particular,

$$\hat{F}_m'\hat{\Sigma}\hat{F}_m = \sum_{q,r=m}^{T-1}\alpha_q\alpha_r\sigma_q\sigma_r v_q' e' D^{-1} e v_r/N + o_{\mathrm{P}}(T).$$

In addition to this equality, we must have

$$\sum_{i=1}^{m-1}\hat{\lambda}_i + \hat{F}_m'\hat{\Sigma}\hat{F}_m \ge \sum_{i=1}^{m} w_i' M U' e' D^{-1} e U M w_i/N = T\sum_{i=1}^{m}\mathbb{E}\mathcal{M}_{ii} + o_{\mathrm{P}}(T).$$

Combining the above two displays, and using the induction hypothesis, this time regarding the validity of the identities $\hat{\lambda}_i = T\mathbb{E}\mathcal{M}_{ii} + o_{\mathrm{P}}(T)$ for all $i < m$, we obtain

$$\sum_{q,r=m}^{T-1}\alpha_q\alpha_r\sigma_q\sigma_r v_q' e' D^{-1} e v_r/N \ge T\mathbb{E}\mathcal{M}_{mm} + o_{\mathrm{P}}(T). \tag{S18}$$

Statements (i), (ii), and (iii) for $k = m$ now follow by arguments that are very similar to those used above for the case $k = 1$.

That is, we represent the sum on the left-hand side of (S18) in the form $A'A$, where $A = \frac{1}{\sqrt{N}} \sum_{k=1}^{T-1} \alpha_k \sigma_k D^{-1/2} e v_k$. Then proceed along the lines of the above proof to obtain an upper bound on $A'A$, similar to the right-hand side of (S15). Then, combining this upper bound with the lower bound (S18), we prove the convergence $\alpha_m^2 \xrightarrow{\mathrm{P}} 1$. Finally, we proceed to establishing parts (ii) and (iii) using part (i). We omit further details to save space.

S3.2.2. *Proof of Lemma S2*

Identity (S9) yields

$$\mathbb{E}\big(v_q' e_{j\cdot}' e_{j\cdot} v_r\big) = -\mathbb{E}\big[(d_{-q} - d_q)(d_{-r} - d_r)\big]/(2T),$$

and

$$\operatorname{Cov}\big(v_q' e_{j\cdot}' e_{j\cdot} v_r, v_p' e_{j\cdot}' e_{j\cdot} v_l\big) = \frac{1}{4T^2} \operatorname{Cov}\big((d_{-q} - d_q)(d_{-r} - d_r), (d_{-p} - d_p)(d_{-l} - d_l)\big).$$

To evaluate the latter expectation and covariance, we use Theorem 4.3.2 of Brillinger (2001) (B01), which describes joint cumulants of finite Fourier transforms. First, we need to represent the expectation and covariance in terms of the joint cumulants. By their definition, and by Theorem 2.3.1 (B01, p. 19),

$$\mathbb{E}\big(v_q' e_{j\cdot}' e_{j\cdot} v_r\big) = -\frac{1}{2T} \sum_{s_1, s_2 \in \{-1,+1\}} s_1 s_2 \operatorname{cum}(d_{s_1 q}, d_{s_2 r}). \tag{S19}$$

Similarly, $\operatorname{Cov}(v_q' e_{j\cdot}' e_{j\cdot} v_r, v_p' e_{j\cdot}' e_{j\cdot} v_l)$ equals

$$\frac{1}{4T^2} \sum_{s_1, s_2, s_3, s_4 \in \{-1,+1\}} s_1 s_2 s_3 s_4 \operatorname{cum}(d_{s_1 q} d_{s_2 r}, d_{s_3 p} d_{s_4 l}).$$

By Theorem 2.3.2 of B01, the joint cumulant of the two products of d, as in the latter display, can be represented in the form of a sum of the products of the cumulants of order 2 and the fourth-order cumulant. Precisely, we have

$$\begin{aligned} &\operatorname{Cov}\big(v_q' e_{j\cdot}' e_{j\cdot} v_r, v_p' e_{j\cdot}' e_{j\cdot} v_l\big) \\ &= \frac{1}{4T^2} \sum_{s_1, s_2, s_3, s_4 \in \{-1,+1\}} s_1 s_2 s_3 s_4 \\ &\quad \times \big\{\operatorname{cum}(d_{s_1 q}, d_{s_3 p}) \operatorname{cum}(d_{s_2 r}, d_{s_4 l}) + \operatorname{cum}(d_{s_1 q}, d_{s_4 l}) \operatorname{cum}(d_{s_2 r}, d_{s_3 p}) \\ &\quad + \operatorname{cum}(d_{s_1 q}, d_{s_2 r}, d_{s_3 p}, d_{s_4 l})\big\}. \end{aligned} \tag{S20}$$

LEMMA S4: *Under assumptions of Theorem OW3, there exists an absolute constant C such that, for any $q, r, p, l = 1, \dots, T-1$, and any $s_1, s_2, s_3, s_4 \in \{-1, +1\}$,*

$$\big|\operatorname{cum}(d_{s_1 q}, d_{s_2 r}) - 2\pi H_{s_1 q, s_2 r} f_j(\varpi_q/2)\big| \le CB^2, \tag{S21}$$

where $H_{s_1 q, s_2 r} = \sum_{t=0}^{T-1} e^{-\mathrm{i}t(s_1 \omega_q + s_2 \omega_r)/2}$, and

$$\big|\operatorname{cum}(d_{s_1 q}, d_{s_2 r}, d_{s_3 p}, d_{s_4 l}) - (2\pi)^3 H_{s_1 q, s_2 r, s_3 p, s_4 l} f_{j4}\big| \le C \varkappa_4 B^4, \tag{S22}$$

where $H_{s_1q,s_2r,s_3p,s_4l} = \sum_{t=0}^{T-1} e^{-\mathrm{i}t(s_1\varpi_q+s_2\varpi_r+s_3\varpi_p+s_4\varpi_l)/2}$ *and* f_{j4} *is the fourth-order cumulant spectrum of the series* e_{jt}, $t\in\mathbb{Z}$ *at frequencies* $s_1\varpi_q/2$, $s_2\varpi_r/2$, $s_3\varpi_p/2$.

PROOF: The proof of Theorem 4.3.2 in B01 implies that the left-hand side of (S21) can be bounded by $C\sum_{k=0}^{\infty}(1+k)|\Gamma_j(k)|$, where C is an absolute constant and

$$\Gamma_j(k)=\mathbb{E}e_{js}e_{j,s-k}=\sum_{t=-\infty}^{\infty}\theta_{jt}\theta_{j,t-k}.$$

Here $\theta_{jt}=(\Psi_t)_{jj}$ for $t\geq 0$ and $\theta_{jt}=0$ for $t<0$. On the other hand,

$$\begin{aligned}&\sum_{k=0}^{\infty}(1+k)\big|\Gamma_j(k)\big|\\&\quad\leq\sum_{k=0}^{\infty}(1+k)\sum_{t=-\infty}^{\infty}|\theta_{jt}||\theta_{j,t-k}|\\&\quad\leq\sum_{k=0}^{\infty}\sum_{t=-\infty}^{\infty}\big(1+|t-k|\big)|\theta_{jt}||\theta_{j,t-k}|+\sum_{k=0}^{\infty}\sum_{t=-\infty}^{\infty}\big(1+|t|\big)|\theta_{jt}||\theta_{j,t-k}|\leq 2B^2,\end{aligned}\tag{S23}$$

where the last inequality follows from Assumption A2b. This yields (S21).

Similarly, from the proof of Theorem 4.3.2 in B01, we know that the left-hand side of (S22) can be bounded by

$$C\sum_{k_1,k_2,k_3=-\infty}^{\infty}\big(1+|k_1|+|k_2|+|k_3|\big)\big|c_{j4}(k_1,k_2,k_3)\big|,\tag{S24}$$

where C is an absolute constant and $c_{j4}(k_1,k_2,k_3)$ is the joint fourth-order cumulant of e_{js}, $e_{j,s-k_1}$, $e_{j,s-k_2}$, and $e_{j,s-k_3}$. By Theorem 2.3.1(i),(iii) of B01, this cumulant equals

$$\begin{aligned}&\sum_{t_1,t_2,t_3,t_4=-\infty}^{\infty}\theta_{j,t_1-k_1}\theta_{j,t_2-k_2}\theta_{j,t_3-k_3}\theta_{j,t_4}\operatorname{cum}(\varepsilon_{j,-t_1},\varepsilon_{j,-t_2},\varepsilon_{j,-t_3},\varepsilon_{j,-t_4})\\&\quad=\sum_{t=-\infty}^{\infty}\theta_{j,t-k_1}\theta_{j,t-k_2}\theta_{j,t-k_3}\theta_{j,t}\big(\mathbb{E}\varepsilon_{j,-t}^4-3\big)\\&\quad\leq\sum_{t=-\infty}^{\infty}|\theta_{j,t-k_1}\theta_{j,t-k_2}\theta_{j,t-k_3}\theta_{j,t}|\varkappa_4,\end{aligned}$$

where the last line follows from Assumption A1. By an argument similar to (S23), expression (S24) can be bounded by $C\varkappa_4B^4$, where C is an absolute constant. This yields (S22). *Q.E.D.*

Returning to the proof of Lemma S2, consider (S19). Inequality (S21) implies that

$$\left|\mathbb{E}\big(v_q'e_{j\cdot}'e_{j\cdot}v_r\big)+\frac{1}{2T}\sum_{s_1,s_2\in\{-1,+1\}}s_1s_2H_{s_1q,s_2r}2\pi f_j(\varpi_q/2)\right|\leq\frac{2KB^2}{T}.\tag{S25}$$

Further, for $q = r$,

$$\sum_{s_1,s_2\in\{-1,+1\}} s_1 s_2 H_{s_1 q, s_2 r} = \sum_{s_1,s_2\in\{-1,+1\}} s_1 s_2 \sum_{t=0}^{T-1} e^{-\mathrm{i}t(s_1+s_2)\pi q/T} = -2T. \tag{S26}$$

For $q \neq r$ and such that $s_1 q + s_2 r$ is even for all $s_1, s_2 \in \{-1, +1\}$,

$$\sum_{s_1,s_2\in\{-1,+1\}} s_1 s_2 H_{s_1 q, s_2 r} = \sum_{s_1,s_2\in\{-1,+1\}} s_1 s_2 \sum_{t=0}^{T-1} e^{-\mathrm{i}t(s_1 q+s_2 r)\pi/T} = 0. \tag{S27}$$

Here, the latter equality holds because $s_1 q + s_2 r$ is an even nonzero integer, such that $|s_1 q + s_2 r| < 2T$ (recall that $1 \le q, r \le T - 1$). For $q \neq r$ and such that $s_1 q + s_2 r$ is odd for all $s_1, s_2 \in \{-1, +1\}$, we have

$$\sum_{t=0}^{T-1} e^{-\mathrm{i}t(s_1 q+s_2 r)\pi/T} = \frac{-2}{e^{-\mathrm{i}(s_1 q+s_2 r)\pi/T} - 1}.$$

Nevertheless, $\sum_{s_1,s_2\in\{-1,+1\}} s_1 s_2 H_{s_1 q, s_2 r}$ still equals zero because

$$\frac{-2}{e^{-\mathrm{i}(q+r)\pi/T} - 1} + \frac{-2}{e^{\mathrm{i}(q+r)\pi/T} - 1} + \frac{2}{e^{-\mathrm{i}(q-r)\pi/T} - 1} + \frac{2}{e^{\mathrm{i}(q-r)\pi/T} - 1} = 2 - 2.$$

Therefore, (S27) still holds. Using identities (S26) and (S27) in (S25), we obtain statement (i) of Lemma S2.

Next, consider (S20). By (S21) and (S22), the difference between

$$\left(\frac{\pi}{T}\right)^2 \sum_{s_1,s_2,s_3,s_4\in\{-1,+1\}} s_1 s_2 s_3 s_4 \big\{2\pi H_{s_1 q, s_2 r, s_3 p, s_4 l} f_{j4} + (H_{s_1 q, s_3 p} H_{s_2 r, s_4 l} + H_{s_1 q, s_4 l} H_{s_2 r, s_3 p}) f_j(\varpi_q/2) f_j(\varpi_r/2)\big\}$$

and $\mathrm{Cov}(v_q' e_{j\cdot}' e_{j\cdot} v_r, v_p' e_{j\cdot}' e_{j\cdot} v_l)$ is no larger by absolute value than

$$\frac{1}{4T^2} \sum_{s_1,s_2,s_3,s_4\in\{-1,+1\}} \big\{C\varkappa_4 B^4 + 2C^2 B^4 + 2\pi C B^2 \big(\big|H_{s_1 q, s_3 p} f_j(\varpi_q/2)\big| + \big|H_{s_2 r, s_4 l} f_j(\varpi_r/2)\big| + \big|H_{s_1 q, s_4 l} f_j(\varpi_q/2)\big| + \big|H_{s_2 r, s_3 p} f_j(\varpi_r/2)\big|\big)\big\},$$

which, in its turn, is bounded from above by $C(1 + \varkappa_4)B^4/T$, where C is an absolute constant (we remind the reader that throughout the paper, the value of the absolute constant C may change from one appearance to another). Indeed, such a bound follows from (S12) and the fact that $|H_{a,b}| \le T$. Further, from the above analysis of $\mathbb{E}(v_q' e_{j\cdot}' e_{j\cdot} v_r)$,

$$\sum_{s_1,s_2,s_3,s_4\in\{-1,+1\}} s_1 s_2 s_3 s_4 (H_{s_1 q, s_3 p} H_{s_2 r, s_4 l} + H_{s_1 q, s_4 l} H_{s_2 r, s_3 p}) = 4T^2 (\delta_{qp}\delta_{rl} + \delta_{ql}\delta_{rp}).$$

Therefore, from (S12),

$$\left| \sum_{s_1,s_2,s_3,s_4\in\{-1,+1\}} s_1 s_2 s_3 s_4 (H_{s_1 q, s_3 p} H_{s_2 r, s_4 l} + H_{s_1 q, s_4 l} H_{s_2 r, s_3 p}) f_j(\varpi_q/2) f_j(\varpi_r/2) \right|$$

is no larger than $4T^2B^4(\delta_{qp}\delta_{rl}+\delta_{ql}\delta_{rp})/(2\pi)^2$. Next, by Theorem 2.8.1 of B01,

$$f_{j4}(\mu_1,\mu_2,\mu_3)=\Theta(\mu_1)\Theta(\mu_2)\Theta(\mu_3)\Theta(-\mu_1-\mu_2-\mu_3)\frac{\mathbb{E}\varepsilon_{jt}^4-3}{(2\pi)^3},$$

where $\Theta(\mu)=\sum_{k=0}^{\infty}\theta_k e^{-\mathrm{i}k\mu}$, and since $|H_{a,b,c,d}|\le T$,

$$\left|\sum_{s_1,s_2,s_3,s_4\in\{-1,+1\}} s_1s_2s_3s_4 2\pi H_{s_1q,s_2r,s_3p,s_4l}f_{j4}\right|\le\frac{TB^4\varkappa_4}{(2\pi)^2}.$$

Overall, we conclude that $|\operatorname{Cov}(v_q'e_{j\cdot}'e_{j\cdot}v_r, v_p'e_{j\cdot}'e_{j\cdot}v_l)|$ is no larger than

$$\left(\frac{\pi}{T}\right)^2\left(\frac{TB^4\varkappa_4}{(2\pi)^2}+\frac{4T^2B^4(\delta_{qp}\delta_{rl}+\delta_{ql}\delta_{rp})}{(2\pi)^2}\right)+\frac{C(1+\kappa_4)B^4}{T},$$

which yields statement (ii) of Lemma S2.

S3.2.3. *Proof of Lemma S3*

Our proof is based on the following theorem, established in Chatterjee (2006).

THEOREM S5—Chatterjee (2006): *Suppose x and y are random vectors in $\mathbb{R}^m$ with y having independent components. For $1\le i\le m$, let*

$$R_i:=\mathbb{E}\big|\mathbb{E}(x_i|x_1,\dots,x_{i-1})-\mathbb{E}(y_i)\big|,$$
$$B_i:=\mathbb{E}\big|\mathbb{E}(x_i^2|x_1,\dots,x_{i-1})-\mathbb{E}(y_i^2)\big|.$$

Let M_3 be a bound on $\max_i(\mathbb{E}|x_i|^3+\mathbb{E}|y_i|^3)$. Suppose $h:\mathbb{R}^m\to\mathbb{R}$ is a thrice continuously differentiable function, and for $r=1,2,3$, let $L_r(h)$ be a finite constant such that $|\partial_i^r h(z)|\le L_r(h)$ for each i and z, where ∂_i^r denotes the r-fold derivative in the ith coordinate. Then

$$\big|\mathbb{E}h(x)-\mathbb{E}h(y)\big|\le\sum_{i=1}^m\left(R_iL_1(h)+\frac{1}{2}B_iL_2(h)\right)+\frac{1}{6}mL_3(h)M_3.$$

Let us denote $(\Psi_k)_{jj}$ as θ_{jk}, as in the previous section. With this notation, we have

$$e_{jt}=\sum_{k=0}^{\infty}\theta_{jk}\varepsilon_{j,t-k}.$$

Let

$$x_i=\begin{cases}\varepsilon_{j,T+1-i} & \text{for } i=1,\dots,2T,\\ \displaystyle\sum_{k=2T+1-i}^{\infty}\theta_{j,k+2T}\varepsilon_{j,T+1-i-k} & \text{for } i=2T+1,\dots,3T,\end{cases}$$

and $m=3T$. Then, for $t=1,\dots,T$, we have

$$e_{jt}=\sum_{k=0}^{T+t-1}\theta_{jk}x_{k+T-t+1}+x_{3T-t+1},$$

so that the R-dimensional vector $\mathbf{X}$ with the kth coordinate $e_{j\cdot}v_k$ can be thought of as a function $\mathbf{X}(x):\mathbb{R}^m\to\mathbb{R}^R$. Further, let

$$y_i=\begin{cases}\text{i.i.d. } N(0,1) & \text{for } i=1,\dots,2T,\\ 0 & \text{for } i=2T+1,\dots,3T.\end{cases}$$

Since x_i, $i=1,\dots,2T$, are independent, we have

$$R_i=B_i=0 \quad \text{for } i=1,\dots,2T. \tag{S28}$$

For $i>2T$, we have

$$\begin{aligned}R_i&=\mathbb{E}\big|\mathbb{E}(x_i|x_1,\dots,x_{i-1})\big|\\&\le\big(\mathbb{E}\big[\mathbb{E}\big(x_i^2|x_1,\dots,x_{i-1}\big)\big]\big)^{1/2}=\big[\mathbb{E}\big(x_i^2\big)\big]^{1/2}\end{aligned}$$

and

$$B_i=\mathbb{E}\big[\mathbb{E}\big(x_i^2|x_1,\dots,x_{i-1}\big)\big]=\mathbb{E}\big(x_i^2\big).$$

On the other hand, for $i>2T$,

$$\mathbb{E}\big(x_i^2\big)=\sum_{k=2T+1-i}^{\infty}\theta_{j,k+2T}^2.$$

By Assumption A2b, $\sum_{k=0}^{\infty}(1+k)|\theta_{jk}|\le B$. Therefore, $|\theta_{jk}|\le B/(1+k)$ and, for $2T<i\le 3T$,

$$\mathbb{E}\big(x_i^2\big)\le\frac{B}{4T+1-i}\sum_{k=2T+1-i}^{\infty}|\theta_{j,k+2T}|\le\frac{B^2}{(4T+1-i)^2}\le\frac{B^2}{T^2}.$$

Hence,

$$|R_i|\le B/T \quad\text{and}\quad |B_i|\le B^2/T^2 \quad \text{for } i=2T+1,\dots,3T. \tag{S29}$$

Further, for $i=1,\dots,2T$,

$$\begin{aligned}\mathbb{E}|x_i|^3+\mathbb{E}|y_i|^3&=\mathbb{E}|\varepsilon_{j,T+1-i}|^3+2\sqrt{2/\pi}<\big(\mathbb{E}|\varepsilon_{j,T+1-i}|^4\big)^{3/4}+2\\&\le(\varkappa_4+3)^{3/4}+2\le\varkappa_4+5.\end{aligned} \tag{S30}$$

Here, the second to the last inequality follows from Assumption A1. For $2T<i\le 3T$, we have

$$\begin{aligned}\mathbb{E}|x_i|^3+\mathbb{E}|y_i|^3=\mathbb{E}|x_i|^3&=\mathbb{E}\left|\sum_{k=2T+1-i}^{\infty}\theta_{j,k+2T}\varepsilon_{j,T+1-i-k}\right|^3\\&\le\left(\mathbb{E}\left(\sum_{k=2T+1-i}^{\infty}\theta_{j,k+2T}\varepsilon_{j,T+1-i-k}\right)^4\right)^{3/4}\\&\le\left(\left(\sum_{k=2T+1-i}^{\infty}\theta_{j,k+2T}^2\right)^2+\varkappa_4\sum_{k=2T+1-i}^{\infty}\theta_{j,k+2T}^4\right)^{3/4}\end{aligned}$$

$$\leq (1+\varkappa_4)^{3/4} B^3/T^3 \leq (1+\varkappa_4) B^3/T^3 \leq \varkappa_4 + 5$$

for all sufficiently large T.

Next, let $h(x) = g(\mathbf{X}(x))$. We have

$$\left|\partial_i^1 h(x)\right| \leq \sum_{t=1}^{R} \left|\partial_t^1 g(\mathbf{X})\right| \left|\partial_i^1 \mathbf{X}_t(x)\right| \leq M_g \sum_{t=1}^{R} \left|\partial_i^1 \mathbf{X}_t(x)\right|.$$

On the other hand,

$$\begin{aligned}\partial_i^1 \mathbf{X}_t(x) = \partial_i^1 (e_{j\cdot} v_t) &= \sum_{s=1}^{T} v_{ts} \partial_i^1 \left(\sum_{k=0}^{T+s-1} \theta_{jk} x_{k+T-s+1} + x_{3T-s+1} \right) \\ &= \begin{cases} \displaystyle\sum_{s=T+1-i}^{T} v_{ts} \theta_{j,i+s-T-1} & \text{for } i = 1, \dots, 2T, \\ v_{t,3T+1-i} & \text{for } i = 2T+1, \dots, 3T. \end{cases}\end{aligned}$$

Since $|v_{ts}| \leq \sqrt{2/T}$ and $\sum_{k=0}^{\infty} (1+k)|\theta_{jk}| \leq B$, we have

$$\left|\partial_i^1 \mathbf{X}_t(x)\right| \leq (B+1)\sqrt{2/T}.$$

Here, we use $B+1$ instead of B to take into account a possibility that $B < 1$. Combining the latter display with the above inequality for $|\partial_i^1 h(x)|$, we obtain

$$\left|\partial_i^1 h(x)\right| \leq M_g C_1 / T^{1/2}, \tag{S31}$$

where $C_1 = \sqrt{2} R(B+1)$.

Further,

$$\begin{aligned}\left|\partial_i^2 h(x)\right| &\leq \sum_{t_1,t_2=1}^{R} \left|\partial_{t_1 t_2}^2 g(\mathbf{X})\right| \left|\partial_i^1 \mathbf{X}_{t_1}(x)\right| \left|\partial_i^1 \mathbf{X}_{t_2}(x)\right| \\ &\quad + \sum_{t=1}^{R} \left|\partial_t^1 g(\mathbf{X})\right| \left|\partial_i^2 \mathbf{X}_t(x)\right| \\ &= \sum_{t_1,t_2=1}^{R} \left|\partial_{t_1 t_2}^2 g(\mathbf{X})\right| \left|\partial_i^1 \mathbf{X}_{t_1}(x)\right| \left|\partial_i^1 \mathbf{X}_{t_2}(x)\right| \\ &\leq M_g \sum_{t_1,t_2=1}^{R} \left|\partial_i^1 \mathbf{X}_{t_1}(x)\right| \left|\partial_i^1 \mathbf{X}_{t_2}(x)\right|,\end{aligned}$$

so that

$$\left|\partial_i^2 h(x)\right| \leq M_g C_2 / T, \tag{S32}$$

where $C_2 = C_1^2$. Similarly,

$$\left|\partial_i^3 h(x)\right| \leq \sum_{t_1,t_2,t_3=1}^{R} \left|\partial_{t_1 t_2 t_3}^3 g(\mathbf{X})\right| \left|\partial_i^1 \mathbf{X}_{t_1}(x)\right| \left|\partial_i^1 \mathbf{X}_{t_2}(x)\right| \left|\partial_i^1 \mathbf{X}_{t_3}(x)\right|,$$

so that

$$\left|\partial_i^3 h(x)\right| \le M_g C_3 / T^{3/2}, \tag{S33}$$

where $C_3 = C_1^3$.

Using inequalities established above in Theorem S5, we obtain

$$\left|\mathbb{E}h(x) - \mathbb{E}h(y)\right| \le \frac{M_g C_1 B}{T^{1/2}} + \frac{M_g C_2 B^2}{2T^2} + \frac{M_g C_3(\varkappa_4 + 5)}{2T^{1/2}}.$$

On the other hand,

$$\begin{aligned}\left|\mathbb{E}g(\mathbf{X}) - \mathbb{E}g(\mathbf{Y})\right| &= \left|\mathbb{E}g\big(\mathbf{X}(x)\big) - \mathbb{E}g\big(\mathbf{X}(y)\big) + \mathbb{E}g\big(\mathbf{X}(y)\big) - \mathbb{E}g(\mathbf{Y})\right| \\ &= \left|\mathbb{E}h(x) - \mathbb{E}h(y) + \mathbb{E}g\big(\mathbf{X}(y)\big) - \mathbb{E}g(\mathbf{Y})\right| \\ &\le \left|\mathbb{E}h(x) - \mathbb{E}h(y)\right| + \left|\mathbb{E}g\big(\mathbf{X}(y)\big) - \mathbb{E}g(\mathbf{Y})\right|.\end{aligned}$$

Note that $\mathbf{X}(y)$ and $\mathbf{Y}$ are normally distributed vectors with zero means but different covariance matrices, which we denote Σ_y and $\Sigma_{\mathbf{Y}}$, respectively. By definition,

$$\Sigma_{\mathbf{Y}} = 2\pi f_j(0) I_R.$$

Define η_{T+1-i}, $i = 1, 2, \dots$ as y_i for $i = 1, \dots, 2T$, and as i.i.d. $N(0, 1)$ random variables independent from $y_1, \dots, y_{2T}$ for $i > 2T$. Then,

$$\begin{aligned}\mathbf{X}_t(y) &= \sum_{s=1}^{T} \sum_{k=0}^{T+s-1} \theta_{jk} y_{k+T-s+1} v_{ts} \\ &= \sum_{s=1}^{T} \sum_{k=0}^{T+s-1} \theta_{jk} \eta_{s-k} v_{ts} \\ &= \sum_{s=1}^{T} \sum_{k=0}^{\infty} \theta_{jk} \eta_{s-k} v_{ts} - \sum_{s=1}^{T} \sum_{k=T+s}^{\infty} \theta_{jk} \eta_{s-k} v_{ts} \\ &= \mathbf{X}_t(\eta) - \sum_{s=1}^{T} \sum_{k=T+s}^{\infty} \theta_{jk} \eta_{s-k} v_{ts} \\ &= \mathbf{X}_t(\eta) - \sum_{\tau=T}^{\infty} \eta_{-\tau} \sum_{s=1}^{T} \theta_{j,\tau+s} v_{ts}.\end{aligned}$$

By Lemma S2,

$$\left|\mathbb{E}\big(\mathbf{X}_{t_1}(\eta)\mathbf{X}_{t_2}(\eta)\big) - 2\pi f_j(\varpi_{t_1}/2)\delta_{t_1 t_2}\right| \le CB^2/T.$$

Further,

$$\begin{aligned}\mathbb{E}\left(\prod_{i=1}^{2} \sum_{\tau=T}^{\infty} \eta_{-\tau} \sum_{s=1}^{T} \theta_{j,\tau+s} v_{t_i s}\right) &= \sum_{\tau=T}^{\infty} \left(\prod_{i=1}^{2} \sum_{s=1}^{T} \theta_{j,\tau+s} v_{t_i s}\right) \\ &\le \sum_{\tau=T}^{\infty} \sum_{s=1}^{T} \theta_{j,\tau+s}^2 \le \sum_{s=1}^{T} \frac{B^2}{(T+s)^2} < \frac{B^2}{T},\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\left(\mathbf{X}_{t_1}(\eta)\sum_{\tau=T}^{\infty}\eta_{-\tau}\sum_{s=1}^{T}\theta_{j,\tau+s}v_{t_2s}\right)\\
&\quad=\mathbb{E}\left(\sum_{s_1=1}^{T}\sum_{\tau_1=-s_1}^{\infty}\eta_{-\tau_1}\theta_{j,\tau_1+s_1}v_{t_1s_1}\sum_{\tau_2=T}^{\infty}\eta_{-\tau_2}\sum_{s_2=1}^{T}\theta_{j,\tau_2+s_2}v_{t_2s_2}\right)\\
&\quad=\sum_{\tau=T}^{\infty}\sum_{s_1=1}^{T}\theta_{j,\tau+s_1}v_{t_1s_1}\sum_{s_2=1}^{T}\theta_{j,\tau+s_2}v_{t_2s_2}<\frac{B^2}{T}.
\end{aligned}$$

This yields

$$\left|\mathbb{E}\big(\mathbf{X}_{t_1}(y)\mathbf{X}_{t_2}(y)\big)-2\pi f_j(\varpi_{t_1}/2)\delta_{t_1t_2}\right|\le\tilde{C}/T$$

for some constant $\tilde{C}$ that depends on B. This and inequality (S13) yield

$$\left|\mathbb{E}\big(\mathbf{X}_{t_1}(y)\mathbf{X}_{t_2}(y)\big)-2\pi f_j(0)\delta_{t_1t_2}\right|\le\hat{C}/T \tag{S34}$$

for any positive integers $t_1, t_2\le R$, where $\hat{C}$ depends on B and R.

Further,

$$\left|g\big(\mathbf{X}(y)\big)-g(\mathbf{Y})\right|\le M_g\left\|\mathbf{X}(y)-\mathbf{Y}\right\|.$$

Therefore,

$$\begin{aligned}
\left|\mathbb{E}g\big(\mathbf{X}(y)\big)-\mathbb{E}g(\mathbf{Y})\right|&\le M_g\mathbb{E}\left\|\mathbf{X}(y)-\mathbf{Y}\right\|\\
&\le M_g\big(\mathbb{E}\left\|\mathbf{X}(y)-\mathbf{Y}\right\|^2\big)^{1/2}.
\end{aligned}$$

On the other hand, we may assume that y and $\mathbf{Y}$ are independent, and thus

$$\mathbb{E}\left\|\mathbf{X}(y)-\mathbf{Y}\right\|^2=\operatorname{tr}(\Sigma_y-\Sigma_{\mathbf{Y}})\le R\hat{C}/T,$$

where the last inequality follows from (S34). Hence, finally,

$$\begin{aligned}
\left|\mathbb{E}g(\mathbf{X})-\mathbb{E}g(\mathbf{Y})\right|\le{}&\frac{M_gC_1B}{T^{1/2}}+\frac{M_gC_2B^2}{2T^2}\\
&+\frac{M_gC_3(\varkappa_4+5)}{2T^{1/2}}+\frac{M_gR^{1/2}\hat{C}^{1/2}}{T^{1/2}}.
\end{aligned}$$

This yields the statement of Lemma S3.

S4. THE "NUMBER OF FACTORS"

S4.1. *Proof of Proposition OW4*

Let us denote $V(k) + k\hat{\sigma}^2 p_j(N,T)$ as $\text{IPC}_j(k)$. For positive integers k, $\text{IPC}_j(k) > \text{IPC}_j(k-1)$ if and only if $\hat{\lambda}_k/T < \hat{\sigma}^2 p_j(N,T)$. The latter inequality is equivalent to

$$\hat{\lambda}_k / \operatorname{tr}\hat{\Sigma} < \left(1 - \sum_{j=1}^{k_{\max}} \hat{\lambda}_j / \operatorname{tr}\hat{\Sigma}\right) p_j(N,T). \tag{S35}$$

On the other hand, by Theorem OW1(iii), for any fixed positive integer k, $\hat{\lambda}_k/\operatorname{tr}\hat{\Sigma} \xrightarrow{\text{P}} 6/(k\pi)^2$. Since $p_j(N,T) \to \infty$ as $N, T \to \infty$ for $j = 1, 2, 3$, inequality (S35) is satisfied with probability arbitrarily close to 1 for all sufficiently large N, T. This yields statement (i) of Proposition OW4.

To establish part (ii), we need the following lemma.

LEMMA S6: *Under assumptions of Proposition OW4, for $k_{\max} = [\gamma\delta_{NT}]$, we have*

$$\hat{\sigma}^2 = O_{\text{P}}\left(\frac{T \operatorname{tr}\Omega}{N\delta_{NT}} + \frac{(T + N_\varepsilon)N^{2\alpha}\delta_{NT}}{NT}\right).$$

PROOF: We rely on notations and definitions from the proof of Theorem OW1. Let $\tilde{V}(k) = \operatorname{tr}\tilde{\Sigma}/T - \sum_{j=1}^{k} \tilde{\lambda}_j/T$. Then,

$$T\tilde{V}(k_{\max}) = \sum_{r=k_{\max}+1}^{T-1} \tilde{\lambda}_r \le \sum_{r=k_{\max}+1}^{T-1} w_r'\tilde{\Sigma}w_r = \sum_{r=k_{\max}+1}^{T-1} \frac{1}{N}\sigma_r^2 v_r'\varepsilon' W\varepsilon v_r.$$

Denote $\sum_{r=k_{\max}+1}^{T-1} \sigma_r^2$ as $s_{k_{\max}}$. Then Lemma OW6 and the fact that $\operatorname{tr} W = \operatorname{tr}\Omega$ yield

$$\sum_{r=k_{\max}+1}^{T-1} \sigma_r^2 v_r'\varepsilon' W\varepsilon v_r = s_{k_{\max}} \operatorname{tr}\Omega + o_{\text{P}}(s_{k_{\max}} \operatorname{tr}\Omega).$$

Since $\sigma_r = (2\sin(\pi r/(2T)))^{-1} \le T/(2r)$ for $r = 1, \dots, T-1$, we have

$$s_{k_{\max}} \le \sum_{r=k_{\max}+1}^{T-1} T^2/\left(4r^2\right) \le T^2/(4k_{\max}),$$

and therefore,

$$\tilde{V}(k_{\max}) \le \frac{T \operatorname{tr}\Omega}{4Nk_{\max}} + o_{\text{P}}\left(\frac{T\operatorname{tr}\Omega}{Nk_{\max}}\right). \tag{S36}$$

Next, similarly to (OW34), we have the following inequality:

$$\left|\left(TV(k_{\max})\right)^{1/2} - \left(T\tilde{V}(k_{\max})\right)^{1/2}\right| \le \left\|\Psi^*(L)\varepsilon M\right\| \min\{1, \sqrt{T/N}\}.$$

Hence,

$$\hat{\sigma}^2 = V(k_{\max}) \le 2\tilde{V}(k_{\max}) + \frac{2}{T}\left\|\Psi^*(L)\varepsilon M\right\|^2 \min\{1, T/N\}.$$

This inequality together with (OW33) and (S36) yield the statement of the lemma. *Q.E.D.*

By Theorem OW1(ii), for any fixed k, $(\hat{\lambda}_k/T)^{-1} = O_{\mathrm{P}}(N/(T\,\mathrm{tr}\,\Omega))$. Therefore, by Lemma S6, $(\hat{\lambda}_k/T)^{-1}\hat{\sigma}^2 = O_{\mathrm{P}}(m_{NT})$ and

$$(\hat{\lambda}_k/T)^{-1}\hat{\sigma}^2 p_j(N,T) = O_{\mathrm{P}}\big(m_{NT}\,p_j(N,T)\big) \xrightarrow{\mathrm{P}} 0.$$

Hence, for any fixed k, $\hat{\lambda}_k/T > \hat{\sigma}^2 p_j(N,T)$ with probability arbitrarily close to 1 for all sufficiently large N, T. This implies that $\mathrm{IPC}_j(k) < \mathrm{IPC}_j(k-1)$ with probability arbitrarily close to 1 for all sufficiently large N, T, and thus, $\hat{k}_j \xrightarrow{\mathrm{P}} \infty$.

REFERENCES

BRILLINGER, D. R. (2001): *Time Series: Data Analysis and Theory*. SIAM. [8,15]

CHATTERJEE, S. (2006): "A Generalization of the Lindeberg Principle," *The Annals of Probability*, 34, 2061–2076. [18]

HANNAN, E. J. (1970): *Multiple Time Series*. John Wiley & Sons, Inc. [9]

KATO, T. (1980): *Perturbation Theory for Linear Operators*. Springer-Verlag. [5]

ONATSKI, A. (2018): "Asymptotics of the Principal Components Estimator of Large Factor Models With Weak Factors and i.i.d. Gaussian Noise," Cambridge Working Paper Economics 1808, University of Cambridge. [6]

ONATSKI, A., AND C. WANG (2021): "Spurious Factor Analysis," *Econometrica*, 89. [1]

SARGAN, J. D., AND A. BHARGAVA (1983): "Testing Residuals From Least Squares Regression for Being Generated by the Gaussian Random Walk," *Econometrica*, 51, 153–174. [1]

SHORACK, G. R., AND J. A. WELLNER (1986): *Empirical Processes With Applications to Statistics*. New York: Wiley. [2]

Co-editor Ulrich K. Müller handled this manuscript.

Manuscript received 26 September, 2018; final version accepted 22 June, 2020; available online 15 July, 2020.